inHarmony: A Digital Twin For Emotional Well-being

by

Amani Albraikan

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the Ph.D. degree in
Computer Science

School of Electrical Engineering and Computer Science
Faculty of Computer Science
University of Ottawa

Abstract

A digital twin is the enabling technology that facilitates monitoring, understanding, and providing continuous feedback to improve quality of life and well-being. Thus, a digital twin can consider a solution to enhance one's mood to improve the quality of life and emotional well-being. However, there remains a long road ahead until we reach digital twin systems that are capable of empowering development and the deployment of digital twins. This is because there are so many elements and components that can guide the design of a digital twin.

This thesis provides a general discussion for the central element of an emotional digital twin, including emotion detection, emotional biofeedback, and emotion-aware recommender systems. In the first part of this thesis, we propose and study the emotion detection models and algorithms. For emotions, which are known to be highly user dependent, improvements to the emotion learning algorithm can significantly boost its predictive power. We aimed to improve the accuracy of the classifier using peripheral physiological signals. Here, we present a hybrid sensor fusion approach based on a stacking model that allows for data from multiple sensors and emotion models to be jointly embedded within a user-independent model.

In the second part of this thesis, we propose a real-time mobile biofeedback system that uses wearable sensors to depict five basic emotions and provides the user with emotional feedback. These systems apply the concept of Live Biofeedback through the introduction of an emotion-aware digital twin. An essential element in these systems guides users through an emotion-regulation routine. The proposed systems are aimed at increasing self-awareness by using visual feedback and provide insight into the future design of digital twins. We focus on workplace environments, and the recommenda-

tions are based on human emotions and the regulation of emotion in the construct of emotional intelligence. The objective is to suggest coping techniques to a user during an emotional, stressful episode based on her or his preferences, history of what worked well and appropriateness for the context.

The developed solution has been studied based on usability studies and extensively compared to related works. The obtained results show the potentials use as an emotional digital twin. In turn, the proposed solution has been providing significant insights that will guide future developments of digital twins using several scenarios and settings.

Acknowledgment

I would like to express my gratitude to my principal supervisor, Prof. Abdulmotaleb El Saddik (University of Ottawa, Ottawa, Canada) for his careful guidance, tremendous patience and eagerness to provide support in the face of uncertainty at any stage of my work on this thesis.

I would like to thank Dr. Diana P. Tob'on for their generous feedback, valuable corrections, provision, and all the invaluable experience and knowledge gained while working with them. I would also like to thank Dr. Hussein Al Osman for their advice on the physiological sensor. To Dr. Haiwei Dong for giving me an example to follow. To all the staff at the Institute of the University of Ottawa and especially to Annik Dion for all her care and attention. Special thanks to Dr. Daniel Amyot for all their time, energy and support. To Dr. Honaida Shahbar for all their sociology knowledge and life coaching experience. To all the students of MCRLab and Discover lab members for all their mistakes and successes that made me grow personally and professionally.

I want to give a very special thanks to (Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia) for sponsoring me through my journey of graduate school. Thanks for all your support, continued generosity, financial contribution, and support in making my college career possible. I hope to give back to the university someday and help a student reach their educational goals as a teacher.

I want to say to my parents that my heart is full of love, respect and sincere appreciation for all you have given me, for helping me to become the person I am, and for always being there to give me a hand. My gratitude also goes to all the members of my family and friends for their support, love and warmth, in particular to my sisters for sharing all my sorrows and joys and for their emotional support.

I would like to express my deepest thanks to my husband, Abdulaziz Alwasel, for all the sacrifices that he had made to support me, for helping me to grow, for sharing my dreams and ideas, and for his love and warmth. This accomplishment would not be possible without him beside me. I could not have imagined having better support during my Ph.D. study. I also would like to express my sincere gratitude to my children, Hamed, Naya, and Rakan, for their understanding and the sparkling brightness of their smiles. Thank you for being part of my life.

Abbreviations

AI	Artificial Intelligence
AR/VR	Virtual reality/augmented reality
DFA	Discriminant Function Analysis
DTW	Dynamic time warping
ECG	Electrocardiography
EDA	Electrodermal Activity
EI	Emotional intelligence
EMDB	Emotional Movie Database
EMG	Electromyographic
ER	Emotion regulation
FLBF	Foreign-Live Biofeedback
GP	Gaussian Process
GSR	Galvanic Skin Response
HCI	Human-computer interaction
HRV	Heart Rate Variability
IAPS	International Affective Picture System
IoT	Internet of Things
IRS	Individual response specificity
K-NN	K-Nearest Neighbor
LBF	Live biofeedback
LDA	Linear Discriminant Analysis
PAD	Pleasure, arousal and dominance
PCA	Principal component analysis

PNS	Peripheral nervous system
Resp.	Respiration
RSs	Recommender systems
SAM	Self-Assessment Manikin
SDNN	Standard Deviation of Normal-to-Normal intervals
SLBF	Self-Live Biofeedback
SMBO	Sequential model-based optimization
SRS	Stimulus-response specificity
SVM	Support Vector Machine
VLF	Vary low frequency of Heart Rate Variability
WHO	World Health Organization
WMD-DTW	Weighted multi-dimensional DTW

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The World Health Organization (WHO) defines health as follows: "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." More recently, a dynamic definition of well-being developed by Dodge et al. highlights that well-being is the balance point between a person's resources and skills and the challenges experienced in life [4]. When people have more challenges than resources, the well-being see-saw dips. People who experience remarkable well-being have a proper balance that enables them to overcome the different forms of challenges while experiencing a measure of happiness, positive emotions, and life satisfaction.

Mental health is a fundamental part of health and well-being. Mental health, like other aspects of health, includes the ability to manage thoughts and emotions and the ability to cope with the normal stresses of life as well as to build social relationships. It also includes the aptitude to learn and to acquire an education, ultimately enabling an individual's full active contribution to society. The WHO, in 2013-2020 mental health action plan, recognizes the essential role of mental health. WHO recommends the development of tools or strategies for self-help and care for persons with mental disorders, including the use of electronic and mobile technologies. The comprehensive WHO action plan stresses the need for early intervention through the identification, prevention and treatment of emotional or behavioral problems, provision of healthy living and working conditions such as work organizational improvements and publicly available evidence-based stress management schemes.

Emotional well-being is a critical component that governs positive mental health. Emotional well-being denotes the ability to achieve a measure of satisfaction with life because of the positive balance of pleasant and unpleasant affect [5]. Individuals experience distress when there is an improper balance between pleasant and unpleasant affect. Stress and fatigue are among the leading triggers that contribute to the development of distress. On the other hand, the failure to deal with challenges that an individual faces each day often leads to distress [6]. In this context, challenges denote various situations that require a response from the individual daily. Individuals need specific resources and skills that can enable them to thrive and overcome daily challenges. One of the most important resources and skills is emotional intelligence (EI). Notably, EI is of critical importance because it helps an individual to register a significant understanding of personal emotions and to extend a level of understanding to the emotions of others [7].

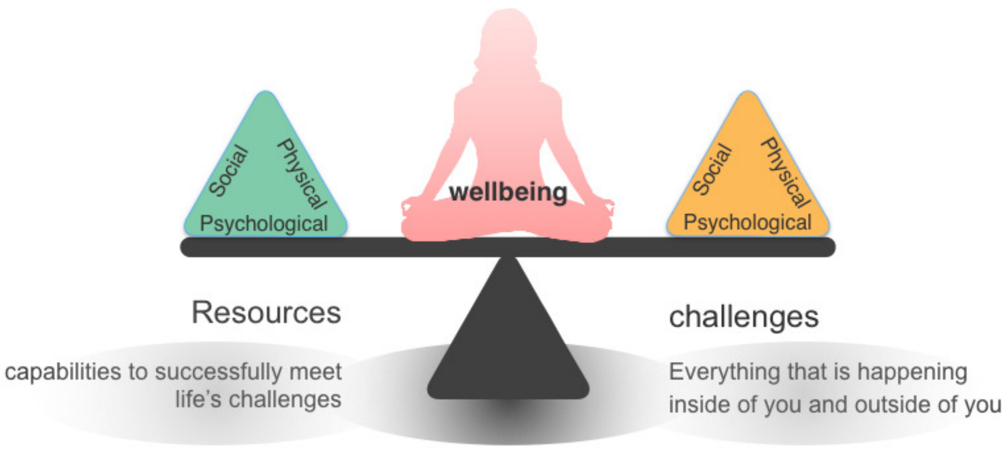


Figure 1.1: The definition of wellbeing (Dodge 2012)

In this content, digital twins have been attracting increasing attention toward improving quality of life and well-being. The use of a digital twin, enabled by EI, has the potential to improve emotional well-being in two ways. First, by recognizing a feeling as it develops, one can properly handle that feeling. Second, identifying others' feelings based on having experienced similar feelings enables an individual to manage emotions in others or at least to show empathy [8]. EI focuses on how individuals identify, understand, express, regulate, and use their own feelings and those of others [9]. In [8], Goleman addressed five domains of EI: knowing one's own emotions, managing emotions, motivating oneself, recognizing emotions in others, and engaging in relationships. He considers self-awareness as a "keystone" of EI. Self-awareness is the foundation of enhancing positive emotions and improving quality of life. From an EI perspective, self-awareness is the first step in exerting emotional self-control for containing, ordering or controlling emotions. The second main element in machine emotional intelligent is manage emotions. Recommender systems (RSs) are a growing area of research. Recently, considerable effort has been made to use emotions in the recommendation process as a way to customize content [10].

The study of emotion and how it influences human decision making and satisfaction focuses on the analysis of data, especially location-based tracking data, social media data, and crowd-sourced geographic information. The role played by EI and emotion recognition (ER) is rarely considered. ER and EI each capture an essential aspect of emotion management. ER focuses on basic emotion regulation processes. EI assesses the consequences of individual differences in ER on social, health, educational, and occupational outcomes. Intelligent ER has a significant influence on human decision making by using emotional information to guide thinking and behavior and to manage and/or adjust emotions to adapt to environments or to achieve one's goals [11]. These skills enable individuals to choose adaptive behaviors as a response. The inclination to attend to one's emotions and the ability to identify one's own emotions are likely to be crucial for adaptively using emotional information and have significant implications for individuals' satisfaction.

1.1 Problem Statement

Research problem: Every day, people are continually exposed to a wide range of potentially arousing stimuli. To seek body-mind harmony, people must engage in some form of ER almost continuously. We would like to explore various methods for assisting individuals to achieve a higher level of emotional well-being.

In the digital world, the digital twin assists in early intervention modes to identify emotional coping strategies and to provide avenues for new interventions through RSs. All the methods explored in this work fall under the general concepts of digital twins and live biofeedback (LBF), and we will investigate various directions stemming from the latter notion. Nonetheless, the general goal of achieving a higher level of emotional well-being remains the underlying theme for all studied methods.

1.2 Motivation

EI empowers an individual to exhibit both personal and social competencies that help in the understanding and regulation of personal emotions. Similarly, personal and social competencies also empower the individual to understand the emotions of others [12]. EI registers positive outcomes in fostering emotional well-being. The digital twin can utilize EI to improve the quality of life and well-being. The balance of positive and negative emotions contributes to judgments of life satisfaction. Positive emotions trigger upward spirals towards enhanced emotional well-being. Individuals use their EI and ER to reinforce emotion states to meet particular goals:

Individual satisfaction: User emotion plays a critical role in the decision process that leads to individual satisfaction. Emotion expiration appears to be a primary human motive [13] and can be considered from a functionalism perspective [11]. EI fosters satisfaction because it helps the individual deal with various situations successfully. According to Thiruchelvi and Supriya, EI enhances an individual's perspective towards challenges and makes it easier for the person to deal with challenges while still experiencing a measure of satisfaction with life [14].

Self-awareness: Most people can regulate their own and others' feelings [15]. However, a common infirmity affects people who cannot recognize emotion in themselves and who are, therefore, unable to experience a life that fulfils them emotionally. For example, individuals reporting greater emotional clarity and a greater ability to repair their emotional states report higher levels of self-esteem and mental health [16]. By contrast, individuals who have lower levels of emotional clarity and individuals who are unable to regulate their emotional states show poorer emotional adjustment and exhibit patterns of response characterized by a mismatch between their goals, responses, and/or modes of expression [11].

Decision Making: EI is a key determinant of better decision making. People with high levels of EI can control their emotions and prevent such emotions from adversely affecting their power of judgment [17]. EI helps individuals in separating the decision-making process from unrelated emotions that may lead to negative outcomes.

Resilience: EI leads to higher levels of resilience. Resilience is the ability to keep going irrespective of challenges in daily life [18]. For this reason, emotionally intelligent people can control their emotions without allowing negative emotions to overload their thinking and behavior. As a result, they become highly resilient. Resilience serves as a prerequisite for mental health because it helps an individual to adapt to adversity irrespective of the risks involved.

Relationships: EI fosters the development of better relationships. Specifically, people with remarkable levels of EI are likely to establish and maintain successful relationships because they can handle different situations [19]. Moreover, emotionally intelligent people are in a position to understand the existing connection between people's emotions and actions, a factor that enables them to establish relationships with the right people.

Communication: EI has a positive impact on communication. Specifically, when individuals manage their emotions successfully, they can communicate successfully [20]. On the other hand, EI helps an individual to react to various situations with a greater understanding of how the communication used will determine the effectiveness of the message delivered.

1.3 Aim and Contribution

In this thesis, we formalize emotion-aware systems through the notions of digital twins and LBF. Moreover, the following contributions are made by this thesis:

Digital twin for emotional well-being: The development of the system and its main components are based on the live biofeedback model presented in this thesis. These methods all have a similar scope and goal in relation to the problem of enabling a higher level of emotional well-being.

An ER algorithm: This contribution concerns the development of an ER algorithm using physiological signals. We defined customizable procedures to limit the divergent dynamics of emotions across individuals. Our aim is to improve the accuracy rate of the classifier using peripheral physiological signals. Here, we present a hybrid sensor fusion approach based on a stacking model that allows for data from multiple sensors and emotion models to be jointly embedded within a user-independent model. We applied a meta-learning approach to the dataset and showed that the ensemble approach outperforms any individual method.

Emotion-biofeedback system: These are methods that we have devised to enable the development of LBF systems.

Opportune moment detection algorithm: These are the methods used to provide the dynamic intervention system and provide personalized recommendations for coping mechanisms.

inHarmony usability study: The usability study provides insight into the perceived usefulness and effectiveness of the digital twin system in the workplace.

Emotion data-set: We collected sufficient physiological data from 18 subjects in five induced emotions (neutral, happy, sad, erotic and horror). We also obtained hundreds of records of physiological signals for each emotion.

1.4 Scholarly Achievements

In the process of completing this work, the following publications have been submitted, accepted or published:

1. Albraikan, A., Tobón, D. P., El Saddik, A. (2018, March). *"Hyper-Parameter Optimization for Emotion Detection using Physiological Signals"*. In 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops) (pp. 836-841). IEEE.

2. Albraikan, A., Tobón, D. P., El Saddik, A. (2018). *"Toward User-Independent Emotion Recognition using Physiological Signals"*. in IEEE Sensors Journal.

3. Albraikan, A., Hafidh, B., El Saddik, A. (2018). *"iAware: A Real-Time Emotional Biofeedback System Based on Physiological Signals"*. in IEEE Access, vol. 6, pp. 78780-78789, 2018.

4. Albraikan, A., Badawi, H., Hamam, A., El Saddik, A. (2013, July). *"Haptibasic: Learning basic concepts of a haptic technology through edutainment games"*. In Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on (pp. 1-4). IEEE.

1.5 Thesis Organization

The remainder of this thesis is organized as follows.

Chapter 2 introduces the basic definition of emotion in psychology and the main categories of emotions. In addition, this chapter presents a two-part literature review.

Chapter 3 discusses our digital twin model along with its various components.

Chapter 4 explains the first layer of the system, including sensing data and a description of the experimental setting.

Chapter 5 presents the second layer of the system, including the design of the proposed emotion estimation system and the live biofeedback.

Chapter 6 presents the third layer of the system and describes the dynamic intervention system, including the opportune moment detection.

Chapter 7 presents the inHarmony case study: a digital twin for emotional well-being.

Chapter 8 provides the results of the experiments and a discussion on the predicted emotions and biofeedback.

Chapter 9 discusses the usability study that was conducted to evaluate the usability and effectiveness of the proposed framework.

Chapter 10 summarizes the work conducted in this research, provides a conclusion, and proposes future research to enhance the present work.

Chapter 2

Background & Related work

No agreement has been reached on the definition of an emotion [21]. We are not attempting to reprove or evaluate emotion theories or to propose a new definition; rather, our objective is to study the new generation of RSs and the effects of emotion in a digital twin. This chapter aims to determine the correct approach to develop an emotion-aware or emotion-representing system; thus, we are interested in knowing, e.g., the best emotion model to use, the different types of emotion, the relationship between stimulus and emotion, and the best type of measurement to use.

In addition, this chapter provides a literature review of two main topics. The first part of the review explores the m-health application approaches used in mental well-being. The second part presents emotion recognition using physiological signal approaches and algorithms in general. These review steps are inspired by the systematic literature review guidelines of Kitchenham (2004)[22]. Moreover, the initial review results were refreshed in 2019.

2.1 Emotion Theories

When comparing the roles of emotional stimuli and perception, different influential psychological theories provide clues about the mechanisms through which emotions serve to produce adaptive responses. The main approaches are basic theories, appraisal theories, and constructivist theories of emotion. Basic emotion theories posit that all feelings can be derived from a limited set of universal and innate basic emotions showing distinct patterns of psychophysiology that are cross-cultural and user-defined. This phenomenon is referred to as stimulus-response specificity (SRS). This view argues that different emotions are associated with unique patterns of physiological changes [23]. Conversely, appraisal theories suggest a more flexible and dynamic mechanism that takes into account the interactions of stimuli and the needs and goals of the observer (top-down processes). Constructivist theories have expanded upon appraisal theories, focusing on the constraining bottom-up effects of mental representations and available concepts on the final emotion decision [24]. The latter two theories consider individual response specificity (IRS).

2.1.1 Cognitive Motivation Emotion Model

In appraisal theory, emotions result from an individual's meaning analysis of the implications of his/her circumstances, and individual differences in emotion arise when individuals appraise similar situations differently. The cognitive motivation emotion model is an appraisal attempt to describe the situational and dispositional antecedents of an individual's circumstances [25].

Acording to Smith (1990), the emotions, in a structural appraisal model, are the product of realtionships between persons and environment [26]. In Figure 2.1, the theoretical model of the cognitive motivational emotive (CME) system shows how the personality and situational variables interact with each other. The correlation between the two loops, the appraisal and these responses, appears to have two distinct levels of organization. The appraisal is itself a continuing component of emotional response [26].

2.2 Emotion Models

The classification of emotion has been studied based mainly on two fundamental viewpoints: SRS and IRS.

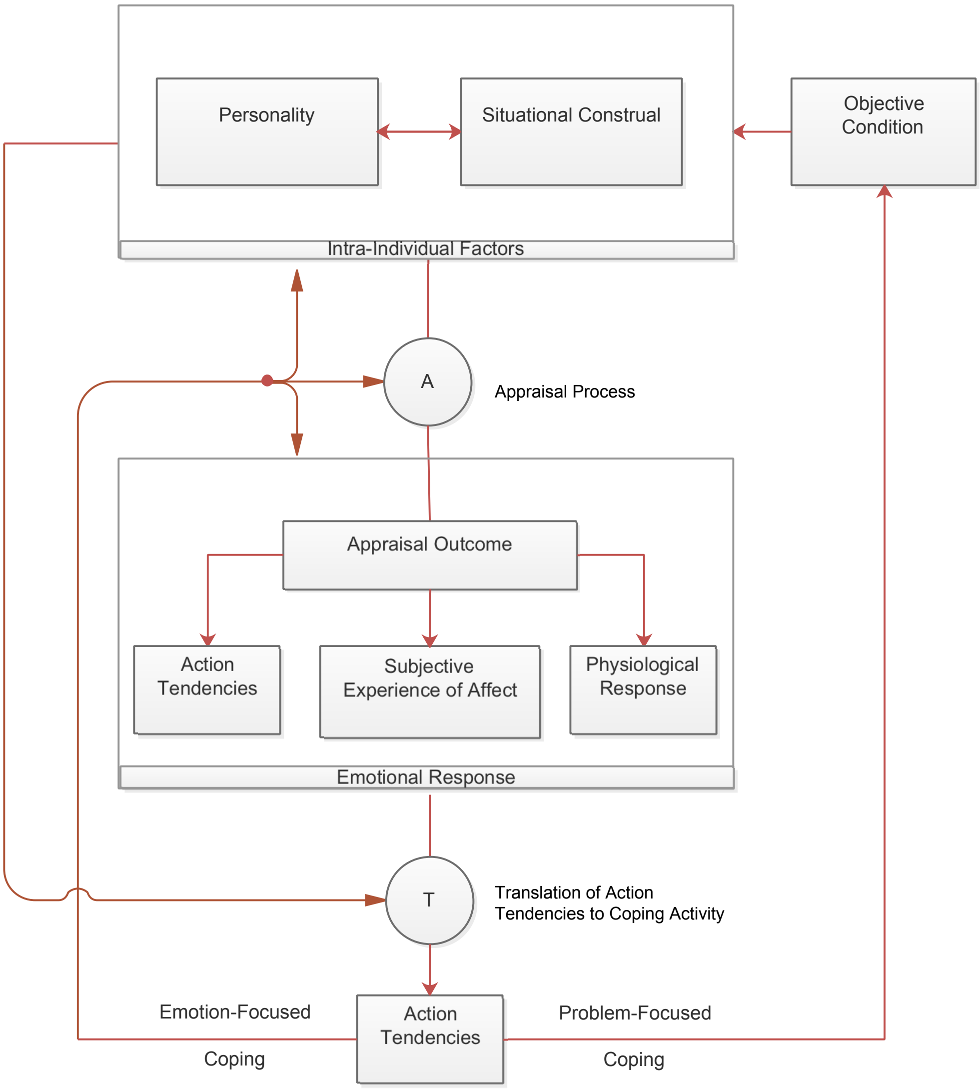


Figure 2.1: The model of the cognitive motivation emotion system [26].

2.2.1 Stimulus Response Specificity

SRS is a discrete autonomic circuit. All components and responses must be aligned with each other and focused on emotional responses that are typically beyond an individual's control [23]. The autonomic measurements include facial expressions (e.g., smiling) and physiological responses (e.g., sweating) caused by changes in the central nervous system or peripheral nervous system (PNS). However, emotions are a complex state of feeling for human beings, and the physiological signals related to these emotions may vary from one person to another. Individual differences represent one of the main challenges in emotion detection. Individual differences in emotional thresholds lead to individual differences in patterns of behavior [27]. The main reasons for individual differences are the dynamic nature of emotions and differences in the perception of a stimulus. The componential architecture depends on the person, the situation, and the time [28].

2.2.2 Individual Response Specificity

Alternatively, the IRS model was proposed by James A. Russell in 1980 [29]. In the dimensional approach, emotions can be described on the basis of three dimensions: pleasure, arousal and dominance (PAD). In 1974, Mehrabian and Russell [30] established a scale to measure the combination of these dimensions. Self-reports focus on emotions experienced in response to a given stimulus by describing the orthogonal continuums from pleasure to displeasure, activation to deactivation, and controlling to dominant. Self-report measures are a necessary tool for evaluating subjective feelings, but self-report methods are vulnerable to validity and reliability issues. Although most authors report that their emotion scales are sufficiently reliable, rating scales are one of the main reliability limitations. The most critical limitation involves response bias; for example, respondents may be unwilling to report their emotions because of social desirability bias, especially for sensitive topics (e.g., erotica and charity) [31]. Respondents may also be unable to report their emotions because they are not aware of exactly how they feel. Furthermore, people may overestimate the negativity of their reactions when they are anticipating negative events [32]. The quality of the self-report will affect the accuracy of the results. Introspective ability may lead to a consistent response to the experiment [33]. Both SRS and IRS can influence physiological response patterns [34]; thus, a physiological response can be viewed as the sum of IRS and SRS [35].

2.3 Emotion Induction Method

A critical step in identifying the emotion of a subject is eliciting emotional responses. The International Affective Picture System (IAPS) is a widely used system for emotion elicitation [36] that has previously been adopted for many psychophysiological studies on emotion elicitation. However, images are not sufficient for evoking powerful emotions [37]. Another approach employs audio, visual, and cognitive stimuli to evoke a particular emotion [38]. This multimodal approach uses controlled illumination as the visual stimulus, background music as the auditory stimulus, and a storyteller to evoke emotion. However, this method was designed for testing children between seven and eight years old. The Emotional Movie Database (EMDB) is a film clip system for emotion elicitation [39] that contains 52 film clips without any sound effects, based on emotional stimuli (i.e., valence, arousal, and dominance). These film clips are divided into five categories, namely, erotic, horror, socially negative, socially positive, and neutral. For the stimulus material herein, we selected the EMDB because videos contain more emotional content than do single images [37]. However, using visual stimulation videos without sound is not sufficient for adequate emotion elicitation [38]. Hence, we added audio to the set of film clips from the EMDB. First, we searched for clips on the internet and then added the sound from the internet sources to the clips from the database. We decided to use movie clip number 1003 for the neutral condition, 2005 for the happy state, 3005 for the sad condition, 4000 for the erotic situation, and 5000 for the horror condition. Relaxing music with a smooth rhythm was played for 2 min between each emotion elicitation to ensure the stabilization of emotions, as previously proposed [38]. Since the discrete model associates unique patterns of physiological changes with emotion, we used the model of Ekman [40] and utilized autonomic measures to map reactions to emotional stimuli.

To determine the correct approach to the development of an emotion-aware or emotion-representing system, we first need to understand how stimulus is perceived based on emotion theories. Basic emotion theories argue that emotion is discrete, as it is an autonomic circuit, and all components and the response must be aligned with each other. Appraisal theories suggest a more flexible and dynamic mechanism that accounts for the interaction of stimuli and the needs and goals of the observer (top-down process). By contrast, constructivist theories expand the appraisal theories by focusing on the constraining bottom-up effects of mental representations and language knowledge on the final emotion decision.

However, in real-life applications, naturalistic data are preferred. A naturalistic database can be produced by the observation and analysis of subjects in their natural context. Such a database should allow the system to recognize emotions based on their context as well as the goals and outcomes of the interaction. The nature of this type of data allows for a more accurate system that can fit real-life implementation because it describes states that are naturally occurring during human-computer interaction (HCI).

2.4 Types of Measurement

One critical factor in the accuracy of an emotion model is the measurement method. The literature on RSs stresses the importance of the emotion in producing affective recommendations. In this section, we will highlight the importance of the available emotion reaction measurements with respect to different types of emotion, namely, lower-order and higher-order emotions. Two main types of method are used for the measurement: self-report measures and autonomic measures. Both methods have been used in RS research to map serene reactions to emotion stimuli. The first type, self-report measures, focuses on emotions felt in relation to a given stimulus, whereas autonomic measurements concentrate on emotional reactions that are beyond an individual's control in most cases.

2.4.1 Self-report Measures

Three main types of self-report measures exist with respect to subjective feelings: verbal self-report, visual self-report, and autonomic measures. In the following sections, we emphasize these measurements and distinguish the links between the measurements and the emotion [41].

Verbal Self-report In verbal self-reports, an individual's emotions are collected either through open questioning or by rating emotions that reflect their feelings, usually after direct exposure to a stimulus [41]. With respect to emotion models, psychological emotion research studies subjective emotions based on two major approaches: the basic emotion and the dimensional emotion models. In the basic emotion model, two important scales are used to measure the occurrence of an emotion: Plutchil's Emotion Profile Index (1980) [42] and Izard's Differential Emotion Scale (1977) [27]. Both scales have been used in psychological studies on emotion. By contrast, the dimensional emotion model approach describes emotions based on two or three dimensions: pleasure, arousal and dominance (PAD). One scale to measure the combination of these dimensions was proposed by Mehrabian and Russell in 1974 [30]. In both approaches, verbal scales are rarely used to capture emotional reaction in the context of recommendation.

Visual Self-report Visual self-report measures subjective feelings based on visual cues, such as using cartoon-like figures to represent different emotions [41]. In the literature, this type of measurement is the most widely recommended for RSs, and the most frequently used method is the recently developed Self-Assessment Manikin (SAM) developed by Lang (1980), which relies on Mehrabian and Russel's PAD dimensions. In the SAM, five figures are used to represent each dimension of PAD. The x-axis corresponds to feelings from negative to positive, and the y-axis represents arousal from low to high. Small figures are placed on the x-axis and y-axis, starting from a frowning, unhappy figure that gradually transforms into a smiling, happy figure and a sleepy figure that gradually transform into an excited figure. A recently developed measure implements moment-to-moment ratings rather than static features [43].

2.4.2 Autonomic Measures

In contrast to self-report measures, autonomic measures consider responses that are beyond an individual's control [41]. Such reactions include facial expressions (e.g., smiling) and physiological reactions (e.g., sweating) caused by changes in the central nervous system or PNS. ER based on physiological features can help identify an emotion and define its type. Upon exposure to an emotional stimulus, physiological changes occur, thus generating an emotional reaction. For example, the sweat glands and blood vessels in the skin are exclusively innervated by the PNS. Thus, when an emotional reaction occurs, the skin response is an ideal indicator of changes in arousal, even though the response is psychologically initiated. The heart and the skin are both linked to the same system [44]. Heart rate (HR) and electrodermal activity (EDA) reliably increase under mental stress in laboratory conditions and in dangerous, novel, or challenging situations [45]. In many studies, these changes have led to the conclusion that both skin conductance and HR increase when subjects are exposed to arousing stimuli [44–46].

2.5 Components of Emotional Intelligence

Salovey defines emotional intelligence as the ability to perceive and express emotions, to understand and use emotions, and to manage emotions to foster personal growth [16]. Goleman addressed five domains of EI: knowing one's own emotions, managing emotions, motivating oneself, recognizing emotions in others, and engaging in relationships [8].

2.5.1 Self-awareness

Self-awareness, a critical component of emotional intelligence, comprises emotional awareness, appropriate self-assessment, and a significant level of self-confidence [47]. Self-awareness is a remarkable skill that allows an individual to recognize and understand personal emotions as they occur.

2.5.2 Self-regulation

Self-regulation is an important skill that comprises self-control, adaptability, trustworthiness, conscientiousness, and innovation and enables an individual to manage emotions appropriately [48]. In each circumstance, an individual experiences a range of emotions that require proper management.

2.5.3 Motivation

Motivation refers to personal drive, initiative, willingness, and the readiness to take advantage of opportunities with the core objective of achieving specific goals. Individuals with high levels of motivation stand a better chance of achieving their goals [47].

2.5.4 Empathy

Empathy is the awareness of the feelings as well as the needs of others in different situations [49]. Empathy empowers an individual to extend understanding towards others with an explicit recognition of their feelings, challenges, and situations [50]. Empathy fosters the ability to leverage diversity and to invest in the development of others [51].

2.5.5 Social Skills

Social skills denote a range of interpersonal and relationship skills as well as the competencies that empower individuals to interact and collaborate effectively with others. Social skills have their basis in confidence and self-esteem, which empower the individual to engage in constructive personal dialogue with a clear understanding of different emotions [47].

2.6 Digital Twin

Twins are two offspring produced by the same pregnancy and can be either monozygotic or dizygotic. Monozygotic twins are sometimes called "identical" twins, meaning that they develop from a single zygote that splits and forms two embryos. Dizygotic, or "fraternal," twins develop from separate eggs, and each egg is fertilized by its own sperm cell [52].

Clone was imported into modern language in the early twentieth century to describe grafting techniques in plants. By the late twentieth century, however, the word had become part of the common language and is used technically, figuratively, and sometimes even pejoratively to describe reproductions or carbon copies [53].

Digital twin is a virtual model of a process, service or product. The digital twin, as introduced by Prof. El Saddik, is envisioned as "a digital replication of a living or non-living physical entity. By bridging the physical and the virtual worlds, data are transmitted seamlessly, allowing the virtual entity to exist simultaneously with the physical entity" [54].

How exactly will digital twins help? The digital twin was named one of Gartner's Top 10 Strategic Technology Trends for 2019 [55]. According to Prof. El Saddik, a digital twin facilitates the means of monitoring, understanding, and optimizing the functions of the physical entity and provides continuous feedback to improve quality of life and well-being. A digital twin hence represents the convergence of several technologies, such as AI, AR/VR and haptics, IoT, cybersecurity and communication networks [54].

In the digital world, we do not attempt to create an exact copy of the user but rather provide the user with an identical twin. Thus, a digital twin can enhance mood to improve the quality of life and decision making. Digital twins can enhance data insights and improve decision making [55]. Thus, in this thesis, we build a digital twin with an emotion recognition model. This pairing of the virtual and physical worlds enables the analysis of physiological data and monitoring of emotions to prevent emotional well-being problems before they occur, avoid downtime, develop new opportunities and even plan for the future by using recommendations.

2.7 Related Work

Mental health problems cost Canadian businesses $ 33 billion per year [1], and the cost of lost productivity due to mental illness in Canadian businesses is equal to $11.1 billion per year [2]. Moreover, Chrysalis Performance Inc. research shows that stress in a business contributes to at least 60% of workplace accidents [3]. Additionally, mental health issues caused by stress were the leading cause of short-term and long-term disability in 2005 [4]. The rapid growth in the use of mobile health (m-health) applications provides an opportunity to increase access to evidence-based mental health care and to reduce health care costs.

The primary key to maintaining healthy behaviors in the long term is to watch for adverse changes of any kind, which is why there is a need for a biofeedback system that helps individuals learn skills to manage these stressors and help reduce emotional stress by combining biofeedback with emotional stress management interventions. M-health applications can improve self-awareness and enable individuals to learn to live in the "low-stress zone" by providing users with stress management recommendations.

However, mental well-being applications are an enormous challenge for therapists and researchers. Thus, it is essential in the design procedure to carefully design the concept requirements, integrate mobile solutions, ensure the support needed for new behaviors, and include quality evaluation procedures [56].

Thus, further research into the effectiveness of different components of emotional biofeedback interventions is needed. The current systematic review aims to synthesize the existing research. The review focuses on creating a better understanding of m-health for emotional well-being and whether there is any evidence for using emotional biofeedback in the context of m-health for emotional well-being. Both the diagnostic and intervention methods are the main focus of the review of spatial m-health intervention using wearable devices. There is a special interest in emotion recognition and physiological indices. Systematic reviews aim to synthesize the existing research.

Our goal is to systematically review the research evidence supporting the efficacy of mental health apps in the context of emotional well-being using mobile devices (such as smartphones and tablets) for all ages. Thus, we formulate the following questions:

RQ1. What are the main components of m-health for emotional well-being?

RQ2. In each system, what are the properties of such systems, including intervention and biofeedback methods?

Since m-health data did not exist prior to 2008, a comprehensive literature search of studies published from 2008 to date was set up through electronic databases, including the Scopus search engine, Web of Science, and ACM Digital Library. In addition, a manual snowballing search based on the referring papers was performed. For the study design, only papers with developed applications and evaluated methods and with both an intervention method and a biofeedback method are included. Any work published in English that meets these criteria is included. Table 2.1 shows a summary of the main search query.

Concepts	emotion	wellbeing	Application
Synonym	mental health, emotion*, stress	Wellbeing, Wellness	mhealth, ”m-health”, ”mobile”, app or apps, application*, system*
Additional keywords	Biofeedback	biofeedback, bio-feedback, wearable	
	Intervention	Intervention, management, therapy, regulation	
Search strings	TITLE-ABS-KEY ((emotion* OR “mental health" OR stress) AND (wellbeing OR wellness OR Well-being) AND (interven* OR manag* OR therap* OR regulat*) AND (mhealth OR “m-health” OR "mobile" OR app OR apps OR application*) AND (biofeedback OR "bio-feedback" OR wearable))		

Table 2.1: Summary of the main search query

A total of 97 papers (articles and conference proceedings) was found using all the search engines: 52 from Scopus, 37 from ACM Digital Library, and 36 from Web of Science. Figure 2.2 illustrates the different phases of the systematic review, while table 2.2 summarizes the systematic review results.

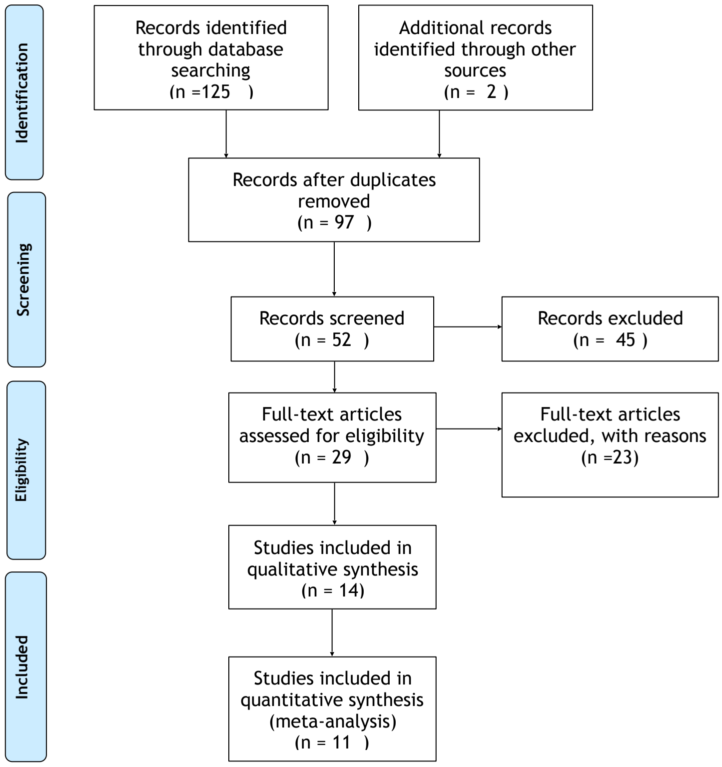


Figure 2.2: Systematic review flow diagram

2.7.1 Sensing

In the literature, several works have used physiological sensors for stress detection. Several methods and sensory data have been used to detect binary stress, such as cStress, in which the authors used a combination of electrocardiograph (ECG) and respiration (RIP) sensors to measure binary stress. Gjoreski et al. used the Empatica E3 wrist device to collect blood volume pulse (BVP), skin temperature (ST), heart rate variability (HRV), galvanic skin response (GSR), and physical activity level using an accelerometer to detect binary stress. Mishra et al. conducted an “in-the-wild” study with 23 participants using physiological data from the users along with “high-level” contextual labels to perceive stress levels. The authors showed that context plays a significant role in the users’ perceived stress levels, and when used in conjunction with physiological signals, this combination leads to much higher stress detection results than relying on just physiological data [57]. Al Osman et al. monitored three mental stress levels via measuring HRV, breathing rate and activity level [58][59].

2.7.2 Biofeedback (BF)

For the feedback, most of the studies in the literature focused on stress management in conditions such as mental stress [60], tension [61] , or depression [62]. Feedback can be provided via one or more of the five traditional senses. The most commonly used form of live biofeedback (LBF) is visual feedback. Al Osman et al. [58] adopted visual feedback in a biofeedback closed-loop system, continuously monitoring mental stress levels via measuring HRV, breathing rate and activity level during serious games for stress management. The feedback was presented in the form of a tree representing the status of the user’s autonomic nervous system. Jercic et al. [63] presented an auction game to train subjects in ER strategies. The game difficulty was adjusted on the basis of the player’s physiologically measured arousal level. The user could gain awareness of his/her emotional state and the influence of his/her emotions on decision making in a financial context using HR data.

The second most common form of feedback is auditory feedback. Millings et al. [62] developed neurofeedback software that transfers data received from a sensor and emits a pleasant, waterfall-like sound. The volume of the sound depends on the power of the alpha frequency band of the user’s EEG spectrum divided by the power of the beta frequency band. The user can try to learn to control the volume of the waterfall sound by increasing his/her alpha power relative to his/her beta power as a stress management strategy.

Other studies have used a combination of sensory data, such as visual and auditory feedback [64] , visual and tactile feedback [65], auditory and tactile feedback [66] or all combined, as in [67] and [68].

The review found no examples of real-time feedback based on emotion recognition using physiological signals. According to our research, no previous studies have reported the use of emotion recognition. Most studies have considered emotions based on the strength of various physiological signals. This gap is due to the challenges related to the complexity of sensor placement, data analysis, and accuracy in emotion recognition systems [69].

2.7.3 Interventions

For the intervention mechanism, the majority of the studies in the review addressed specific ER strategies. For example, deep breathing, as in [58], [65], [66], and [70], or guided relaxation, as in [64] and [71]. Other studies have used random methods through microinterventions, as in [60], [61], and [67]. Very few papers have addressed dynamic interventions, such as [72], by using context-based stress detection and machine learning techniques, as in [68], or using both the recommender system and Q-learning methods, as in [72].

The main factors of dynamic treatment regimes and adaptive interventions (DTRAIs) are the decisions about whether to intervene, the timing, and the intensity needed to maximize the intervention effectiveness [73]. For mental well-being applications, there are two main challenges: the appropriate timing of the effective intervention and the frequency of sending messages to the users. Opportune moment detection based on user biofeedback and behaviors could help us to understand when and for how long to intervene and to better assess the effectiveness of the intervention [74].

For mental health interventions, Matthews et al. stressed ethical requirements. According to Matthews, positive computing systems for mental health care should be informed by accepted scientific models, designed in collaboration with therapists to be integrated with existing clinical practices and used by clients under the supervision of a professional therapist [75].

	Sensing				Biofeedback		Intervention		
ID	# sbject	Signals	pre_meature	post_meature	Focus	Type	Strategy	Duration	Method
[60]	30	GSR	No	Likert questionare	negative stress	Active	distraction	10 mins	a mobile robot
[70]	7	HR, HRV	physical exam	average blood glucose	negative stress	Passive	deep breathing (yogic)	8-16 weeks for 30 mins a day	Assisted relaxation
[64]	63	HR	Stress self tracking and stress level	stress level	negative stress	Passive	guided relaxation, biofeedback, and stress self taking	Using the app for at least 120 second	Assisted relaxation
[71]	36	HRV	No	stress level self report and cardiac activity data	negative stress	Passive	relaxation technique, Bio-feedback	N/A	one method deep breathing
[58]	3	ECG	stress level	stress level and likat questionnaire	negative stress	Passive	stress management exercise deep breathing	13-17 days	one method deep breathing
[72]	N/A	HR	stress level	stress level	positive	Passive	micro- intervention	simulation	Recommender system and Q learning
[61]	12	HRV	NASA task load index	NASA task load index	negative stress	Active	micro- interv relaxation method deep breathing, eye exercise	9am -5pm	Random
[68]	N/A	HR	stress level	stress level	negative stress	Passive	micro- intervention	N/A	context based stress detection using ML

[67]	20	GSR , ECG	STAI, LEQ	stress level	negative stress	Active	micro- intervention	two sections two weeks a part	ramdom
[65]	25	HRV	WAIS	stress level and Likert questionare	negative stress	Active	stress management exercise deep breathing	2 days 3 sessions for 25 mins	one method deep breathing
[66]	24	ECG, EDA	stress level	stress level	negative stress	Active	stress management exercise deep breathing	one session	one method deep breathing

Table 2.2: Summary of the systematic review result

The following acronyms are used in the table. State-trait anxiety inventory (STAI), life events questionnaire (LEQ), adult intelligence scale (WAIS), recommender system (RS), and reinforcement learning algorithm (Q- Learning)

2.7.4 Human State Recognition

Measuring the objective correlatives of subjectively reported emotional states is a major concern in research and clinical applications. Physiological and physical activity information provide mental health professionals with integrative measures that can be used to improve the understanding of patients' self-reported feelings and emotions. The combination of wearable biosensors and smartphones can offer a great opportunity to collect, elaborate and transmit real-time body signals to a remote therapist, who then can help develop a reasonable health plan [56]. To the best of our knowledge, no works exist in the literature that facilitate the use of emotion detection; therefore, we conducted a separate review to address the state of the art in emotion recognition using physiological signals.

Emotions can be recognized based on four categories: audiovisual information, physiological signals, tactile perception, and multimodal form [76]. For real-world scenarios, using physiological signal-based emotion recognition is a promising approach because it is computationally less complex than other methods [76]. This method can provide real-time monitoring of emotions at any time and in any place, providing the user with more mobility and freedom. Table3.1 summarizes the related works that use physiological signals.

In the related works, most studies explain accuracy using pleasure and arousal [77], [78], [79], [80]. Since this method was first proposed, it has been widely employed in many studies on emotion. The highest accuracy was approximately 96% for arousal and 94% for valence using very high volumes of features extracted from heart rate variability (HRV), respiration (RSP) and electrocardiogram-derived respiration (EDR) signals and through the use of principal component analysis (PCA) for dimensional reduction [81]. However, when using the dimensional approach, it is not necessary to map dimensional spaces for a specific emotion, as this approach captures what emotions have in common but not what is unique to a specific emotion. The dimensional approach can capture a particular aspect of an individual's internal state but not the overall emotion. To overcome this problem, a hypertheory was postulated that combines both dimensions and emotion categories [82]. By using the hypertheory, most of the current emotion recognition systems map the two dimensions into two or three classes within arousal-valence areas. When using two classification schemes, the classes are "high" and "low" for arousal and "positive" and "negative" for valence. When there are three classes, they are "calm", "medium" and "activated" for arousal and "unpleasant", "neutral" and "pleasant" for valence [83], [84]. For example, fear and anger share unpleasant and aroused feelings but differ in the external causes and behavioral reactions, yet the model may not capture the difference between them.

Most hyperapproaches use four classes to map emotions into each quadrant of the two dimensional planes [77], [37], [85], [86]. Other works have reported an accuracy of 70% for classifying three emotions (calm, positively excited, and negatively excited) [87], 75% for classifying four emotions (neutral, sadness, fear and pleasure) [85], 45% for classifying six emotions (amusement, contentment, disgust, fear, neutral, and sadness) [88], and 50% for classifying nine emotions (anger, interest, contempt, disgust, distress, fear, joy, shame, and surprise) [89]. Clearly, the complexity is greater when the model addresses more emotions, and thus, the accuracy is reduced. Table I summarizes related works using physiological signals.

To understand individual differences in human emotions and affective phenomena, we must understand and unravel the componential and dynamic nature of emotion and how the componential architecture is dependent on time, the individual, and the situation [28]. In [90], the authors highlight the influence of brain activity, and the individual differences include personality, dispositional effect, biological sex, and genotype. Recent studies have focused on how emotion components and their constellations unfold dynamically over time [91], [92]. In [93], the participants showed sizable individual differences in the duration of their emotions. In [94], the authors showed individual differences in the variability of emotional intensity, while in [95] the authors showed the degree of synchronicity displayed by emotion components over time.

Related work	Predicted emotions	Physiological signal	Stimuli Used	No. of Subjects	Devices	No. of Features	Classification	Accuracy (%)
[87]	Three emotions: calm, positively excited, negatively excited	EEG, GSR, RSP, ST, BVP	IAPS	5	64 electrodes, Galvanic Skin Resistance	384	KNN, LDA	50 % (KNN), 70%(LDA)
[77]	Four emotions: high stress, low stress, disappointment, euphoria	EMG, ECG, RSP, and EDA	Film clips	12	4 sensors including wearable balaclava, EMG, & wireless communication	total of 152 EMG, 3 ECG, 3 Respiration, 7 EDA	SVM, ANFIS	79.3 % (SVM), 76.7 % (ANFIS)
[37]	Four emotions: joy, anger, sadness, pleasure	ECG, EMG, RSP, SC	Musical induction	3	ProComp Infiniti	total of 110 53 ECG, 10 EMG, 37 RSP, 10 SC	PLDA, EMDC	65% (PLDA), 70% (EMDC)
[88]	Six emotions: amusement, contentment, disgust, fear, neutral, sadness	BVP, EMG, SC, RSP	IAPS	10	5 sensors	total of 306 BVP, 6 EMG, 6 SC6 SKT6 RESP	FD, SVM	0% (FD) 45% (SVM)
[89]	Nine emotions: anger, interest, contempt, disgust, distress, fear, joy, shame, surprise	BVP, ECG, EMG, SC, RSP	IAPS	28	ProComp Infiniti	total of 36 6 features for each signal	KNN	50.3% (KNN)

Related work	Predicted emotions	Physiological signal	Stimuli Used	No. of Subjects	Devices	No. of Features	Classification	Accuracy (%)
[83]	Three arousal and valence levels	GSR, RSP, EEG, ECG, and ST	video clips	25	Biosemi active II system	20 GSR, 14 RSP, 64 ECG, and 4 ST	SVM	46.2% for arousal, 45.5% for valence
[81]	Neutral, arousal and valence levels	HR, RSP, EDR	IAPS	35	BIOPAC MP150	very high with the use of PCA	QDC	96% for arousal, 94% for valence
[85]	Four emotions: neutral, sadness, fear, pleasure	ECG, GSR and PPG	Film clips	11	BIOPAC MP150	total of 11867 ECG21 GSR30 PPG	KNN, DT, NB, MLP and RF	68% (5-NN) 66.14% (DT) 29.13% (NB) 64.37% (MLP) 75.69% (RF)
[80]	Level of excitement	ECG, EDA, video and audio	IAPS	27	Biopac MP36, HQ headset + LQ microphones and HD 720p webcam	total of 22540 (video), 65 (audio), 54 (ECG), 66 (EDA)	LSTM network	80% for arousal, 52% for valence
[79]	Neutral and arousing levels	EDA	IAPS	40	Textile electrodes	4	KNN	71.67%
[86]	Four emotions: joy, pleasure, sadness, anger	EQ Radio, IBI	multiple stimuli	12	FMCW radio	27	SVM	72.3%

Related work	Predicted emotions	Physiological signal	Stimuli Used	No. of Subjects	Devices	No. of Features	Classification	Accuracy (%)
[78]	Neutral and arousal and valence levels	EDA, ECG, HR, SCL, SCR	film clips	27	Biopac MP36	total of 11053 (ECG), 37 (RSP), 10 (SC), and 10 (EMG)	CCC	46.3% for arousal, 40.7% for valence
[84]	Three arousal and valence levels	GSR, RSP, EEG, ECG, and ST	video clips	24	Biosemi active II system	20 GSR, 14 RSP, 64 ECG, and 4 ST	SVM	59.57% for arousal 57.44% for valence

Table 2.3: Related works using physiological signals.

The following acronyms are used in the table. For the physiological signals: electroencephalography (EEG), galvanic skin response (GSR), electrodermal activity (EDA), photoplethysmography (PPG), blood volume pulse (BVP), interbeat interval (IBI), heart rate variability (HRV)), electrocardiography (ECG), electromyography (EMG), skin conductance (SC), skin conductance level (SCL), skin conductance response (SCR), skin temperature (ST) and respiration (RSP). For the stimuli: The International Affective Picture System (IAPS) is a database of pictures designed to invoke emotions. For the classification methods: k-nearest neighbors (KNN), linear discriminant analysis (LDA), support vector machine (SVM), adaptive neuro-fuzzy inference system (ANFIS), probabilistic linear discriminant analysis (PLDA), emotion-specific multilevel dichotomous classification (EMDC), Fisher discriminant (FD), principal component analysis (PCA), quadratic discriminant classifier (QDC), decision tree (DT), naive Bayes (NB), multilayer perceptron (MLP), random forest (RF), recurrent neural network models for sequence classification (LSTM), and concordance correlation coefficient (CCC).

One of the main challenges in emotion detection is individual differences. The key point to understanding individual differences in human emotions and affective phenomena lies in unraveling and understanding the componential and dynamic nature of emotion, including how specific emotion components occur and fluctuate and how the outlook of the componential architecture is dependent on the individual, the situation, and time [28]. Recent studies have focused on how emotion components and their constellations dynamically unfold over time and how the phenomenology and componential architecture of emotions fluctuate along with these changes [91], [92]. In [93], the participants showed sizable individual differences in the duration of their emotions. The authors in [94] show individual differences in the variability of emotional intensity over time. [95] studied the degree of synchronicity displayed by emotion components over time.

Time is the main element that explains the divergent dynamics of emotions across individuals. This result highlights the importance of segmentation in the implementation process, in feature selection, and in a classification method that can address the dynamic nature of emotion. In [85], the authors improved the accuracy by clustering the users into three groups based on their IRS and SRS. The corresponding emotion recognizers were built according to these groups, and a new user would be generalized to the recognizer constructed by the similar group response. In [89], the system trained the data into separate subject models and then used KNN to classify the testing data samples into the corresponding subject model according to the similarity of their inner structures. The grouping showed an average improvement of 90%. However, to cluster the users in a real-time application, sufficient and reliable data are needed to assign a user to the appropriate group. Therefore, in this thesis, we use the DTW algorithm, which can address individual differences without the need for user clustering. DTW is a handy tool for matching two time series in order to calculate an optimal path between the two sequences [96]. It has been applied in many time series analyses, and the classification includes gesture analyses [97], speech recognition [98], and video and graphical data [99]. A complete description of the methods and materials is provided in the next section.

For emotion recognition problems using physiological signals, several databases have been established to test the pertinence of this modality, such as MAHNOB-HCI [83], DEAP [100], MIT [101], and HUMAINE [102]. The signals for these databases come from ECGs, GSR, skin temperature (Temp) and respiration volume (RESP), all sampled at 256 Hz. In [103], the authors conducted a comparative study between the DEAP and MAHNOB datasets. The study concluded that the stimulus videos in the MAHNOB dataset were more powerful in evoking emotion than the video clips in the DEAP dataset. Moreover, the results obtained using the recorded signals in the MAHNOB dataset were better than those obtained using DEAP. Among the related studies, we compared our results with those using the MAHNOB dataset [83]. The accuracy was 46.2% for arousal and 45.5% for valence when classifying the affective states into three classes. Another similar work applied early fusion for all features and an SVM classifier [84]. The accuracy was 59.57% and 57.44% for arousal and valence, respectively.

Both SRS and IRS emotion models are considered. On one hand, large differences among individuals exist in terms of how these two fundamental dimensions of affect are related to a person's experience [104]. On the other hand, the quality of self-reports may affect the accuracy of the results [33]. There is a need to use customizable procedures to limit the divergence of emotion dynamics across individuals.

2.8 Concluding Remarks

In this chapter, we have introduced many concepts that we will refer to throughout this thesis, such as digital twin, emotion, and LBF. Most importantly, this chapter has highlighted two major issues for application in emotional well-being, emotion detection and broadcasting, as well as the intervention methods. The chapter has also discussed emotion recognition models using physiological signals. The aforementioned studies support the importance of the segmentation process in selecting the required features. There is a need to use customizable procedures to limit the divergence of emotion dynamics across individuals. To the best of our knowledge, minimal work has been conducted to simultaneously optimize the window frame size, the delay time, and the feature selection. For emotion, which is known to be highly user dependent, improvements to the emotion learning algorithm can greatly boost predictive power.

In addition, there is a lack of intervention methods that address opportune detection to deliver intervention as needed, and the current approaches to providing a comprehensive method to improve a just-in-time mechanism are limited. Also, the literature review revealed a lack of analysis methods to estimate the effect of changes on the current processes and on the satisfaction goals of emotional well-being applications.

Thus, the next chapter proposes a framework that attempts to fill that gap and contributes to solving these problems.

Chapter 3

inHarmony Architecture

This chapter presents a conceptual model of the relevant entities and their relationships in the context being studied. It also discusses the positive computing services, including detecting human behaviors that can signal well-being problems, delivering therapeutic interventions in a timely fashion, and tracking responses to assess the effectiveness of the interventions. Furthermore, the chapter provides illustrative design scenarios and a discussion.

3.1 Design requirements

In the following sections, we highlight the most frequently reported design requirements, standardized guidelines, and methods based on the three-layer design architecture for m-health and mental well-being. The three layers are sensing, detection and broadcasting, and intervention. Table 3.1 summarizes the main design requirements.

3.1.1 Sensing

The sensors should measure physical and biological features that can capture arousal and adaptability [56]. The hardware should be comfortably wearable, mobile, unobtrusive, and easy to connect and should perform its functions with minimal or no maintenance requirements. The provided data should be reliable, robust and consistent with minimal personal variation. Moreover, the design should target continuous and long-term use [105]. It is essential to use offloading computation via the cloud to make the digital twins more scalable and available anywhere and at any time [106].

3.1.2 Human State Recognition

Reflection is one of the stages in the model of personal informatics systems. The design should be hedonic because hedonic techniques are used to induce positive and pleasant experiences [107]. Hedonic technologies also help avoid the risk of focusing too much on negative emotions [73]. By focusing on both positive and negative emotions, the system model can help users to regain balance in the emotional well-being seesaw between personal demands and resources. In addition, the system should be flexible to handle individual differences in both the causes of stress and the user reaction to stress [107] while allowing the user to experience a sense of control and empowerment. Moreover, the system should be fluent. Fluency is obtained by avoiding discrete states so that changes can always be viewed as a process. For example, the transition from one color to another is always performed by blending the two colors [107].

3.1.3 Biofeedback

The biofeedback information is useful not only to allow users to collect real-time information related to their health conditions but also to identify specific trends using historical data that can empower users to self-engage and self-manage their health status [56]. Through the finding patterns and behavior, users can start to determine both what stresses them and how to cope [108]. Positive computing system feedback content and interactions should be offered through more than one of the senses by augmentation to achieve multimodal and mixed experiences. For example, multisensory experiences technology should be used to overlay virtual objects on real-life scenes [109] so that the biofeedback offers direct information based on the stress level. Moreover, recent work in interactive body-centered art has demonstrated that it is crucial to create an immersive experience through live interactions. For example, a widely used metaphor for stress in western cultures is the heart rhythm and pulsation in the interface [107].

3.1.4 Intervention

The intervention methods can be single, random, or dynamic. On the one hand, just-in-time suggestions increase the user's resources for dealing with negative emotions at an early stage, when intervention is most needed, and help achieve the appropriate balance between the timing and frequency of sending messages [73]. On the other hand, the intervention delivery time can delay highly sensitive responses. Thus, it is crucial to provide ultra-low-delay and ultrareliable communications via the use of 5G and the tactile Internet [106] to enable high-quality service. Moreover, pliability is a sensuous quality [107]. Pliability is the degree to which the user is involved and perceives the interaction as a mediator between action and outcome [110]. The system should have some degree of pliability with respect to ambiguity and openness to interpretation that helps the user perceive the reality of what is happening gaggioli2013mobile. Thus, the interventions are tailored according to the characteristics of each user and reinforcement learning [111]

Overall, the digital twin positive technologies must be easy to use for diverse needs [111]. The design must also be simple, as people with symptoms of stress may have certain physical and mental constraints [107]. Moreover, the system should provide the user with history data to allowing him/her to compare and find patterns in different parts of the data [107]. In addition, the system should support and improve the connectedness between individuals, groups, and organizations [107].

Sensing	Effect detection	Biofeedback	Intervention	Ethical	Overall
Robust Long-term use Continuous Easy to connect Comfortably wearable Measure bodily features	Fluency Both positive and negative emotions Flexibility Eudaimonic	Multisensory experiences Highly delay sensitive Aliveness Perceivable Empower Pliability	Dynamic notification Highly delay sensitive Direct information Just-in-time personalized History data	Informed by accepted scientific models Designed in collaboration with therapists Designed for integration with existing clinical practices Used by clients under the supervision of a professional therapist	Social/ Interpersonal Openness Hedonic Simplicity Easy to use Scalable

Table 3.1: Summary of the main design requirements

3.2 Conceptual Model

Lee et al. [112] proposed the "five I's" of intelligent, positive computing using mobile, wearable IoT devices: in-site, intelligent, in-time, intimate, and incorporating services. In-site refers to sensing in a user's daily life to provide quantified self- and contextual data collection. Intelligent identification refers to mining this big data of problematic situations. In-time refers to an intervention that can be supplied to the user through always-on mobile, wearable, and IoT devices in an intimate, personalized fashion. Incorporating refers to using continuous feedback from the user to improve the service experience.

A recent survey carried out by Lee et al. [112] addressed a conceptual framework and highlighted the key components of guidelines for intelligent, positive computing systems research for emotional well-being. Six critical core areas for intelligent, positive computing systems are design methodology, mobile platform design, behavior marker detection, opportune moment detection, device and modality selection, and evaluation methodology.

3.2.1 Design Methodology

We have employed psychological and physiological theories and psychological microintervention methods to support the behavioral lifestyle changes of the users. According to David et al. [113], positive computing system designers should extract principles from a theoretical framework and translate them into the critical technical features of the system using an evidence-based design that is clearly illustrated in the behavioral intervention technology model.

In the context of positive computing for emotional well-being, we used a scenario-based design with scenarios focused on emotional well-being in the workplace. We addressed three types of stress: eustress, neustress, and distress [114]. Neustress is a kind of sensory stimulus or information that is regarded as unimportant, while eustress is good stress that motivates a person to an optimal level of health or performance, and distress is bad stress caused by the negative interpretation of an event that leads to continuous threatening feelings of anger or fear [114]. Distress can be further divided into two categories: acute and chronic. Acute stress is considered severe, but only for a short duration (such as a driver's fear upon seeing flashing police lights in the rearview mirror). Chronic stress remains with the individual for a longer time (such as sadness over personal finances). In the following sections, we will discuss detailed scenarios of the use of an emotional well-being digital twin.

Maya is a college student who came to Canada to pursue her dream of earning a PhD. She is enthusiastic and energetic with excessively high expectations. A few weeks after starting the program, she begins to feel a lack of development. She pushes her limits and works even harder, yet all the hard work has no tangible outcome. Showing acts of denial and avoiding venting, she tends to overload herself and work more hours to make extra efforts to catch up. She gradually develops signs of social withdrawal. Finally, she feels stuck. She cannot do any better, but she does not want to give up her dream. By this time, a close friend has started to notice signs of physical exhaustion. Only then does Maya sense that something may be wrong. Finally, she collapses, loses her belief in a better future and becomes depressed. At this stage, life can seem pointless with no hope for improvement. It has taken her less than two years to develop stress burnout. This final stage is severe and, without intervention, could end in serious chronic illness or even death.

Based on Maya's case, we extract four scenarios that highlight the main components of positive computing to demonstrate the system design and implementation.

Scenario 1 is related to chronic stress. Burnout is a process that usually occurs sequentially and progresses through stages. According to Borritz et al. [115], a low perception of opportunities for personal development in a job is a predictor for burnout in three years. When Maya opens the inHarmony application for the first time, she will be prompted to fill out a survey that includes biographical information as well as burnout symptoms and stage questionnaires. Based on her information, the system will start to cue her with a set of interventions that will interrupt her based on her emotions and the severe level of burnout following the consensual burnout mode. As she is in stage three of burnout, she is feeling sad most of the time. The system first sends her a motivational quotation: "The pain you feel today will be the strength you feel tomorrow.—Livinpalo." She feels better, and the biofeedback shows her a natural emoji face. Maya continues to do some research. The system then attempts an urgent interaction so that she will rest her eyes for 20 seconds. Not long afterward, the system detects the dominant emotion of sadness for ten minutes and sends her another microintervention that includes a link to a PDF file with educational information about burnout as well as an email so that she can read the information in her spare time. Later, the system prompts her with a microintervention and requests her feedback to start a four-minute meditation. She is in need of the intervention, and she clicks yes. The system then sends her meditation cards that contain the setup and a link to a progressive office-appropriate meditation video. She feels much better now, and her biofeedback shows her a natural emoji face.

Scenario 2 is related to acute stress. Maya is nervous about public speaking, and she starts to panic and feels afraid. The inHarmony digital twin application connected to an E4 sensor detects that she is severely stressed, with an approximately 80% feeling of fear over one minute, and it activates when it senses the fear emotion. The system immediately sends a microintervention of guided breathing for one minute. Once the feeling of fear reaches the critical threshold level, the system sends a motivational quotation just in time to help Maya calm down and restore harmony. Moreover, it provides her with continuous emoji expressions of biofeedback. The digital twin is able to provide direct intervention using the just-in-time mechanism without the need for a third party (e.g., a friend) who may not be available when the intervention is most needed.

Scenario 3 is related to neustress. Maya is talking with a friend about her dog, and the system immediately detects the "love" emotion. This is the ideal case for the system to reach out in response to excitement and happiness. The system also provides Maya with visualization over time that can help her realize what triggers her positive emotions.

Scenario 4 is related to eustress. During this type of stress, the system continues to watch for, sense and detect any changes.

3.2.2 Behavior Marker Detection

The automatic identification of well-being problems is the key enabler of the just-in-time intervention [116]. Thus, for behavior detection, we use emotion recognition with physiological signals and tracking of therapeutic responses through a biofeedback system that enables the proactive management of emotional well-being. The detailed model and implementation will facilitate the emotion recognition and biofeedback system, as discussed in (chapter 5).

3.2.3 Opportune Moment Detection

After a behavior marker is detected, it is essential to identify the proper opportune moments for interruption. The interventions must be delivered in a timely fashion [116]. By using both the collected sensor data and user feedback, we can select the automatic identification of opportune moments as occasions for interruption [112]. Chapter 6 illustrates the proposed algorithm for the identification of opportune moments.

3.2.4 Device and Modality Selection

The interruption of users can be achieved through a combination of visual, haptic, and sound notifications for information delivery based on the context and the preferred mood. Once an event is detected, the interruption mechanism starts to work; some interruptions are scheduled regularly (e.g., the 20-20-20 rule). However, it is well known that off-task interruptions sometimes result in productivity loss, increased stress, and time pressure [117]. Therefore, it is crucial for the disruption mechanism to consider the user's receptiveness to interruption or perceived burden of interruption. Thus, it is recommended that the user's feedback be used to improve the accuracy of the detection algorithms and accommodate the user's preferences.

3.2.5 Evaluation Methodology

A randomized controlled trial (RCT), a popular approach, is used as an evaluation method that randomly assigns participants to experimental groups without revealing

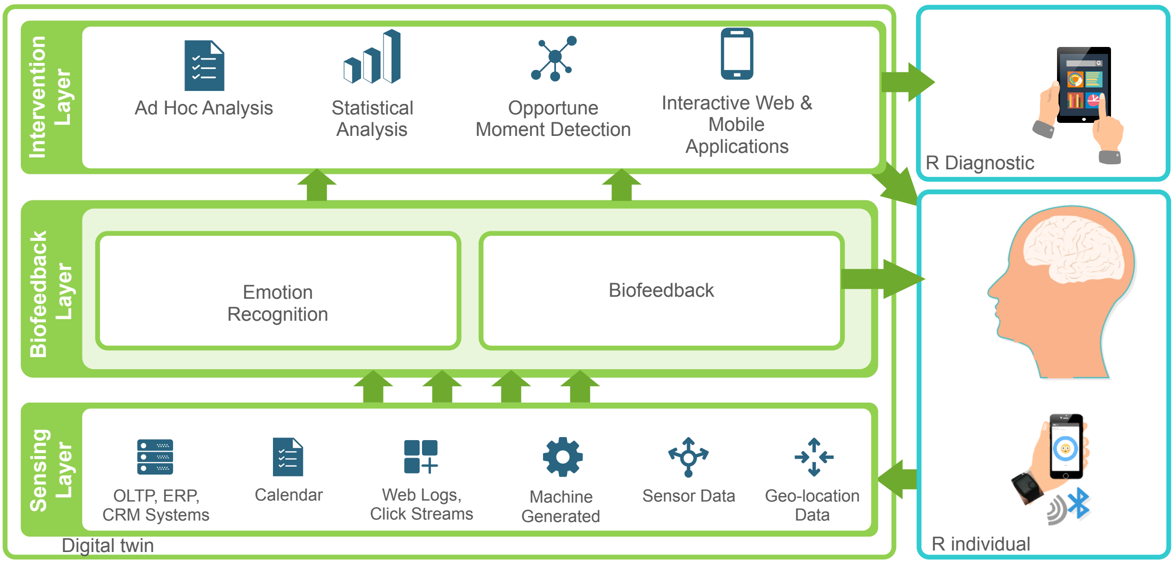


Figure 3.1: An emotional well-being digital twin framework

the assignment information. The usability study provides us with quantitative measures extracted from the usage logs and qualitative findings from the pre- and postquestionnaires that can help us draw conclusions regarding how the proposed strategies affect personal emotional well-being. Chapter 11 evaluates the usefulness of the proposed methods.

3.2.6 inHarmony Platform Design

The digital twin system is a closed-loop feedback system in which information taken from the human body is translated into a language perceivable by any of the human senses (see Figure 3.1). The loop begins with human sensory information from the body. The physiological signals are then interpreted and converted to a recognizable emotion state. Feedback is provided to the individual through recommendations while monitoring the user behavior and action tendency. If the desired quality of service is not achieved, then the system loops back to the matching process to recalculate the feedback. Once the human brain consumes the biofeedback information, a change in the mental state will occur, which will cause a change in the human physiological state.

The cycle then starts again.

Following the three-layered architecture provided by [112], the mobile platform design includes sensing, biofeedback, and intervention layer. Figure 3.1 depicts the platform design.

For the first layer, a data source and communication interface (e.g., row sensor data include EDA, IBI, Temp, and GPS signals) sensing data from sensors and smartphones. The communication interfaces deliver the data collected by the sensors to smartphones via Bluetooth and then to the cloud for further processing. Online transaction processing (OLTP) is supporting transaction-oriented applications on the Internet. An essential attribute of an OLTP system is its ability to maintain concurrency. To avoid single points of failure, OLTP is designed to instantly recording events and reflecting changes as they occur. The data acquisition layer handles the essential functions (e.g., device registration, connection, the sampling rate of raw sensor data, and connection failure handling). Machine-generated data is generated as a result of GPS information in order to draw decision whether or not it is a good time to intervene. We present the details of the sensing layer include data acquisition in the next section.

The next layer, data processing, and analysis implement all the algorithms or machine learning methods for processing the raw sensor data to extract the contextual features and behavioral markers. The system is based on the theoretical model of the cognitive-motivation-emotion (CME) system in [26]. The awareness process in the system modules begins with the human body (physiology), which is connected to various sensors that measure physiological parameters. The captured data are fed into the human state recognition component, which performs multiple signal operations, such as the signal preprocessing and signal analysis. The conclusions are communicated to the human mind using the intervention component through a recommender system. For the feedback, we use a combination of sensory data and multimedia recommendations. Visual feedback is provided through an emoji-based version of the emotions wheel [118]. We also employ tactile feedback through haptic feedback.

Finally, the application and participant interface includes interventions and interactions with the participant. Ad hoc analysis is done in response to an event, such as a sudden dip in positive emotion or an increase in negative emotions. Ad hoc analysis can enable users to get up-to-the-minute insights into data not yet analyzed by a scheduled report. All the captured information is stored in a data repository for short- and long-term trend analysis. Furthermore, the collected data can be optionally transferred to a remote system for health monitoring or statistical analysis by interested parties. We present the details of the biofeedback system in chapter 5, while chapter 6 manifests the intervention system.

3.3 Concluding Remarks

In this chapter, we propose inHarmony, the digital twin system, and its various components: the emotion detection, intervention system, and multimedia response modules. We followed the accepted design principles, standardized guidelines, methods and evaluations of positive computing systems for mental well-being. The next chapter will explain the first layer of the system, including sensing data, and a description of the experimental setting.

Chapter 4

Sensing and Data Acquisition

In this chapter, we present the first layer in the system architecture: the sensing layer. The sensing data include static and continuous data. In addition, this chapter describes data acquisition and presents the experiment for the emotion recognition task.

4.1 Introduction

Measuring the objective correlatives of subjectively reported emotional states is a major concern in research and clinical applications. Physiological and physical activity information provide mental health professionals with integrative measures that can be used to improve the understanding of patients' self-reported feelings and emotions. The combination of wearable biosensors and smartphones can offer an excellent opportunity to collect, elaborate and transmit real-time body signals to a remote therapist, who then can help the individual develop a reasonable health plan [56].

4.2 Methods and Materials

In this section, we describe the data acquisition and present the experiment for the emotion recognition task. We describe the participants, the hardware used, the emotion induction method employed, and the experimental setup. The MAHNOB dataset was selected for this study using peripheral physiological signals.

4.2.1 Static Sensing Data

4.2.1.1 Premeasure Stress Survey

In order to avoid the problems associated with a cold start, we used a premeasure stress survey. The selected survey serves as a tool for risk assessments of health and well-being based on a review of the available tools to measure stress. Moreover, the premeasure stress survey helps to assess the need for action, helps with the decision-making process and preventive measures, and hence helps with personal coping style. Moreover, the survey recognizes the signs and symptoms of stress and helps us understand the causes and effects of stress by recognizing the nature, location, and extent of stress problems for employers. Biographical information (e.g., age, education, experience) is included. The system provides qualitative and comprehensive data in order to develop an action plan and provides the users with recommendations for how to cope with the issues that have been identified and information on practical techniques and strategies to deal with stress.

4.2.2 Continuous Sensing Data

4.2.2.1 Tracking Progress

As an outcome measure, self-reporting questionnaires were used before and after the use of the application to collect demographic data from the participants, and data were collected on the users' satisfaction, implementation of the application and changes in their daily and professional life after use of the application (through, e.g., accelerometers, GPS, smartphone calendars). Self-reflection was chosen as another outcome measure to evaluate the effects of the intervention. The participants were given the mindfulness intervention and asked to write a daily reflection to summarize their work and thoughts on the current treatment day while visualizing the daily emotion percentages for that day. Additionally, biofeedback was used as an essential target for reducing burnout in the workplace.

4.2.2.2 Human Physiology

Using physiological sensors is very common for stress detection. Upon exposure to an emotional stimulus, physiological changes occur, thus generating an emotional reaction. The most common physiological information is blood volume pulse (BVP), skin temperature (ST), heart rate variability (HRV), galvanic skin response (GSR), and physical activity level using an accelerometer to detect binary stress. Both heart rate (HR) and electrodermal activity (EDA) reliably increase under mental stress in laboratory conditions and under dangerous, novel, or challenging situations [45].

4.2.3 Descriptions of E4-dataset

4.2.3.1 Participants

A total of 24 healthy participants took part in this experiment (12 males and 12 females). Their ages ranged from 18 to 40 years, with an average age of 25 years (SD=2.3). All participants were informed about the content and potential risks of the experiment. To assure that the emotional ratings given to the stimuli were due to psychological effects arising at the time of the experiment rather than due to the specific autobiographical memories associated with the particular film clip used [119], we eliminated four participants since they had seen at least one of the film clips. We also eliminated two participants for whom we had insufficient data. None of the remaining 18 participants reported having seen excerpts of any of the clips. Our experiment received ethical approval from the Research Ethics Board of the University of Ottawa.

4.2.3.2 Hardware

An Empatica E4™ sensor was used to collect the subjects' physiological signals [120]. The motivation is that E4 is a multisensor device that is wireless, flexible, and easy-to-use; it is worn on the wrist for real-time computerized biofeedback and data acquisition. Figure 1a shows the sensor. The E4 uses low-energy Bluetooth. Bluetooth is designed to be an ultra-low-power (ULP) protocol to service short-range wireless devices that may need to run for months or even years on a single coin-cell battery. Using a simple stack that enables asynchronous communication with low-power devices, the sensor can send low volumes of data at infrequent intervals. Thus, the connections can be established quickly and released as soon as the data exchange is complete, minimizing both time and thus power consumption [121].

Moreover, the multisensor data fusion represents a practical solution to infer high-quality information even in the presence of noisy signals, data loss, or inconsistency [122]. The E4 wristband has four embedded sensors: a photoplethysmography (PPG) sensor, an electrodermal activity (EDA) sensor, a 3-axis accelerometer, and an infrared temperature sensor. These sensors collect and report information from the sympathetic branch of the autonomic nervous system, including the galvanic skin response (GSR), blood volume pulse (BVP), acceleration, heart rate (HR), and skin temperature (ST).

For data synchronization, Empatica E4 includes an event marking button that when triggered causes a time stamp in the session archive's "tags.csv" file. This event mark then can be used to align E4 data for the analysis stage. Figure 4.1(a) Empatica E4™ sensor for the acquisition of physiological signals, figure 4.1(b) shows a participant wearing the sensor, and figure 4.1(c) shows a subject in the emotion induction experiment.

4.2.3.3 Emotion Induction Method

One of the critical steps in identifying the emotion of a subject is to elicit emotional responses. For the stimulus material herein, we selected the emotional movie database (EMDB) because videos contain more emotional content than a single image. EMDB is a film clip system for emotion elicitation [39]. The database consists of 52 film clips without any sound effects based on emotional stimuli (i.e., valence, arousal, and dominance). These film clips are divided into five categories: erotic, horror, socially negative, socially positive, and neutral. However, using visual stimulation videos without sound was not sufficient for the adequate induction of emotion [38]. Thus, we added audio to the same set of film clips in the EMDB for this study. We decided to use movie clip number 1003 for the neutral condition, clip number 2005 for the happy state, clip number 3005 for the sadness condition, clip number 4000 for the erotic sit-

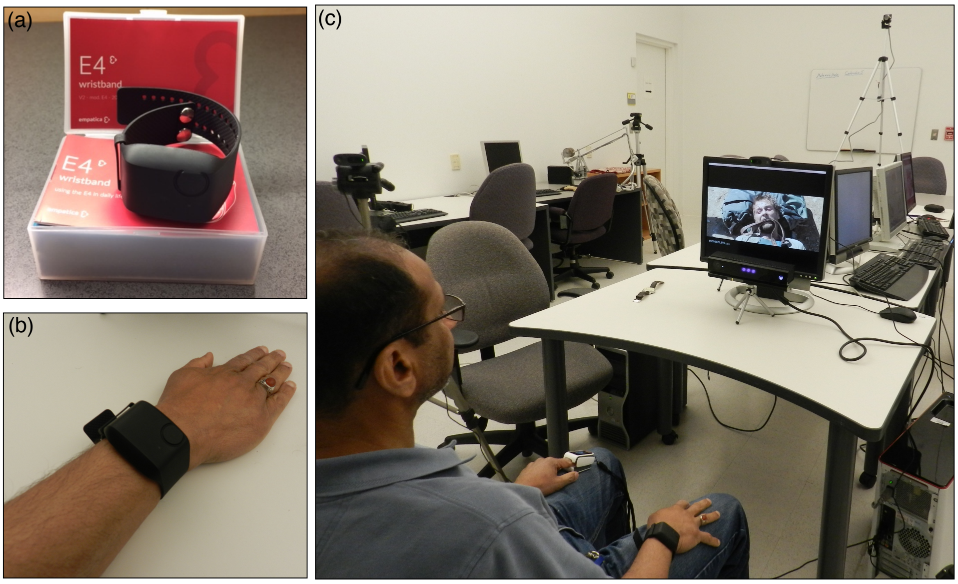


Figure 4.1: Experiment setup; (a) Empatica E4™ sensor for the acquisition of physiological signals. (b) A participant wearing the sensor. (c) The participant during the emotion induction experiment

uation, and clip number 5000 for the horror condition. The source movie files were found on the internet. Subsequently, we extracted the proper audio from the sources and then synchronized the audio with the clips in the database. Relaxing music with a smooth rhythm was played for two minutes between each elicitation of emotion to ensure emotion stabilization.

4.2.3.4 Experimental Setup

For each participant, the experiment involved a session lasting approximately 30 min in a quiet room. Figure 4.2 shows the experimental protocol. Before the experiment, a questionnaire was completed by each subject to record personal information related to the experiment. Then, the researchers explained the experimental procedure and any risks that may occur. Before the trial was initiated, each participant was required to wear the E4 wristband on his or her non-dominant hand. Then, the participants began an active task (briskly cycling for 3 minutes on a laboratory bicycle), as recommended by [120]. An active task helps to build up adequate moisture where the skin contacts

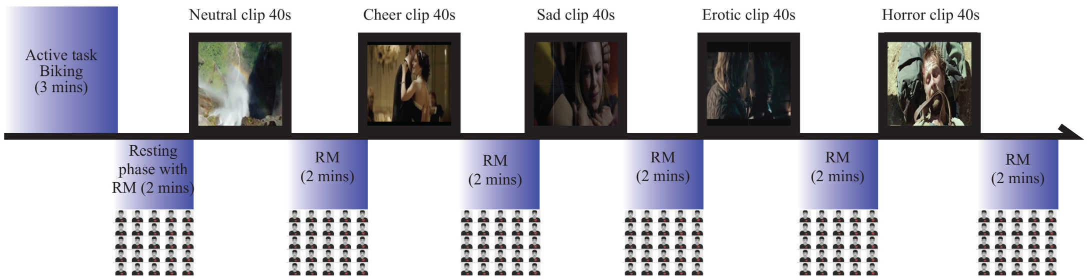


Figure 4.2: Experimental procedure for the data acquisition involving the resting phases and visual stimulation periods. RM corresponds to relaxing music.

the electrodes to allow the electrodes in the EDA sensor to record sensitive changes [120]. This task serves the same purpose for the PPG sensor, which is necessary to obtain reliable HRV data. Next, there was a resting phase in which the participants listened to 2 min of relaxing music. Thereafter, the emotion induction clips were played for the subjects in a sequence, with relaxing music between each set to assure emotional stabilization. The design of the experiment was as follows: an initial maximal expiration task phase lasting approximately 3 min, followed by a resting phase of 2 min, a sufficient visual stimulation period lasting 40 seconds and then a 2 min resting phase and self-report. During the experiment, each participant watched five film clips played in the same sequence to increase the intensity of perception, one for each category of emotion (i.e., neutral, cheer, sadness, erotic, and horror). After each stimuli clip, the participants were asked to report their feelings immediately using AniAvatar; an effective feedback tool for affective states [123].

For data synchronization, Empatica E4 includes an event marking button that, when triggered, creates a timestamp in the session archive “tags.csv" file. This event mark can then be used to align E4 data for the analysis stage. For data acquisition and monitoring, we employed a mobile application to collect and monitor the data during the experiment; the application received the signal data via Bluetooth. The software uploaded the securely encrypted data to a cloud service for data analysis. Herein, four channel signals, such as EDA, HR, inter-beat interval (IBI), and temperature, were used.

4.2.4 Descriptions of the MAHNOB dataset

The MAHNOB dataset is a multimodal database recorded in response to the affective stimuli of emotion recognition. A multimodal setup was arranged for the synchronized recording of face videos, audio signals, eye gaze data, and peripheral physiological signals. A total of thirty participants of both genders and from different cultural backgrounds participated in the experiments. During the emotion recognition experiment, 25 participants completed the experiment by watching 20 emotional videos and self-reported the emotions they felt using arousal, valence, dominance, and predictability as well as emotional keywords [83].

4.3 Concluding Remarks

In this chapter, we collected sufficient physiological data from 18 subjects using five induced emotions (neutral, cheer, sad, erotic and horror) to investigate the performance of physiological responses toward emotion recognition in a user-independent scenario. We conducted a reasonable experiment for emotion elicitation and data collection. In our design, we obtained hundreds of records of physiological signals for each emotion. The next chapter addresses the biofeedback system and includes both the human state recognition and the biofeedback module.

Chapter 5

A Real-Time Emotional Biofeedback System

In this chapter, we propose a real-time mobile biofeedback system that uses wearable sensors to depict five basic emotions and provides the user with emotional feedback. Here, we present a hybrid sensor fusion approach based on a stacking model that allows for data from multiple sensors and emotion models to be jointly embedded within a user-independent model. We applied a meta-learning approach to the dataset and showed that the ensemble approach outperforms any individual method. We also compared the results using two datasets.

5.1 Introduction

Many techniques have been developed to improve the flexibility and the fit of detection models beyond user-dependent models, yet detection tasks continue to be complex and challenging. For emotion, which is known to be highly user-dependent, improvements to the emotion learning algorithm can greatly boost predictive power. Our aim is to improve the accuracy rate of the classifier using peripheral physiological signals.

On one hand, there are large differences among individuals in how these two fundamental dimensions of affect are related to a person's experience [104]. On the other hand, self-report quality may affect the accuracy of the results [33]. Therefore, we proposed a method that uses both emotion models. To obtain the best of both worlds, we proposed multimodal emotion detection using both discrete and dimensional models. The first model used was the model by Ekman [40], which utilizes the physiological signals to map reactions to the emotion's stimuli labeling. The second model used was a hypermodel by Russell [82] [124], which utilizes self-report labeling for induced emotion annotation. WMD-DTW, which is a weighted multidimensional DTW, and the k-nearest neighbors algorithm (k-NN) are used to classify the emotions as a base model. The ensemble methods were used to learn a high-level classifier on top of the two base models.

5.2 Motivations

The understanding of human behavior and cognition has begun to influence the fields of economics, finance, accounting, law, and marketing, among others. When one experiences the cognitive awareness that emotions can affect decisions, better decisions will be made, thus leading to fewer negative consequences. Human decision making often includes a complex combination of emotions and rationalizations. According to the Gallup 2017 Global Emotions Report, 70% of human behavior is based on emotions, not reasoning [125]. Cognitive psychology studies the mental processes that underlie behavior, including thinking, deciding, and reasoning. These topics cover a wide range of research domains, including the workings of memory, attention, perception, knowledge representation, reasoning, creativity, and problem solving [126]. Statistics show the importance of developing the emotion recognition system, which can improve quality of life and decision-making ability.

5.3 Methods and Materials

In this section, we present a description of the methods used in the data analysis.

5.3.1 Human State Recognition Module

5.3.1.1 Preprocessing

There are three main steps involved in the preprocessing stages: filtering, denoising, and segmentation. E4 has two built-in algorithms for the artifact removal process for the BVP signal. The reader can refer to [120] for more details about these algorithms. For the EDA, the Z-score normalization method was used to smooth the signal [127]. For segmentation, we used a meta-optimization task with frame size, delay time and feature as hyperparameters. The process has three steps: define an objective function, define a configuration space, and choose a search algorithm. The objective function finds the best segmentation and features that maximize the accuracy. The configuration space object describes the domain over which the optimizer is allowed to search. We select the shortest and the longest frame sizes in the literature, which are between 10 and 120 seconds. For delay time, in which the measured signals start to be collected, we set the delay parameter to be 0-120 seconds, where 0 sec represents the starting time of the emotion induction procedure, that is, the start of watching each clip. We add the constraint that maximizes the frame size (Fsize) and the delay time (Dt) at 120 seconds (Fsize + Dt $\leq$ 120 seconds). The grid search algorithm is used herein. This algorithm is a standard approach that takes a set of values for every parameter to try and simply enumerates all combinations of the parameter values. The advantage of using grid search over other methods, such as random or Bayesian search, is that it is more accurate [128]; however, it is expensive to run since we have a relatively small dataset. We use Scikit-learn, which is an open-source package for hyperparameter tuning methods. The Scikit-Optimize or skopt library provides algorithms and a parallelization infrastructure for performing hyperparameter optimization. The optimization is performed in three layers. For the first layer, the frame size and delay are set up from 10 to 120 seconds with an increment of 10 seconds in each iteration. Based on the result of the previous step, we performed a second layer optimization by increasing the search around frame size and delay time. We optimized the frame size, delay time and feature selection 10 seconds before and after with increment of 2 seconds at each iteration. Finally, the last layer was performed with an increment of

1 second before and after the parameters from the previous step. For each step, the training dataset was recomputed based on the segmentation results for each emotion class and generated a unique training data frame. We employed one-user-out cross-validation with the F-score of the classification as our metric to identify the best tag delay parameters yielding the highest value. The optimization process is illustrated in the following equation (5.1):

$$w^*(\alpha, \beta) = arg\ min_w\ P(w, \alpha, \beta, S_{train}) \tag{5.1}$$

where w^* is the optimal hyperparameter values and P is the optimization problem, a function of the model parameters w. The function minimizes training error plus a regularization term through cross-validation using a training sample (S_{train}) that is a function of the model parameters w. α and β are a fixed set of hyperparameter values. The function takes as inputs the hyperparameters α and β and returns the validation error corresponding to w^*.

5.3.1.2 Feature Extraction

Multisensor data fusion is a well-established research area in the emotion detection domain. There is wide literature addressing sensor fusion at different levels and using diversified approaches; readers can follow the survey in [122]. Depending on the processing level, data fusion approaches can be divided into three strategies: centralized, distributed and hybrid. The centralized approach relies on a fusion center in which the processing is performed. The distributed approach executes each sensor independently to form a global decision. However, the hybrid data fusion approach benefits from both the centralized and distributed techniques. In this strategy, data collection and preprocessing are performed with a distributed strategy; then, a centralized strategy for fusing data performs the decision-level computation [122]. In this study, we used a hybrid data fusion approach. The following section highlights a detailed explanation of the data analysis methods.

Dynamic Time Warping (DTW) DTW is a useful tool for matching two time series. The algorithm measures the similarity between two sequences that may vary in size in order to calculate an optimal path between the two sequences [96]. Moreover, to cluster users in a real-time application, sufficient and reliable data are needed to assign a user to the appropriate group. Therefore, we use the DTW algorithm, which can address individual differences without the need for user clustering. It has been applied in many time series analyses, and the classifications include gesture analyses [97], speech recognition [98], and video and graphical data [99]. The multidimensional DTW (MD-DTW) algorithm is an extension of the conventional DTW algorithm in which all dimensions are taken into account when identifying the optimal matching path between two series. There are two ways of computing DTW in multidimensional time series: DTW_D [99] and DTW_I [97], where $_D$ denotes dependent, and $_I$ denotes independent. Similar to one-dimensional DTW, DTW_D is calculated by the cumulative squared Euclidean distances of d-dimensional data points T_i and S_j instead of the single data point used in the more familiar one-dimensional case, as shown in (5.2).

$$d(T_i, S_j) = \sum_{d=1}^{k} (T_{i,d} - S_{j,d})^2. \tag{5.2}$$

DTW_I is the cumulative distance of all dimensions independently measured under DTW, as shown in (5.3). Other methods are similar to DTW_I, such as adding up all the dimensions and dealing with a single dimension time series [129].

$$DTW_I(T, S) = \sum_{d=1}^{k} DTW(T_d, S_d). \tag{5.3}$$

DTW_D has only one path to pick for all dimensions to measure their distances. Therefore, DTW_D may not be a suitable distance measure for instances with lag. DTW_I is capable of measuring the distances independently and is invariant to lag. However, since DTW_I is a combination of d paths, it eventually produces a more significant distance score than DTW_D [130]. Similar to one-dimensional weighted dynamic time warping (1D-WDTW) [131] and weighted DTW_D ($WDTW_D$) [132], we introduced a weighted multidimensional DTW (WMD-DTW) to reduce the overall distance score, as shown in the equation (5.4). The optimization method is used to minimize the warping windows of each path and segmentation of the time series and to assign a weight for each diminution to overcome the disadvantage of DTW_I. The w_d is the weight that was assigned to each diminution d. We can write WMD-DTW as:

$$WMD - DTW(T, S) = \sum_{d=1}^{k} w_d \times |DTW(T_d, S_d, ww_d)|. \tag{5.4}$$

where WMD-DTW is the cumulative distance of both dimensions of a two-k-dimensions time series, T and S, measured independently under DTW. $DTW(T_d, S_d)$ is defined as the DTW distance of the d^{th} dimension of T and the d^{th} dimension of S. In equation (3), each dimension is considered to be independent, and DTW is allowed the freedom to warp each dimension independently of the others with respect to the warping window parameter ww_d. The WMD-DTW classifier calculates a distance matrix between two k-dimensions time series T and S according to a warping window parameter ww_d and dimension weight w_d. WMD-DTW uses a 1-NN algorithm to find the smallest distance in WMD-DTW and return the corresponding label associated with that S_d sample within S. The optimization process is illustrated in the next section.

5.3.1.3 Classification and Feature Selection

Tuning the hyperparameters of an estimator Recently, researchers in machine learning have treated the hyperparameter optimization strategy as an important component of all learning algorithms [133]. The most commonly employed optimization algorithms are known as sequential model-based optimization (SMBO) algorithms, also known as Bayesian optimization. In SMBO approaches, there are three main tuning methods: grid searches, random searches, and Bayesian optimization searches. The grid search is the most widely used method; however, grid searches retain the best combination after all the possible combinations of parameter values have been evaluated [134]. As samples are expensive to acquire, the increased precision comes at the cost of a higher runtime. Other search methods have more favorable properties than an exhaustive search, such as Gaussian process (GP)-based optimization, which samples based on educated guesses that are more effective in high-dimensional spaces [135]. However, GP-based algorithms are more complicated than grid searches as they have several free hyperparameters themselves; Bayesian optimization has been shown to obtain better results in fewer experiments than grid search and random search [134].

Thus, GP-based optimization is used for precisely finding the regularization parameter optima in the WDTW.

Gaussian Process-Based Optimization Gaussian optimization is a technique for the global optimization of noisy black-box functions. GP-based optimization is a statistical model in which a GP prior distribution is chosen to describe the unknown function under optimization using an acquisition function [135]. The acquisition function is typically an inexpensive function that can be evaluated at a given point. A GP surrogate model is maintained for the unknown function to determine the optimal next sampling location. The model is updated with every newly obtained sample value [136]. We used Scikit-learn [137], an open-source package for hyperparameter tuning methods. The Scikit-Optimize library provides algorithms for performing hyperparameter optimization. For Gaussian priors, initial points were used to find local minima. The optimal form of these local minima is used to update the prior. The GP forms the posterior distribution over the objective function. Then, the posterior distribution is used to construct an acquisition function for querying the next point. Each acquisition in the equation (5.5) function is optimized independently to propose a candidate point x_i [136].

$$
\begin{aligned}
UCB(x) &= \mu_{GP}(x) + \kappa\sigma_{GP}(x) \\
EI(x) &= E[f(x) - f(x_t^+)] \\
PI(x) &= P(f(x) \geq f(x_t^+) + \kappa)
\end{aligned}
\tag{5.5}
$$

We used upper confidence bound (UCB), negative expected improvement (EI), and negative probability of improvement (PI) as acquisition functions to minimize over the Gaussian prior, where $\mu_{GP}(x)$ is high (exploitation) and $\kappa\sigma_{GP}(x)$ is high (exploration). κ is used in acquisition functions to control the exploration-exploitation trade-off, while x_t^+ is the best point observed so far.

The Optimization and Cross-Validation By using optimization, the hyperparameter values function evaluates a validation set. Cross-validation methods are used to evaluate the machine learning model at each iteration of the GP. The aim is to reveal the highest accuracy possible at the lowest number of iterations possible. We employed two-user-out cross-validation with the F-score of the classification as our metric to identify the best parameter optima in the WMD-DTW. At each iteration, one user's data were used for optimization validation, while the other user's data were used for the meta-learning approach to prevent overfitting.

Ensemble Learning The ensemble method is a powerful technique for improving the modeling of the learning algorithm and significantly boosting predictive power [138]. It is a linear model in a very high-dimensional space where each point in the space represents the model weights. It can improve the classification accuracy by canceling out independent errors made by individual classifiers [139]. The most popular methods include Bayesian model averaging, bagging, boosting, and stacking. The Bayesian model uses a vote proportional to the likelihood [140], bagging uses an unweighted majority vote [141], boosting uses a weighted majority vote, and stacking learns an ensemble classifier based on the output of multiple base classifiers [138]. Some studies [140–143] show that stacking has robust performance. In fact, stacking performs better than both boosting and bagging [141]. It also outperforms the Bayesian model averaging scheme [140]. Thus, we used stacking to learn a high-level classifier on top of the base classifiers, called first-level classifiers.

Stacking Classifier Stacking is a successful technique to combine multiple models in the ensemble model to achieve better performance. It is a meta-learning approach in which the base classifiers are called first-level classifiers, and a second-level classifier is learned to combine the first-level classifiers [139], as shown in Figure **??**. At the first-level classifiers, we apply a bootstrap sampling technique to learn independent classifiers, adaptively learn base classifiers based on data with a weight distribution, tune parameters in a learning algorithm to generate diverse base classifiers (homogeneous classifiers), and apply different classification methods and/or sampling methods to generate base classifiers (heterogeneous classifiers). At the next stage, the second-level classifiers, we construct a new data set based on the output of base classifiers. Here, we used class probabilities from the first-level classifier as new features in the

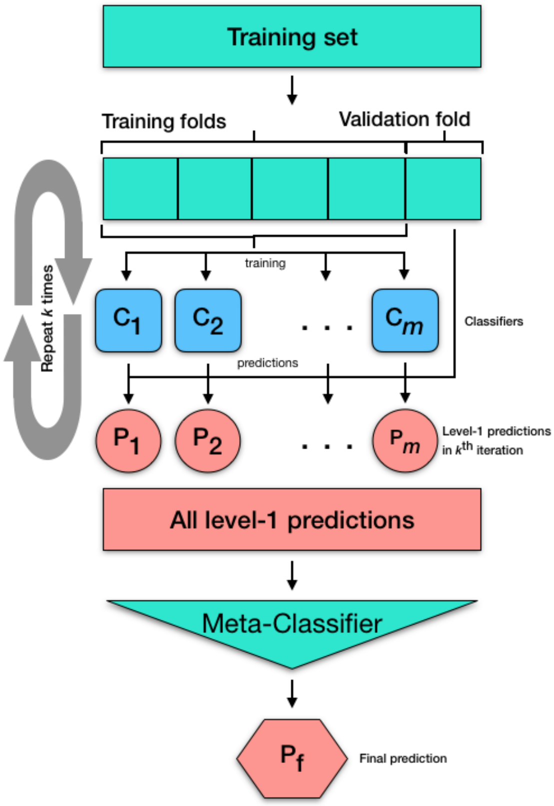


Figure 5.1: The Stacking Cross-Validation Procedure [139].

second-level classifier to train the meta-classifier for successful stacking [141]. The advantage of using conditional probabilities as features is that the training set of the second-level classifier will include not only predictions but also prediction confidence of first-level classifiers. We used an open-source python library, StackingCVClassifier [139].

Stacking and Cross-Validation To avoid overfitting, a stacking cross-validation method within first-level classifiers is used [143]. Similar to the base classifiers, we used one-user-out cross-validation to evaluate classification performance. The user data used in this validation are not used in the model generated in the first-level classification. The first-level classifiers are fit to the different data set that is used to prepare the inputs for the second-level classifier.

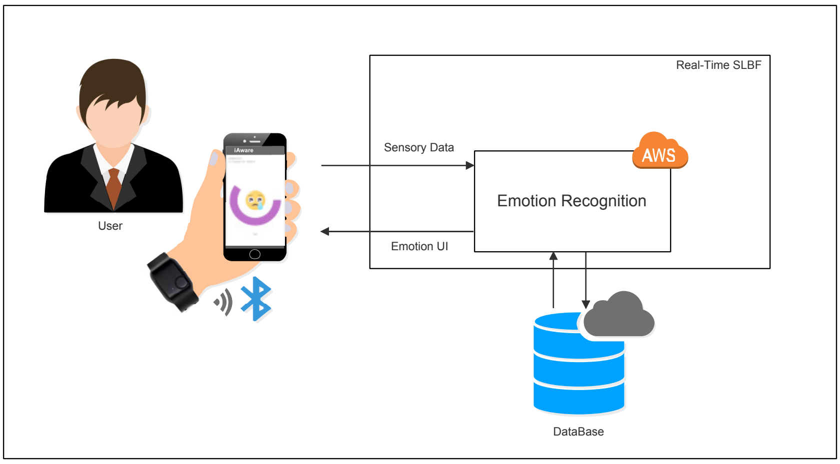


Figure 5.2: The iAware emotion monitoring system

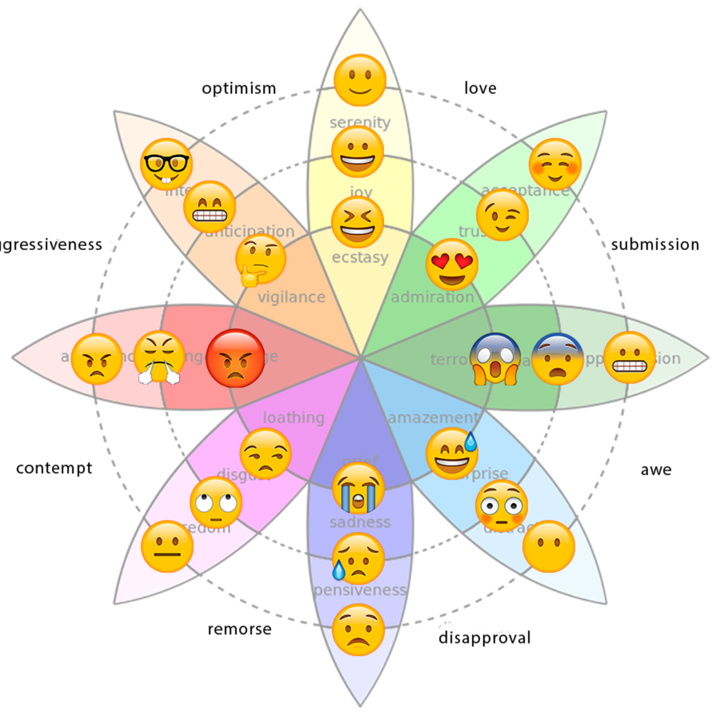


Figure 5.3: Emoji-based version of the emotion wheel [118]

5.3.2 Emotional Biofeedback Module

By considering the biofeedback system of Al Osman et al. [59], Figure 5.2 shows the continuous monitoring of emotion using physiological signals.

For the feedback, we used a combination of the sensory data. Visual feedback was provided using an emoji-based version of the emotions wheel [118]. Kazim et al. developed an emoji-based version of the emotions wheel for Emotive UI [144]. The wheel is a combination of the standard Plutchik wheel and the emotional values associated with each color using emojis that reflect the emotional spectrum; see Figure 5.3. We also employed tactile feedback using haptic feedback. The vibration intensity was set from low to high based on the emotional intensity. The fear emotion was set to high, and the happy emotion was set to low, while the neutral state was set to zero. We designed two experiments: one was a controlled experiment, and the other was a noncontrolled test.

5.4 Concluding Remarks

The implementation of a physiological, signal-based emotion recognition system involves several stages: physiological signal data acquisition, preprocessing, feature extraction, feature selection, and classification. Each step has been discussed and summarized in this chapter. We improved the emotion elicitation by using both discrete and dimensional emotion models. Moreover, we proposed a real-time emotional biofeedback system based on physiological signals. The design and implementation of the system allow it to be easily adapted for SLBF or FLBF using a cloud service. In the next chapter, we will describe the intervention recommendations and the opportune moment detection.

Chapter 6

Intervention system

In this chapter, we present the motivation for the development of a dynamic intervention algorithm. We also present the intervention and recommendation methods, including the opportune moment detection. In addiction, we describe the dynamic mindfulness model by adopting the basic elements of a generic four-step action plan.

6.1 Introduction

Emotional intelligence focuses on how individuals identify, understand, express, regulate, and use their own feelings and those of others [9]. It is important to determine what type of intervention to provide and how and when to intervene. By knowing the burnout type and subtype, coping strategies can be used to improve emotional regulation to contribute to reducing the psychopathological symptoms of burnout. This aspect is reflected in the positive relationships between coping style and burnout symptoms and in the tendency to rely on coping strategies and practices to relieve stress [145].

In the literature of stress management interventions (SMIs), mindfulness seems to be a promising intervention for reducing stress and enhancing well-being in a workplace [146]. Kabat-Zinn defined mindfulness as paying attention to the present in a conscious and nonjudgmental way [147]. According to Shapiro et al., this definition embodies the three axioms or building blocks of mindfulness, intention, attention, and attitude, depending on individual preferences, burnout subtypes, and the working environment [148]. By recognizing a feeling as it develops, one can then properly handle the feeling. Employing coping strategies during times of stress appears to be beneficial [149]. However, studies have shown that attempting to express emotions and understanding a stressful event increase the symptoms of burnout [150]. Counterintuitively from a clinical standpoint, other research suggests that by actively trying to problem solve in a situation, an individual feels that he/she has more control over the situation, giving him/her a sense of control and power [149]. Passively attempting to understand the feelings related to a stressful event may offer more introspection. The solution for this is not only to help users identify the stressors causing burnout but also to provide the impact of this identification on their emotions as a form of emotional biofeedback similar to that in [151]. Therefore, it is crucial to examine the coping mechanisms that provide the most significant reductions in stress in the workplace; therefore, we developed a dynamic mindfulness model by adopting the basics of a generic four-step action plan by Leiter et al. [152].

6.2 Methods and Materials

In this section, we present a description of the methods used for intervention development. We developed a dynamic mindfulness model by adopting the basics of a generic four-step action plan by Leiter et al. [152]. The steps are as follows: defining the problem, setting objectives, taking action, and tracking progress. The present study aimed to investigate the feasibility and effects of mindfulness recommendations that are primarily developed and implemented to better cope with stress.

6.2.1 Defining the Problem

Research shows that stress appears to manifest in different ways. Therefore, it requires specific evaluation and various intervention approaches. These dimensions may include the core definition of burnout. The participants' levels of stress were classified into three phases: overload, lack of development and neglect. These classes were created based on the cut-off scores on the BCSQ-12 (overload 3.38; lack of development 3.63;

and neglect 2.63) to differentiate between 'exposed' and 'nonexposed' in each of the conditions [153]. Participants not meeting these criteria were considered to suffer from a lower severity of symptoms.

6.2.2 Setting Objectives

According to Montero et al., each prolonged stress profile shows different features and relationships with other health-related symptoms that can help treat each case differently [154]. The typology can also identify targets of intervention for each case. It is here that a coping strategy is used as a coping mechanism for self-regulation, value clarification, and psychological flexibility. The coping strategy plays an important part in mediating variables in the development of burnout [146]. Thus, we used the weighted factors of the burnout syndrome from the BCSQ-36 [154] to address the specific sources of distress along with the user negative feelings (e.g., sad and scared), which can then be used to provide convenient and real-time intervention.

6.2.3 Taking Action

With inspiration from the self-help book Banishing Burnout: Six strategies for improving your relationship with work by Leiter et al. [152], smartphone-based mindfulness intervention in the form of microintervention is used. While being used for a maximum time of 80mins per participant per day for one week, the application provides the user with mindfulness intervention consisting of a short web-based psychoeducation, a step-by-step mindfulness practice program and motivational quotes administered via a smartphone application. For example, the participants were advised to begin with short mindfulness exercises, such as a guided 3min mindfulness exercise, which was one of the audio tracks in the application.

6.2.4 Tracking Progress

As an outcome measure, self-reporting questionnaires before and after the use of the application were used to collect demographic data from the participants, and data were

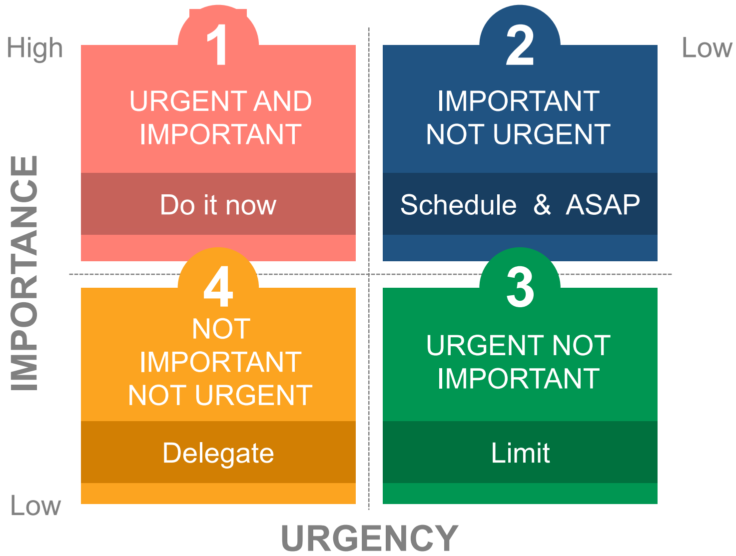


Figure 6.1: Eisenhower Matrix for task management [1]

collected on the users' satisfaction, implementation of the application and changes in daily and professional life after use of the application.

Self-reflection was chosen as another outcome measure to evaluate the effects of the intervention. The participants were given the mindfulness intervention and asked to write a daily reflection to summarize their work and thoughts on the current treatment day while visualizing the daily emotion percentages for that day. Additionally, biofeedback was used as an essential target for reducing burnout in the workplace.

6.2.5 Opportune Moment Detection

The selection of the content and the time of the intervention are based on judgments regarding the set of criteria under the notion of decision-making methods. These methods include ranking, rating, and pairwise comparison (eigenvector) methods. Since we have two main factors that can determine the intervention content and include the time and burnout level, we adopted the time management decision-making method called the Eisenhower decision matrix [1]. Using the Eisenhower decision principle, the content was evaluated using the criteria important/unimportant and urgent/not urgent. Figure 6.1 illustrates the Eisenhower Matrix for task management. The system provides immediate intervention in urgent cases while waiting for user feedback in not urgent cases. By using the user feedback, the system gives the user some space to decide whether to receive interventions. Moreover, the user can have some degree of pliability with respect to ambiguity and openness to interpretation.

6.3 Concluding Remarks

In this chapter, we describe the dynamic intervention system, including the opportune moment detection, with automatic identification of opportune moments as an interruption instance. The type of intervention is taken into account by the design requirements, including pliability, ambiguity, and openness. The next chapter covers a case study on workplace stress management techniques for both individuals and organizations.

Chapter 7

inHarmony for the Workplace: A Case Study

This chapter presents a case study in which the digital twin system is used in a real context in order to assess the effectiveness of inHarmany, a digital twin for individuals and organizations to help improve well-being in the workplace. The system addresses the five principles of emotional intelligence by providing the user with emotional biofeedback for self-awareness, stress management techniques for self-regulation as well as motivation and social skills. Recognizing and sensing how other people feel will improve empathy skills and power relationships.

7.1 Introduction

Workplace stress is a pressing issue that concerns Canadian employers. The costs of workplace stress are not limited to the individuals who experience stress. The Journal of Occupational and Environmental Medicine reports that health-care expenditures are nearly 50% greater for workers who report high levels of stress. There is substantial evidence that ignoring an unhealthy, unsafe workplace is costly. Prolonged stress can be costly for employers since it can result in increased absenteeism or a decline in productivity. For example, research shows that stress in the workplace contributes to 19% of absenteeism costs [155]. The Canadian Policy Research Networks estimate that stress-related absences cost Canadian employers approximately $3.5 billion each year [156]. Mental health problems cost Canadian businesses $33 billion per year [157]. The cost of productivity lost due to mental illness in Canadian businesses is equal to $11.1 billion per year [158]. Moreover, Chrysalis Performance Inc. research shows that stress in a business contributes to at least 60% of workplace accidents [155]. Additionally, mental health issues caused by stress were the leading cause of short-term and long-term disability in 2005 [157].

Stress is perceived differently by everyone. As a result, not all individuals who experience stress notice the stress experience. Therefore, it is important to learn to discern our body's subtle cues and how the body reacts to stress. For example, we tend to become stressed when we experience overwhelming situations, are pressured by high demands, have minimal control and few choices, feel poorly equipped and harshly judged, and have steep or unpredictable consequences for failure.

How does stress affect us? There are several different types of stress, ranging from eustress, which is a positive form of stress from excitement, to chronic stress. The changes caused by stress may be emotional, physical or behavioral or a combination of all three. The pain from emotional stress can have a greater impact than some other types of stress. Therefore, it is important to be able to manage emotional stress in practical ways. Moreover, there is a long list of commonly experienced effects from stress that range from mild to life-threatening. When experiencing stress, bodily functions change; for example, the heart rate increases, muscles tighten, blood pressure rises, sweating increases and breathing quickens. When stress levels become very high, many people experience stress symptoms such as, headaches, irritability, and ' fuzzy thinking,' which are all symptoms that a person is under too much stress. One of the main consequences of work-related stress is burnout [159].

The primary key to stress management is to watch for adverse changes of any kind. This is why there is a need for a biofeedback system that helps individuals learn skills to manage these stressors and help reduce experiences of stress. Such an application can improve self-awareness to transform our experience of chronic stress to pleasure, productivity, and creativity, allowing individuals to learn to live in the " low stress zone " and providing a stress management recommendation when needed. Thus, this chapter proposes a digital twin for individuals and organizations to help improve well-being in the workplace. For individuals, the system can help the user recognize the signs and symptoms of stress, understand the causes and effects of stress and provide an action plan by providing users with recommendations on how to address the issues that have been identified and practical techniques and strategies to deal with stress. Moreover, the system provides the organization with qualitative data and comprehensive reports while respecting employee confidentiality. By recognizing the nature, location, and extent of stress problems for employers, directors can then help employers provide various resources to create a resilient and healthy workforce environment.

7.2 Background

7.2.1 Conceptual Stress Models

Burnout models have been developed based on theories about job stress and the notion of imbalances leading to strain. According to this theory, burnout is defined as a persistent physical, mental or emotional exhaustion caused by long-term stress, usually as a result of excessive stress from the workplace, personal responsibilities, or both [160].

There have been various conceptual models about the development of stress and its subsequent impact. One of the well-established models is based on phases and transactional burnout, which serves as the conceptual bridge between sequential stages and imbalances [153]. The model's three stages are job stressors (e.g., an imbalance between work demands and individual resources), individual strain (e.g., emotional response of exhaustion and anxiety), and defensive coping (e.g., changes in behavior, such as greater cynicism) [160]. The model with the three dimensions hypothesized a sequential progression over time, in which the occurrence of one dimension precipitates the development of another. According to this model, exhaustion starts first, leading to the development of cynicism, which subsequently leads to inefficacy [161].

More recently, Maslach and Leiter developed a model that is based on engagement and burnout as the opposite poles of a continuum of work-related well-being; burnout represents the negative pole, and engagement represents the positive pole

[162]. This model incorporates more information related to antecedents of classic or standard symptoms of burnout. Based on the phase models, researchers have established tools to measure the progression of phases from low to high burnout and the effect of burnout on both work and personal well-being.

7.3 Methods and Materials

In this section, we describe emotion recognition methods and present the experiment for the emotion recognition task. We describe the participants, the hardware used, the emotion induction method employed, and the experimental setup. Finally, we present a description of the data analysis in terms of preprocessing, feature extraction, and feature selection.

7.3.1 System Architecture

One of the core benefits of real-time technology is that it improves and automates real-time analysis. Thus, this study explores digital twin application with stress management themes for the workplace environment. It aims to determine the desired emotion and well-being, as well as their potential contribution to behavioral change. By considering the architecture of the emotional communication system, Figure 7.1 illustrates the digital twin architecture of stress and emotion communications in the workplace. The system consists of three modes: individual, mutual, and diagnostic.

7.3.1.1 Individual Mode

The benefits of such an application can be realized in two-stream biofeedback and stress management recommendations. Biofeedback helps the user become more aware of their stress and emotions, enables them to enjoy small moments of happiness and provides them with self-monitoring tools. The recommendations promote relaxation, which can help relieve a number of conditions that are related to stress. The suggestions are based on the emotion, stress trigger type, and user coping style. Eventually, users can learn how to control these functions on their own.

7.3.1.2 Mutual Mode

Similar to the individual mode, in the mutual mode, the user can select an option to share the emotional biofeedback. The biofeedback then serves as a tool to help the user become more aware of other stressors and emotional states and provides support and sympathy for friends and colleagues.

7.3.1.3 Diagnostic Mode

For the organizational side, the system provides directors with access to real-time analytics to understand what is happening in the workplace and exposure to real-time data to help improve the understanding of stress in the work environment. Moreover, the system offers dashboards and data visualizations for stress competency review for managers. The dashboard shows a highly informative view of the status of workplace stress to allow the organization to identify potential trends and react to potential issues. The activity streams help human resources be aware of what is happening within the organization so that changes can be identified and solutions quickly implemented. The data visualizations help extract more information to identify trends and changes in the data that would be very difficult to identify using only raw data. All of these tools are great visual indicators and triggers that can help interpret historical information.

Figure 7.2 demonstrates the communication and execution sequence diagram between users, terminals, and the cloud.

7.3.2 Sensing Module

Several questionnaires can be used as a diagnostic tool to diagnose burnout based on self-report transactional levels. The most widely accepted standard for burnout assessment is the Maslach Burnout Inventory (MBI) [163]. The MBI is made up of three subscales: personal accomplishment, emotional exhaustion, and depersonalization. The items are answered on a seven-point Likert-scale, ranging from 0 (never) to 6 (every day).

In contrast, the Utrecht Work Engagement Scale (UWES) assesses a mental state of accomplishment as an antithesis to burnout [164]. The most commonly used UWES-17

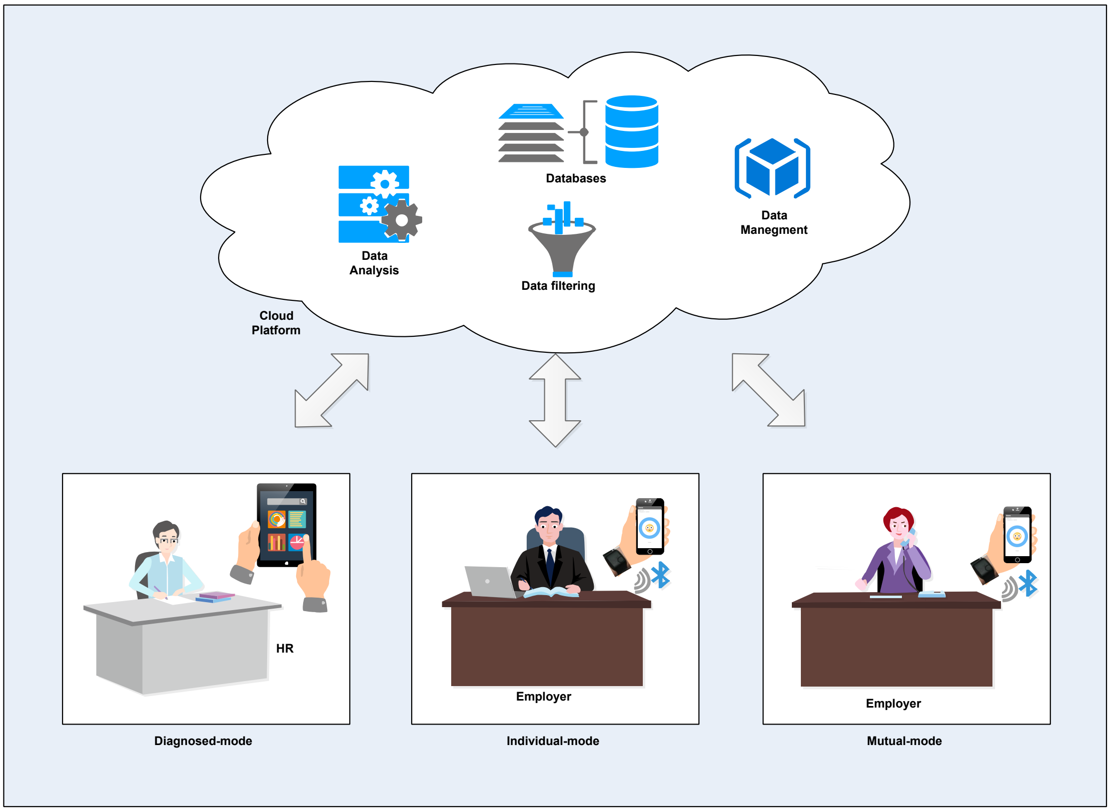


Figure 7.1: The digital twin architecture of stress and emotion communications in the workplace

Figure 7.2: The execution sequence diagram of the system

questionnaire consists of 17 items equally distributed between three dimensions: vigor, dedication, and absorption. The questions are answered on a seven-point Likert-scale, ranging from 0 (never) to 6 (every day). The Burnout Clinical Subtype Questionnaire (BCSQ-12) represents both the negative and positive affects [153]. The BCSQ consists of 12 items equally distributed between three dimensions: overload, lack of development, and neglect. The items are answered on a seven-point Likert-type scale, scored from 1 (totally disagree) to 7 (totally agree).

The BCSQ-12 presents a higher explanatory power than the standard MBI-GS. Compared to other measurements, the BCSQ-12 assumes that engagement and burnout constitute opposite poles of a continuum of work-related well-being, with burnout representing the negative pole and engagement representing the positive pole [153]. Thus, we used BCSQ-12 as a measurement tool because not only is the BCSQ-12 a valid and useful tool for organizational evaluation, but the subtype levels of the BCSQ-12 are more comprehensive and can suggest a pattern of burnout contributions or stressors that can then be used for more specific treatments to increase the efficacy of interventions [154].

7.3.2.1 Summary of the Sensing Information

- Static sensing data
  - Burnout Clinical Subtype Questionnaire (BCSQ-12)
  - Biographical information (e.g., age, education, experience)
- Continuous sensing data
  - Human physiology using E4 sensor
  - Accelerometers, GPS, smartphone calendars

7.3.3 Emotion Recognition Module

In this study, we employed a hybrid data fusion approach using a weighted multidimensional DTW (WMDDTW). The optimization method was used to minimize the warping windows of each warping path and segmentation of the time series and to assign a weight to each diminution. The reader can refer to the methods provided by Albraikan et al. [165].

7.3.4 Intervention Delivery Module

We developed a dynamic mindfulness model by adopting the basics of a generic four-step action plan by Leiter et al. [152]. Table reftable:stressManaInter shows the typology of stress management interventions for both the individual and organizational levels based on the three stress levels. Figure 7.3 illustrates the flowchart of the time and content of the interventions.

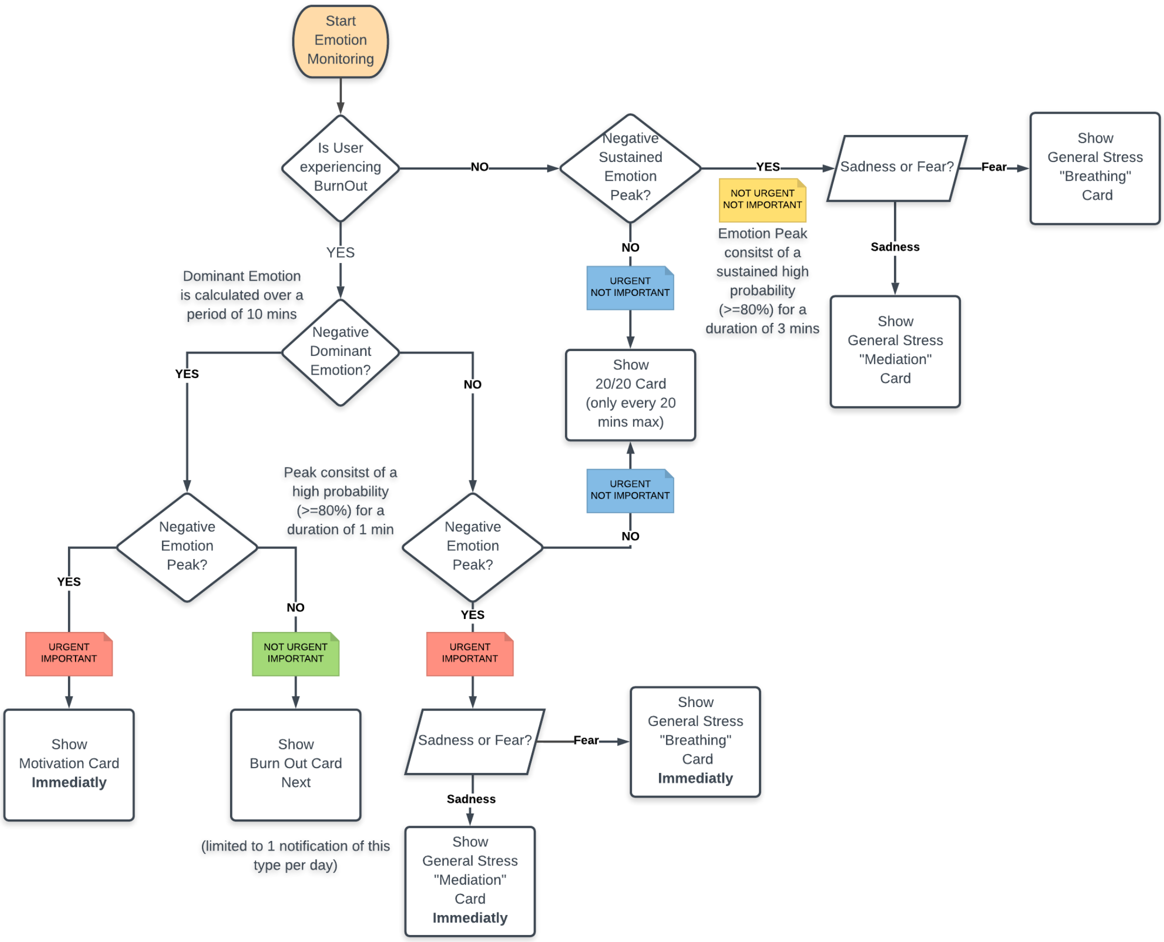


Figure 7.3: Flowchart of the time and content of the interventions

Intervention type	Individual	Organizational
Primary	Selection & Assessment Pre-employment medical examination	Job RedesignWorking time and schedules Management training, e.g., mentoring
Secondary	Mindfulness training Health promotion, e.g., exercise Cognitive behavioral therapy Relaxation Meditation Personal and interpersonal skill training Acceptance and commitment therapy Psychosocial intervention training Coping skills training Resilience training	Improving communication and decision making Conflict management Peer support groups Coaching & career planning
Tertiary	Employee Assistance Programs Counselling Post-traumatic stress assistance Disability management	Vocational rehabilitation Outplacement

Table 7.1: A Typology of Stress Management Interventions [2]

The selection of the content and the time of the intervention are based on judgements regarding the set of criteria under the notion of decision-making methods. These methods include ranking, rating, and pairwise comparison (eigenvector) methods. Since we have two main factors that can determine the intervention content and include the time and stress level, we adopted the time management decision-making method called the Eisenhower decision matrix [1]. Using the Eisenhower decision principle, the content was evaluated using the criteria important/unimportant and urgent/not urgent. Based on the importance level of the intervention, another card will override. Figure 7.3 illustrates the time and content of the interventions. The levels are

1. For important/urgent intervention, immediately intervene with the user when burnout stages are identified using the BCSQ-12 survey, and the dominant emotion over the past 10 minutes is a negative emotion or an instance of negative emotion. The intervention content is short (e.g., a motivation card that includes a quotation). For the quotations, we used the Goodread dataset, which contains the 2999 most popular recent quotations shared on Goodreads. Each quotation includes tags, the number of likes, and the author of the quotation[1].

[1]https://www.kaggle.com/pramud/most-popular-quotes-on-goodreads

2. Important/not urgent intervention is used in cases of identified burnout stages and dominant negative stress (e.g., an educational material card based on the stress phase is shown within the next 10 minutes). Table reftable:stressManaInter shows intervention types based on the stress level.

3. Unimportant/urgent is used in cases of no burnout; only symptoms of burnout and some instances of negative emotion are identified (e.g., the 20/20/20 rule card is shown to the user every 20 minutes).

4. Otherwise, for unimportant/not urgent intervention, stress prevention techniques are used (e.g., mindfulness and meditation practices, stress education). For the experiment, we have the additional constraint that a maximum of five cards per day is the limit for every participant, and no more than one card is shown every 10 minutes. This can be reset by the user at any time. Table 1 provides a summary of the intervention typology of the SMI.

7.4 Concluding Remarks

The current chapter presented a case study used to assess the effectiveness of the inHarmony application for the workplace. We proposed a convenient and easy method for managing and scoring stress levels in a workplace context for both individuals and organizations. The solution can help increase productivity and efficiency, boost staff morale and reduce absenteeism. The next chapter addresses the results for the main component of the inHarmony system. Moreover, chapter 10 covers the usability study for workplace stress management techniques for both individuals and organizations.

Chapter 8

Results and discussion

This chapter discusses the results to evaluate the effectiveness of the emotional well-being digital twin. The evaluation includes emotion recognition methods, live biofeedback and the intervention. We end this chapter with some concluding remarks.

8.1 Emotion Recognition

For the analysis, we used two datasets: the E4 dataset that we collected and the MAHNOB dataset. For the E4 dataset, Figure 8.1 shows the confusion matrix[1] for both models and the stacking model. Table 8.1 shows a summary of the results of the various models that were used for the E4 dataset. The accuracy is calculated using the WMD-DTW and 1-NN classifiers. For the MAHNOB dataset, Table 8.2 shows the comparative summary results of the various models that were used for the MAHNOB dataset. Figures 8.2 shows the confusion matrix[1] for the various base models with and without padding, where Figure 8.2.a for arousal and 8.2.b for valence. Figure 8.3 shows the overall staking results for both arousal and valence.

[1] The detailed confusion matrix results include the overall accuracy, the precision, and the recall. The column on the far right of the plot shows the positive predictive and the false discovery rate. The row at the bottom of the plot shows true positive and false negative rates. The cell in the bottom right of the plot shows the overall accuracy. The diagonal cells correspond to observations that are correctly classified. The off-diagonal cells correspond to incorrectly classified observations. In each cell, both the number and the percentage of observations are displayed

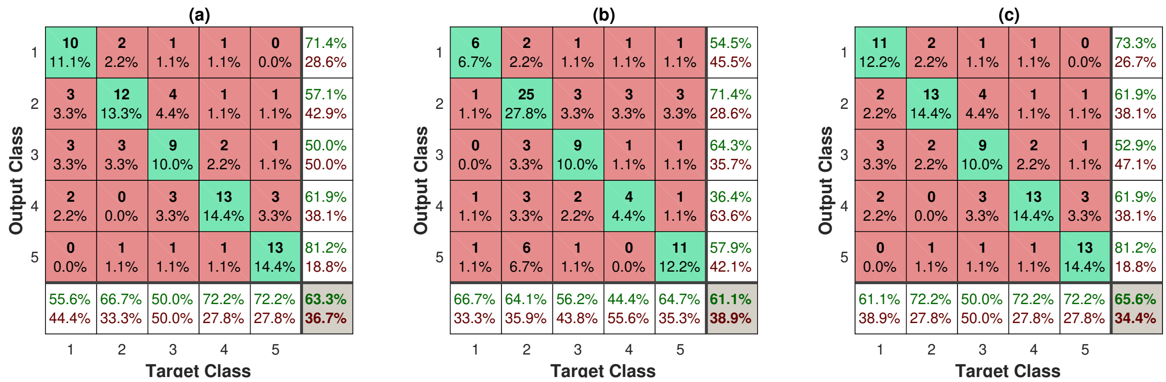
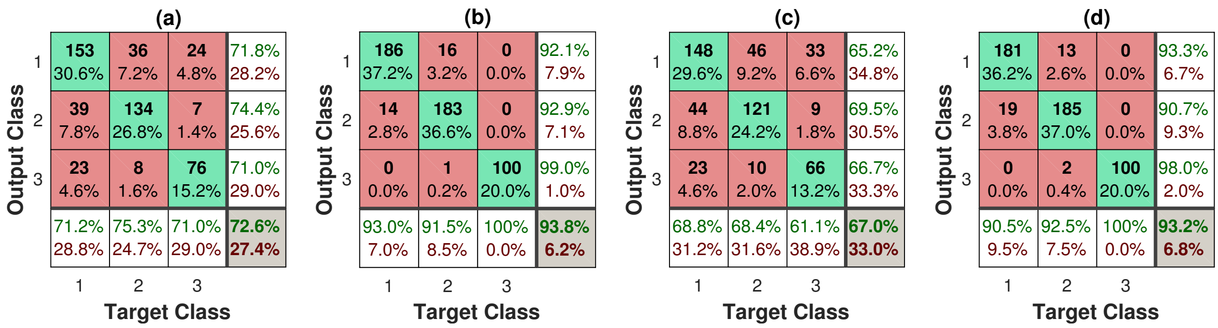


Figure 8.1: The confusion matrix for the E4 dataset. (a) using objective model [40], (b) subjective model [124], and (c) stacking result. The classes are: 1–neutral, 2–cheer, 3–sadness, 4–erotic, and 5–horror [1].

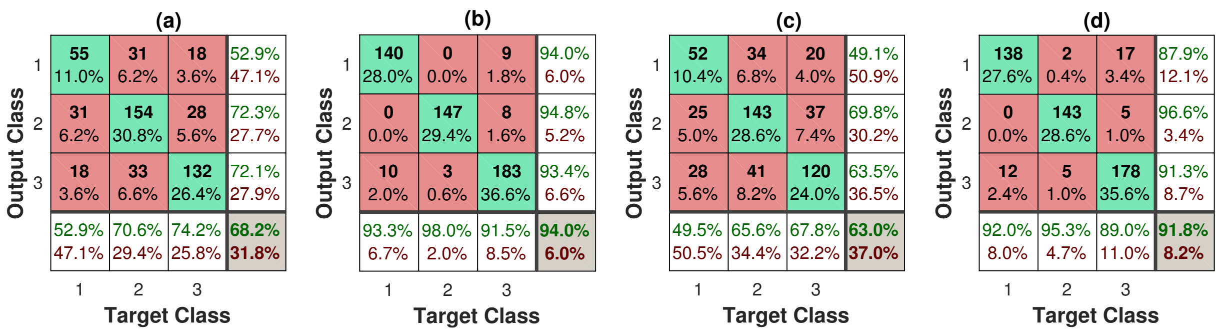
		Parameters				% Accuracy		
DTW Algorithms	Signal	Max Warping	Delay	Frame Size	Weight	F1-score	Precision	Recall
MW-DTW using objective model [40]	EDA	29	34	44	1.00	63.44	64.35	63.33
	HR	89	4	44	0.53			
	Temp	39	4	44	0.24			
MW-DTW using subjective model [124]	EDA	34	68	10	0.55	62.46	63.12	62.22
	HR	5	68	10	1.00			
	Temp	57	48	30	0.33			

Table 8.1: A summary of the comparative results of the various models using the E4 dataset.

From this analysis, it is clear that the performance varies slightly from one model to another. Table 8.1 shows the summary results using the E4 dataset. It includes the optimized parameters for each signal, such as the max warping window, delay time, frame size, and weight parameters. Each sensor makes an independent report based on its weight, as reported using the optimization process. The EDA signal has a significant influence in the first model compared to the second model. However, the HR signal shows a higher weight in the second model compared to the first model. In Figure 8.1, the basic model works by finding a discriminate pattern that can distinguish between five different emotion classes: neutral, cheer, sadness, erotic, and horror. This model performs the best for the horror and neutral emotions but lacks accuracy for the sad emotion. The main disadvantage of this model is user perception deference. To overcome this issue, the diminution model uses self-report [29]. This model shows adequate performance for the sad and cheer emotions and deficient achievement for the erotic emotion. Since self-report can measure only the perception of emotional reaction after the end of a stimulus, the emotional measurement may be affected due to many reasons. During the experiment, some participants were confused about whether they felt happy or neutral. Another limitation is that participants may be unwilling to report their emotion because of shame or concerns about social desirability. This was very obvious in the erotic emotion. There was an overall accuracy of 63.3% for the first base model and an accuracy of 61.1% in the second base model. The WMD-DTW of the objective model presented a slightly better performance than the subjective model. In Figure 8.1, The stacking WMD-DTW outperformed both models, as the average accuracy was 65.6% for all users.



(a) The confusion matrix for arousal; the classes are 1–calm, 2–medium and 3–activated.



(b) The confusion matrix for valence; the classes are 1–unpleasant, 2–neutral and 3–pleasant.

Figure 8.2: The confusion matrix for arousal using the MAHNOB dataset, (a) MW-DTW using 9 felt emotion keywords without padding, (b) MW-DTW using stimuli labeling without padding, (c) MW-DTW using 9 felt emotion keywords with padding, and (d) MW-DTW using stimuli labeling with padding [1].

DTW Algorithms		F1-score	Precision	Recall	F1-score	Precision	Recall
		with padding			without padding		
MW-DTW (using 9 emotion Keywords)	Arousal	67.48	70.35	67.06	71.00	71.00	71.09
	Valence	63.15	68.55	63.18	65.50	65.62	65.60
MW-DTW (using stimuli labeling)	Arousal	93.04	93.46	93.06	93.76	93.21	93.20
	Valence	90.84	91.82	90.74	93.79	93.80	93.80
Arousal stacking result					**94.00**	**95.04**	**93.05**
Valence stacking result					**93.69**	**93.78**	**93.68**

Table 8.2: A summary of the comparative results of the various models using the MAHNOB dataset.

The second experimental result was obtained using the MAHNOB dataset. Each trial contained 30 seconds before the beginning of the effective stimuli and another 30 seconds after the end. We used the first 30 seconds for the normalization step; the remaining parts were used in the classification models. Since the 20 emotional videos were of varied length, a percentage delay and a frame size were selected based on the following equation: $F_s + D_t = S_l + \alpha$, where (F_s) is the frame size, (D_t) is the delay time, (S_l) is the stimuli length, and α is a percentage of the stimuli length. Moreover, due to the variation of the length of the emotional videos, the segmentation resulted in unequal frame sizes. Therefore, we attempted two different techniques to calculate the WMD-DTW. The first technique used the exact length without any padding. The second technique used the padding technique to achieve the same frame size for the entire dataset. In the time series analysis, for different length input sequences, it is a common practice to use zero padding such that each sequence has the same length [166]. Padding can be applied to either the beginning or at the end of the sequences. We used post-sequence padding that was applied to the end of each sequence. We presented the emotional states in three defined classes. The classes were "calm," "medium" and "activated" for arousal and "unpleasant," "neutral" and "pleasant" for valence.The classes were defined using nine emotional keywords (happiness, amusement, neutral, anger, fear, surprise, anxiety, disgust and sadness) as in [83] and [84]. Figure 8.2 shows that the results of both models were better without the use of the padding technique.

In Table 8.2, we note that the WMD-DTW of the objective model presented a significantly better performance than the subjective model. We achieved 71.00% for arousal

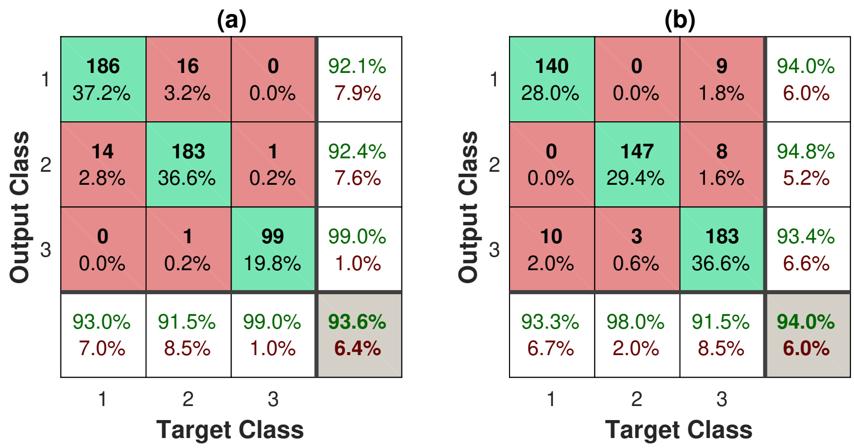


Figure 8.3: The stacking confusion matrix for the MAHNOB dataset, (a) valence and (b) arousal [1].

Model	Accuracy	Precision	Recall	F-Measure
SVM (Linear)	0.34	0.47	0.34	0.37
SVM (Poly)	0.36	0.70	0.37	0.42
SVM (rbf)	0.41	0.53	0.42	0.45
Random Forest	0.64	0.65	0.65	0.65
Naive Bayes	0.27	0.43	0.27	0.33
K-NN	0.72	0.73	0.73	0.72
CNN	0.78	0.78	0.78	0.78
Proposed WMD(DTW)	0.93	0.94	0.93	0.93

Table 8.3: Performance metrics using MAHNOB dataset compared with the reported accuracy by Al Machot et al. [3]

and 65.50% for valence for the first base model. We also achieved 93.76% for arousal and 93.79% for valence for the second base model, which we consider very promising compared to related studies. Our proposed system achieved overall accuracies of 94.0% and 93.6% for recognizing valence and arousal emotions, respectively. In the comparative study, these values outperform the two related works: 46.2% for arousal and 45.5% for valence [83] and 59.57% for arousal and 57.44% for valence [84]. Table 8.3 shows the performance metrics using MAHNOB dataset compared with the reported accuracy by Al Machot et al. [3]. The analysis of the results of the state-of-art, for subject-independent emotion detection, using proposed model does lead to higher performance.

The MAHNOB dataset shows higher improvement for two reasons. First, the data were obtained using different methods, such as the user numbers, the hardware sensors, and the emotion induction methods. Second, the emotion recognition systems for the MAHNOB dataset map the two dimensions into three classes within arousal-valence areas. The classes are "calm," "medium" and "activated" for arousal and "unpleasant," "neutral" and "pleasant" for valence. It is not necessary to map dimensional spaces into a specific emotion. For example, fear and anger share unpleasant and activated aroused feelings but differ in the external causes and behavioral reactions, yet the model cannot capture the difference between them. However, the E4 dataset maps the dimensional spaces into a specific emotion despite the fact that they share the same characteristics. For example, the neutral and sad emotions share the same calm arousal, and sadness and fear share the same unpleasant valence. The complexity is higher for the E4 dataset than for the second dataset, resulting in the accuracy reduction.

Visual inspection of the time series recorded during the experiment shows that the EDA performance differs from one user to another. We identify three groups of subjects with similar emotion-relevant physiological response patterns. The first group, the standard group, shows a strong relationship between the arousal level and the phasic driver peak amplitude of the EDA signal, as we expect [167]. A small portion of the general population shows changes in the phasic peak amplitude of the EDA signal. These minimal changes can be due to many factors, such as medications that suppress the sympathetic nervous system response. Another factor is user personality; one subject from the second group indicated that she does not show emotion and typically refers to herself as "a cold person," and another subject also mentioned that he hardly ever becomes emotional. The third group, however, shows significant changes in the driver peak amplitude of the EDA.

For the MAHNOB dataset, the result in Figures 8.2 and 8.3 suggest that most of the individual differences were based on the dynamic nature of the emotion. As a result, the objective model (stimuli labeling model) shows a significant improvement over the subjective model. The result proves that the stimulus videos in the MAHNOB dataset were more influential in evoking the emotion. This result also suggests that the self-report may have some bias; as a result of the subjective model (9 emotion keywords model), there is some insignificant improvement over the objective model. The average agreement using the multilayer Cohen's kappa test was $k = 0.32$ for arousal and valence ratings. The correlation values were $m = 0.40(SD = 0.26)$ for arousal and $m = 0.71(SD = 0.12)$ for valence [83]. In Figure 8.3, the objective model and subjective model within the ensembling process outperformed any individual model and performed at least as well as the best of the base models. This outcome will most likely depend on the size and quality of the dataset used.

The hyperparameter tuning proved to be the vital step and had a large impact on the end performance of the machine learning models. The data parameter tuning proved to have a large impact on the end performance of the Ekman models. Notably, the optimizer gives more weight to EDA for the first base model. However, the second base model gives greater weight to HR. The third reason for the superiority of the achieved accuracies is the use of the ensemble approach, which allows multiple emotion models and learning algorithms to be jointly embedded within a metalearning model. The staking result captures the individual differences in both emotional perception and the component architecture. As seen in the experiment, the stimulus interpretation is heavily dependent on the user, which appeared obvious when one of the users felt excited and overjoyed when seeing a horror clip.

8.2 Live Biofeedback

Figure 8.4 shows the visual feedback for each emotion, including (a) natural, (b) happy, (c) sad, (d) love, and (e) fear emotions. The intensity of the emotion is shown in a circular progress view with different colors that match the emotion wheel [118], including light gray for natural, yellow for happy, purple for sad, green for love and blue for fear.

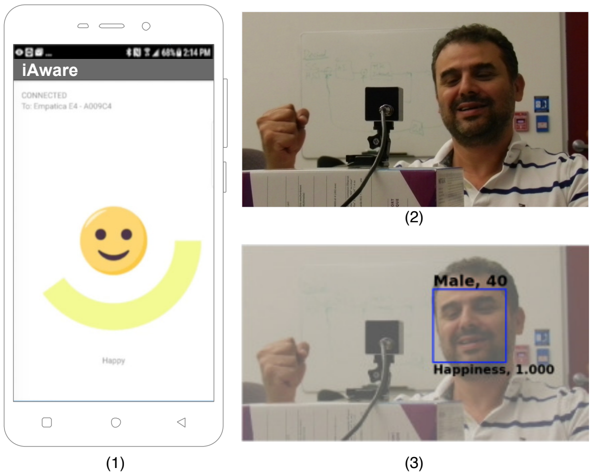
To perform statistical analysis in a natural setting, we used the 2018 FIFA World Cup analysis as a benchmark to compare the iAware results. Therefore, in the following, we analyzed one subject for 45 minutes while watching the Argentina vs. Nigeria match during the 2018 FIFA World Cup. Figure 8.5 shows video images captured for emotion recognition using the Microsoft Azure API. Figure 8.6 shows the results of the analysis

Figure 8.4: iAware visual feedback for (a) natural, (b) happy, (c) sad, (d) love, and (e) fear emotions

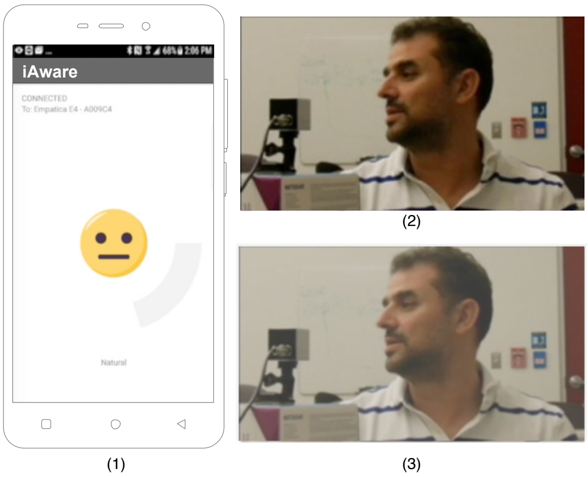
comparing iAware with the half-time statistical analysis.

Figure 8.5 shows example results of the analysis comparing video detection with iAware detection. In general, iAware performs similarly to video detection; however, in some cases, iAware shows substantially better results than the model based on video detection in natural settings, allowing the user more mobility and freedom. The Azure API was not able to analyze some cases: for example, when the user began jumping for joy after the first goal scored by Argentina; when the user cupped his hand over his mouth, which occurred several times during the match; and when the user turned his face and moved extensively during conversations.

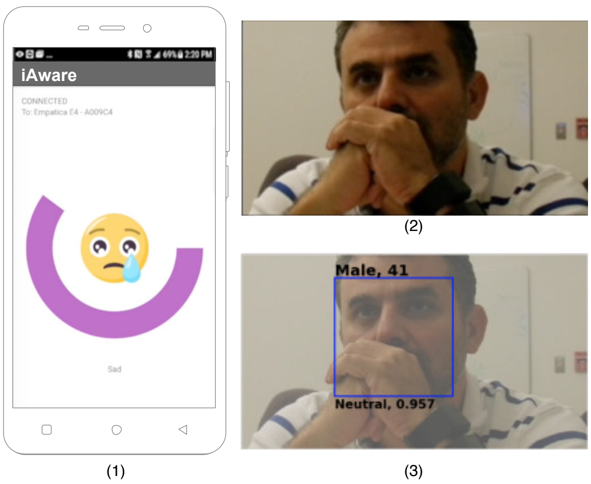
Figure 8.6.a highlights the emotion percentage throughout the match, during which he felt 22.66% neutral, 44% sad, 9.33% fear, 22.66% joy, and 1.33% love. Figure 8.6.b shows the emotion heat map based on ball possession using [168]. For Argentina, the user was 37% happy each time that the team attacked and 4% neutral. He also felt 36% neutral most of the time during the midfield battle. There was one case (2%) during which the user felt love. This case occurred during the time when he saw his favorite player, Mr. Diego Maradona, on TV. In contrast, the user felt 17% sad when Argentina was defending and 4% fear when Nigeria almost scored a goal. For Nigeria, the user felt 12% fear each time the players almost scored and 21% sad when they attacked. He then felt neutral 33% during the time the ball was in the midfield, He felt 2% happy when the defender missed the ball, allowing Argentina to take a good



(a) Stimuli: Argentina scored the first goal

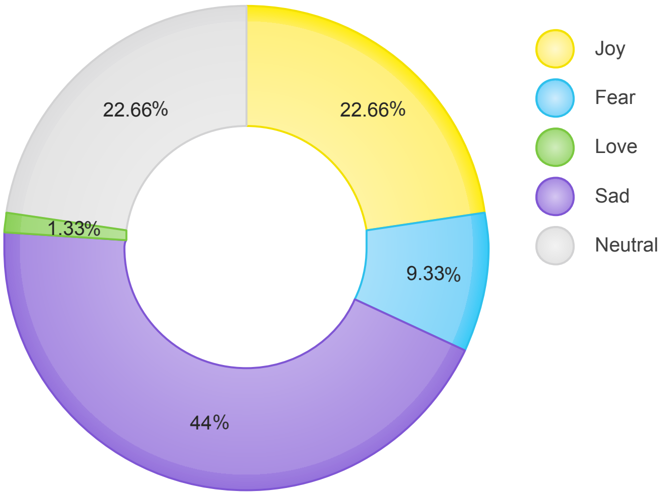


(b) Stimuli: Conversations with a friend

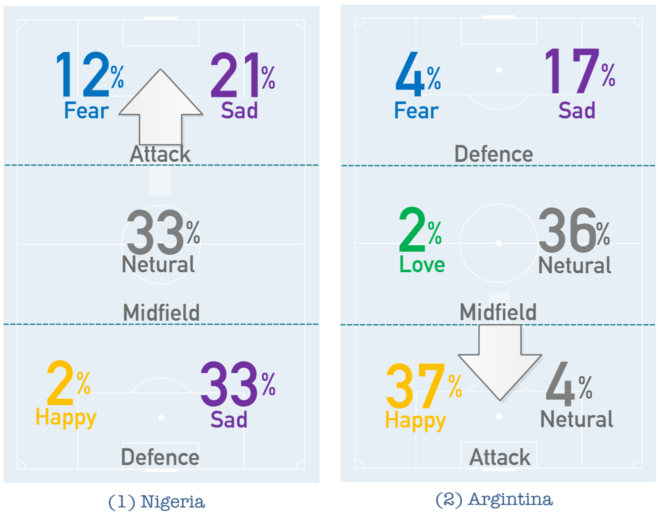


(c) Stimuli: Argentina's player injured

Figure 8.5: Example of the comparative results between iAware and Azure: (1) the iAware, (2) the original video, and (3) Microsoft Azure



(a) Emotion percentage



(b) Emotion heat map based on the ball possession

Figure 8.6: Half-time statistic of the soccer match between Argentina and Nigeria; the total may not be 100% due to rounding

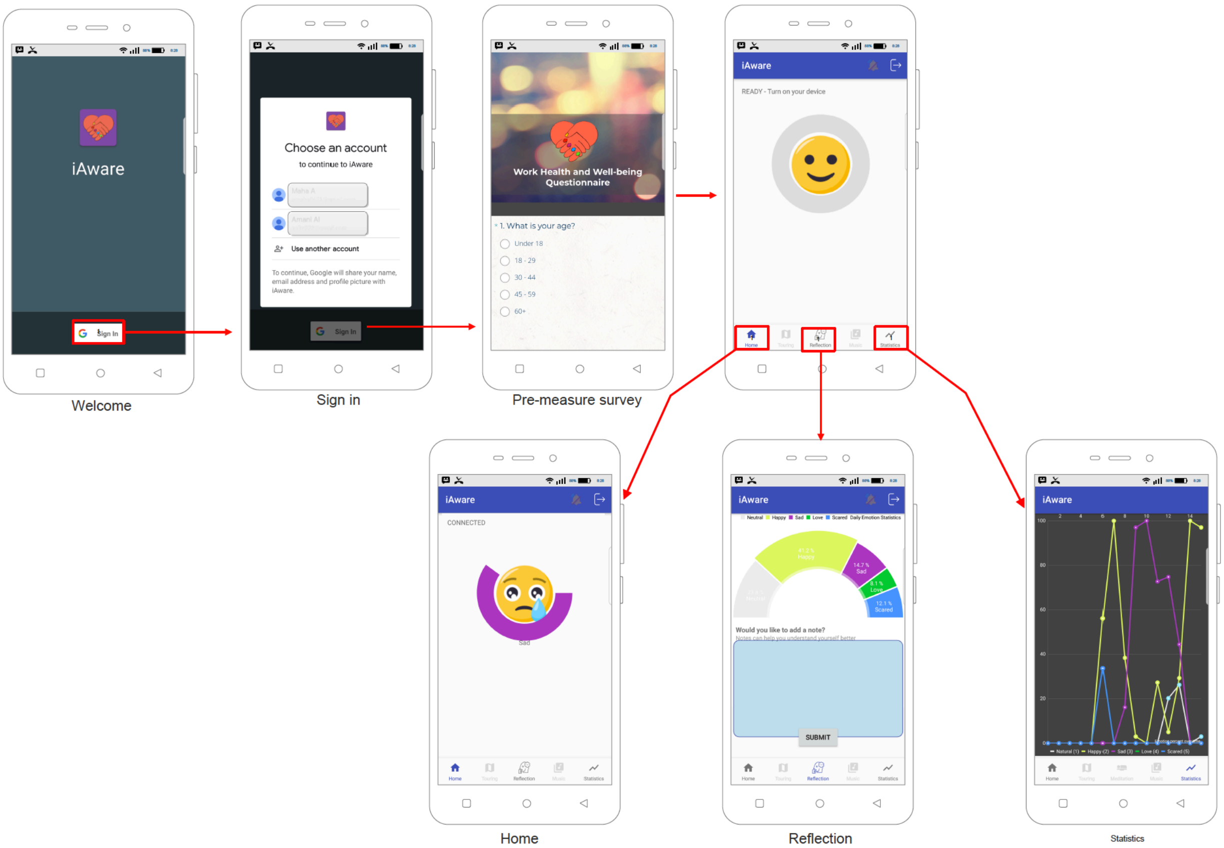


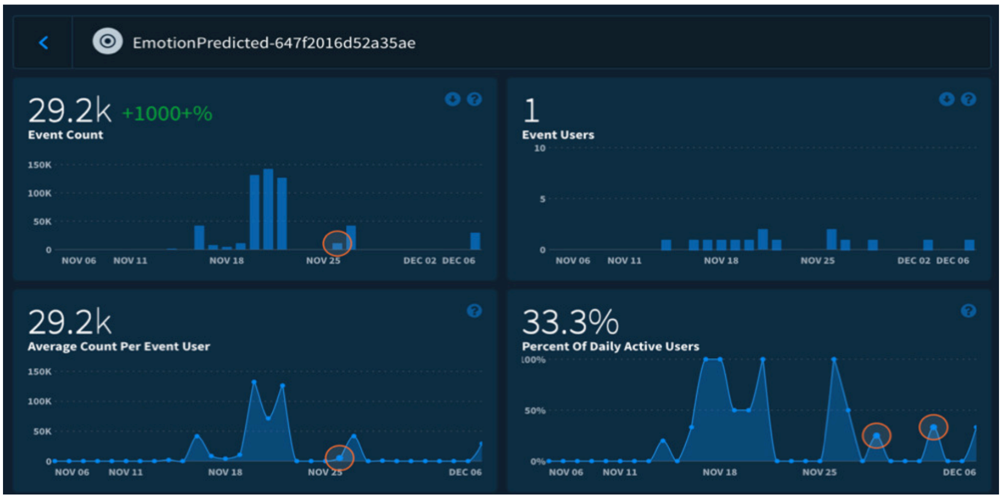
Figure 8.7: Individual mode interface including the wireframes diagram

shot; however, he felt sad most of the time (approximately 33%) during defending, as Argentina primarily lost good opportunities. This result suggests that iAware could be a useful tool for detecting emotion in a noncontrolled environment.

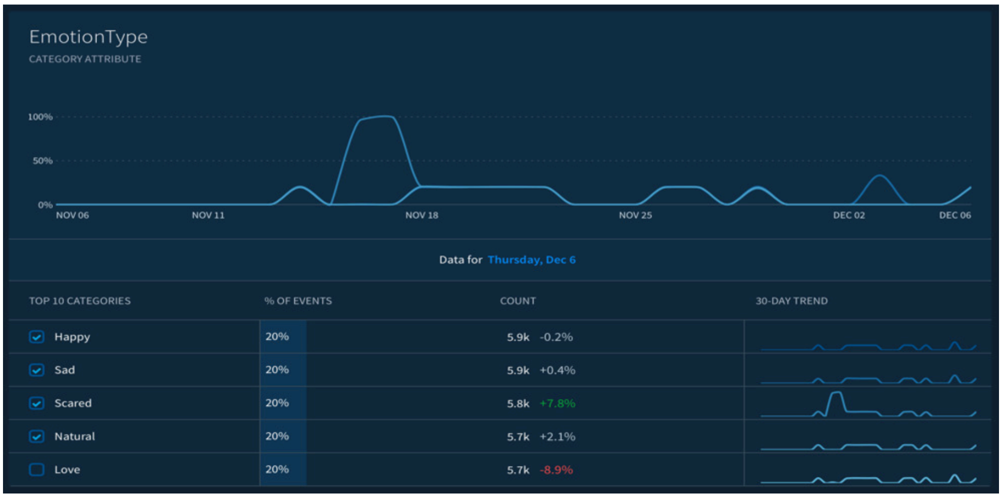
8.3 Intervention

Figure 8.7 shows the individual user interface wireframes diagram. Figure 8.8 and 8.9 show the diagnostic mode dashboards of the APP for all the organization users.

The individual mode includes system login, premeasure stress survey, biofeedback and stress management recommendation. Self-reflection users can write a daily reflection to summarize their work and thoughts on the current treatment day while visualizing the daily emotion percentages for that day. In addition, the user can view a real-time statistical visualization of the emotions; see Figure 8.7.

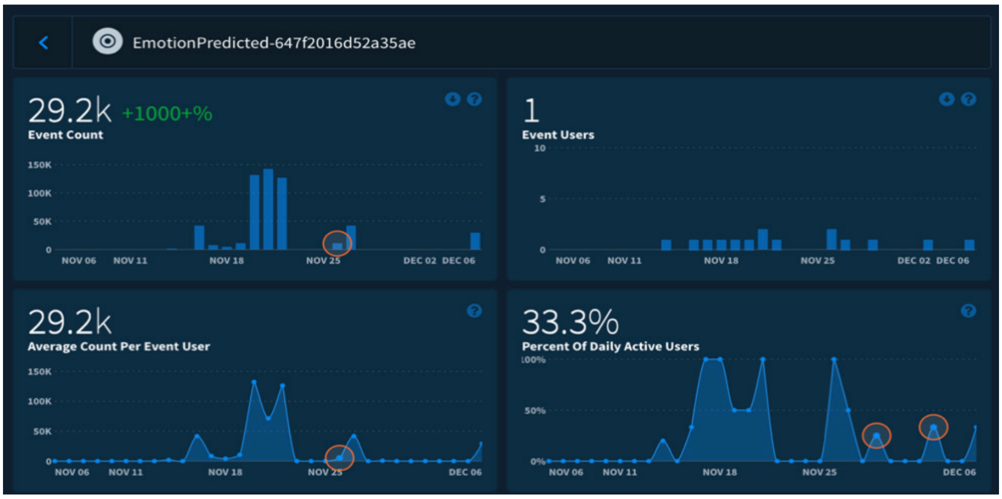


(a) View1: over all diagnostic for one user

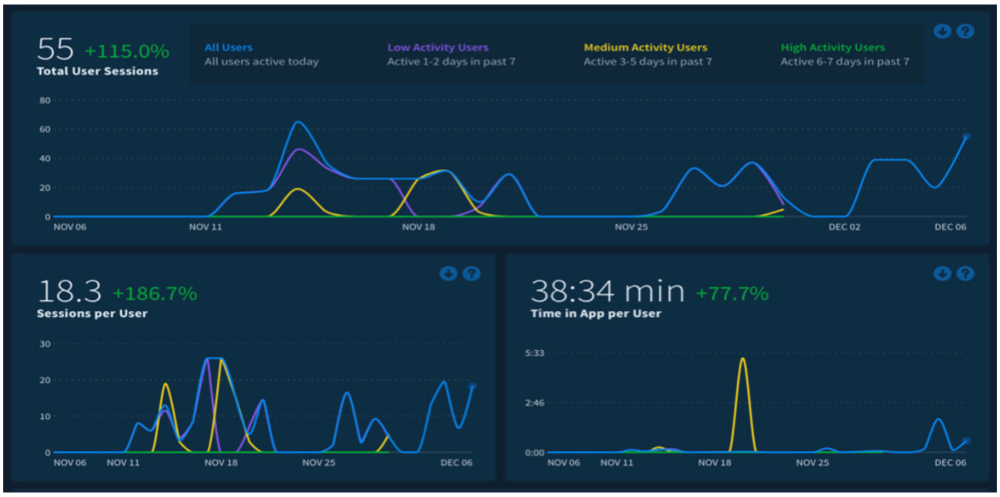


(b) View2: emotion events percentage for one user

Figure 8.8: Diagnostic mode dashboards of the APP for one user



(a) View1: overall diagnostic for all the users



(b) View2: breakdown diagnostic for all users

Figure 8.9: Diagnostic mode dashboards of the APP for all the organization users

Figure 8.8 presents the diagnostic mode dashboards of the APP for one user, while Figure 8.9 presents historical information for all users. At iAware, we used Fabric to monitor a whole host of information using an activity streaming mode in real time including the number of persistent connections maintained and the number of emotional events per user to identify the cause of a problem. A stream of sequential activity for a diagnostic plan was created using historical information.

8.4 Concluding Remarks

In this chapter, we proposed and evaluated the inHarmony system, including the emotion recognition and the biofeedback system based on physiological signals.

Overall, it can be seen that the discrimination among physiological patterns was improved using the ensembling process. This superiority is attributed to three reasons. First, the optimization methods help select the best parameters for the signals to have the most significant information possible. The selection of the delay component reduced the degree of synchronicity that the emotional components display over time, and the frame size address the individual differences in the duration of the participant emotions. The second reason is that the WMD-DTW learned the dynamic nature of the emotion and the individual differences in the variability of emotional intensity over time without the need for group-based models.

In our evaluation, we show that the system can provide real-time monitoring of emotions at any time and in any place, providing the user with more mobility and freedom. The next chapter covers the results of the usability case study on workplace stress management techniques.

Chapter 9

inHarmony Usability Study

Subjective evaluation is an essential component of the assessment of mobile application usability. Psychological test construction methods were applied to ensure the proper construct and empirical validity of the items and to assess their reliability. The chapter presents the usability study of inHarmony, an emotional well-being digital twin. It discusses the study objectives and plan. In addition, it shows how the study was conducted along with the results. In addition, the chapter evaluates the usefulness of the proposed methods, and the primary results suggest the need for such a system.

9.1 Motivation

The main aim of the usability study is to answer the second research question of the thesis: 'What is the usefulness and usability of the proposed solution for emotional well-being?". The following section presents the feasibility assessment with two examples, one for the feedback and the other for comparison between three different intervention modes.

The study has two parts: evaluating the biofeedback effectiveness and usefulness and evaluating the usability of the suggested interventions, using three modes in real practice in order to assess how usable the live biofeedback system is and, more importantly, the chances for inHarmony to be used in practice. The answers to the questions are mostly scored on a Likert-type scale, with a few open-ended questions to collect feedback, comments, and improvements.

9.2 Ethics Approval

Ethics approval was acquired from the University of Ottawa Research Ethics Board for the usability study. The certificate is included in Appendix D: Ethics Approval.

9.3 Evaluation

9.3.1 Live Biofeedback

In our view, the validity of self-reports of emotion is questionable. Here, we follow Brody et al. [169], who concluded that men and women display gender-stereotypical expressions. There are individual differences in awareness of and willingness to report on emotional states, which potentially compromises the emotional experience. Men exhibit restrictive emotionality [170]. Restrictive emotionality refers to a tendency to inhibit the expression of certain emotions and an unwillingness to self-disclose intimate feelings [171]. Moreover, women report more intense emotional experiences and more overt emotional expressions across 37 cultures [172]. We hypothesize that the user will experience an increase in emotion self-awareness while using the SLBF from the inHarmony system by comparing the discordant pairs resulting from the system and the self-report.

9.3.1.1 Hypothesis

The goal of this study was to test our hypotheses that the iAware system can improve self-awareness by measuring the clarity of emotions and attention to emotions of users according to an emotion self-report.

We formulate the null hypothesis that the probabilities are the same or, in simplified terms, that neither of the two models performs better than the other. Thus, the alternative hypothesis is that the performances of the two models are not equal. Therefore, if there is no association between emotion detection based on automatic detection by the system and the user's perceived emotion reported by the self-report methods, what is the probability of observing a significant discrepancy between the two methods based on the discordant pairs?

9.3.1.2 Participants

A total of twenty healthy participants, ten males and ten females, took part in this experiment. This sample size is recommended as an appropriate size for quantitative usability studies [173]. Their ages ranged from 18 to 58 years old. None of the participants had experienced symptoms of excessive sweating (hyperhidrosis) or hypokalemia (bradycardia or tachycardia) or had a known history of heart disease. None were experiencing any mental health problems or taking antianxiety or antidepressant medications. All participants were informed about the procedures and potential risks before beginning the experiment.

9.3.1.3 Noncontrolled Experiment

Experimental setup We evaluated the users' self-experience in a natural setting for 20 participants. While the participants were using the application and data were being collected, each participant was asked to report their feeling in the moment, first without feedback and then with the real-time feedback. Finally, they were asked to what extent the application was able to explain their inner feelings and represent their emotions.

9.3.1.4 Controlled Experiment

Experimental setup Figure 9.1 shows the experimental procedures. For each participant, the experiment involved a session lasting approximately $15min$. The experiment was designed with the following steps: an initial phase, emotion detection without biofeedback, and emotion detection with biofeedback. First, the researchers explained the experimental procedures and any risks associated with participation. Before the trial was initiated, each participant was required to wear the E4 wristband on his or her nondominant hand. Then, the participant completed a consent form. After that, benchmark data were collected. This step was optimized to last approximately $110s$, according to previous work, and was followed by a sufficient visual stimulation period

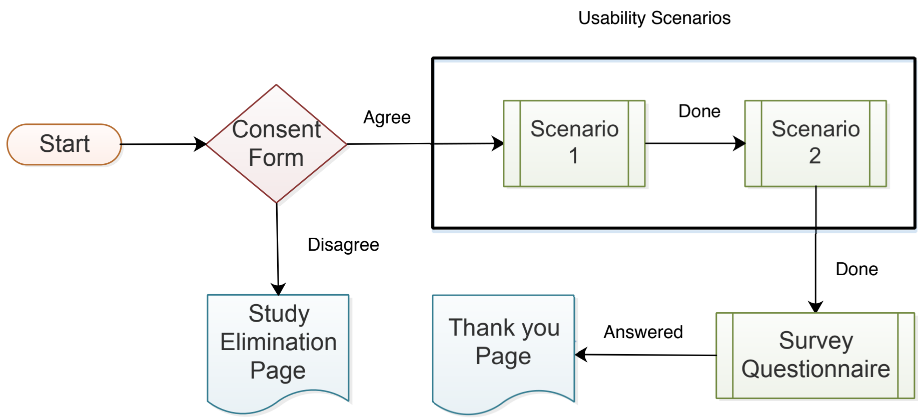


Figure 9.1: The controlled experiment procedures

lasting 40s for each targeted emotion and the self-report using AniAvatare [123].

For the stimulus material, we chose EMDB [39] because videos contain more emotional content than a single image. We randomly selected two clips from the three categories of happiness, sadness, and fear. We used movie clips 4007 and 4009 for the happiness condition, 3009 and 3008 for the sadness condition, and 1007 and 1008 for the fear condition. All clips had sound added to maximize the emotional experience. We randomly chose one of these clips for each scenario.

This test has two scenarios while watching the emotional movie clips: detection of emotion without any feedback and detection of emotion with feedback using the system. Each participant was asked to watch one 40 s movie clip in each of the three categories, sad, happy, and afraid. After each clip, they reported their emotion using AniAvatare [123] and then completed a satisfaction survey. The questionnaire serves to evaluate the system and capture the user's feedback after using the system.

9.3.1.5 Evaluation Method

For the statistical analysis, we used the nonparametric method. We used the McNemar test, also referred to as the within-subjects chi-squared test, for two reasons. First, the McNemar test is a useful tool for comparing two different models to test the significance of the difference between two paired results of matched individuals. Second, the test is recommended for small sample sizes [174]. We study the effect with respect to the responses of gender groups and emotion types for cases with and without using the system. The P value was calculated using a 2×2 contingency table from the McNemar test with the continuity correction [175]. The McNemar test approximates the binomial exact P value using χ^2, as in equation (9.1):

$$\chi^2 = \frac{(|B - C| - 1)^2}{B + C} \tag{9.1}$$

where B is the total cases where the self-report fails to detect the emotion but the system passes and C is the total cases where the system fails to detect the emotion but the self-report passes for the same person.

We also checked the plausibility of the system using an offline textual questionnaire organized into two sections, with the first section including multiple-choice questions and the second section including open-ended questions. In the first section, all six questions were answered on a discrete scale of 1-5, where 1 denotes the lowest plausibility and 5 denotes the highest plausibility.

9.3.2 Randomized Controlled Trial

9.3.2.1 Hypothesis

A randomized controlled trial (RCT), which randomly assigns participants to experimental groups without revealing the assignment information, was used as an evaluation method. The usability study provides us with quantitative measures extracted from usage logs and qualitative findings from postquestionnaires that can help us draw a conclusion regarding how the proposed strategies affect personal emotional well-being.

We formulate the null hypothesis that the satisfaction level is the same or, in simplified terms, that neither of the three models performs better than the other, which proves the alternative hypothesis that the performances of the models are not equal.

9.3.2.2 Experimental Setup

For evaluation purposes, we developed three versions of the application. Version A is the baseline silent mode, which collects the users' data, provides emotional feedback and sends all the information to the cloud without providing the user with intervention.

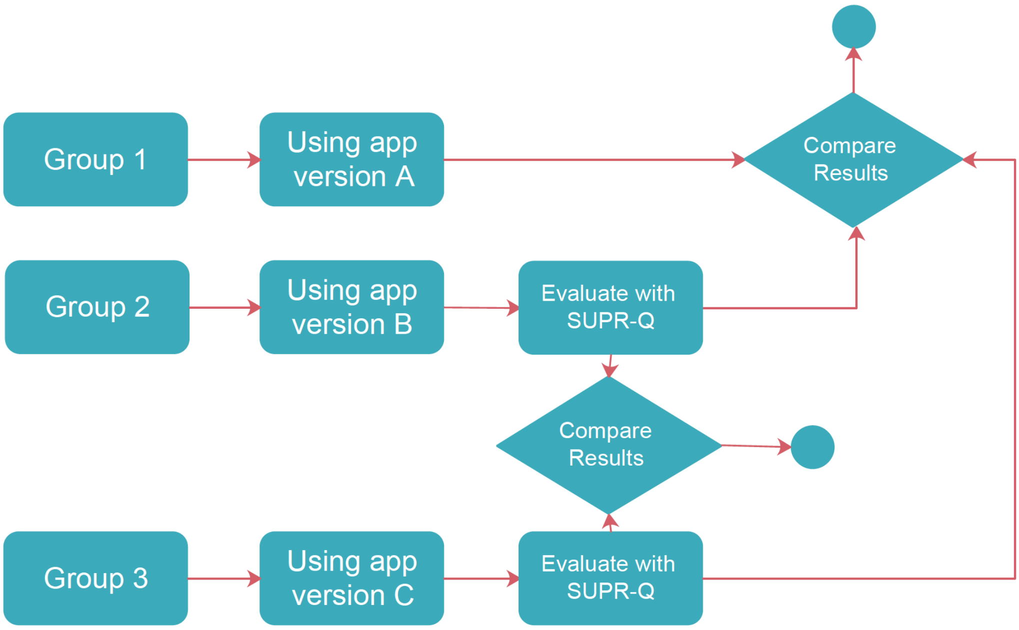


Figure 9.2: The experimental procedures

Version B is the action card mode, which adds more features, such as microintervention and self-monitoring. Finally, version C is the stress feedback mode, which uses user feedback and input about the situation in some cases. The test subjects were randomly divided into three groups. We asked one-third of the users, or the first sample group, to use version A. The second and third sample groups were asked to use version B and version C of the application, respectively. The participants then took the standardized user experience percentile rank questionnaire (SUPR-Q) after using the assigned version.

Group 1: Baseline silent mode, Group 2: Action card mode, and Group 3: Stress feedback mode

9.3.2.3 postquestionnaires

Among the popular questionnaires used to measure the perceptions of usability are website questionnaires such as the WAMMI [176] and the SUPR-Q [177]; both provide relative rankings expressed as percentages. Other questionnaires include the software usability measurement inventory (SUMI) [178], which is a lengthy 50-question survey.

The computer system usability questionnaire (CSUQ) [179] is composed of 19 questions. In contrast, the system usability scale (SUS) [180] consists of a 10 item questionnaire that measures the perceptions of usability. The SUS was developed over 20 years ago and provides a reliable estimate of usability. Its questions are generic enough to use on a wide variety of software and hardware applications. Both the CSUQ and SUS are overall satisfaction questionnaires. However, a well-standardized questionnaire provides both a valid and reliable instrument, as well as a relative ranking based on a dataset. Although some of the existing surveys share some of these aspects, most notably, the generalizable and multidimensional elements, none contain all four (i.e., brief, generalizable, covering multiple constructs including trust, and a normative database).

In order to study the satisfaction level and compare the three different intervention modes, we adapted the SUPR-Q. The SUPR-Q uses both overall satisfaction and a hierarchical approach divided into subcomponents consisting of independent psychometric scales, e.g., usability, credibility/trust, loyalty, and appearance [177]. The SUPR-Q is an 8 item questionnaire that measures the quality of the website user experience. The SUPR-Q provides measures of usability, credibility/trust, loyalty and appearance.

We had to modify the questions slightly, in some places, to suit the nature of the unit under test. The term "website and web page" was replaced by "application or feedback". In addition, a good questionnaire is a mixture of closed- and open-ended questions. Open-ended questions are valuable tools for identifying a range of possible responses and give the respondent an opportunity to state their opinions about a topic [181]. Thus, we added two open-ended questions: "What are two things about the application that you really liked?" followed by "What are two things about the application that you did not like?” The questions of the survey were also checked by experts in the usability domain to ensure reliability and suitability. Both surveys are included in Appendix B: Evaluating the usability of inHarmony.

9.3.2.4 Evaluation Method

For the statistical analysis, we used repeated-measures multivariate analysis of variance (rMANOVA). This approach is also known as multivariate growth-curve analysis. The advantage of using rMANOVA is the ability to detect treatment differences between different groups [182].

We also compare the plausibility of the system using an offline textual questionnaire organized into two sections, with the first section including multiple-choice questions and the second section including open-ended questions. In the first section, all six questions were answered on a discrete scale of 1-5, where 1 denotes the lowest plausibility and 5 denotes the highest plausibility.

9.4 Results and Discussion

9.4.1 Life Biofeedback

9.4.1.1 Noncontrolled Experiment

We evaluated 20 participants who felt emotions in a natural setting. Table 9.1 shows the results per subject for a noncontrolled experiment without feedback. Then, we compared these results with the results obtained after the participants saw the feedback.

Table 9.1 shows the results per subject. There were five discordant pairs for emotion detection between the system and the self-report. After receiving the feedback, participants 4, 13, and 15 reported that they were more nervous during the experiment. By contrast, participant 20 reported feeling sad, which was not expected to be caught by the application. Moreover, participant 9 stated that she had more overt emotional expressions, as she was feeling positive neutral. This evidence may support the gender-stereotypical expression of emotions; therefore, the SLBF can potentially increase self-emotional awareness. Participants 4, 13, 15 and 20 are examples of inhibiting the expression of fear and sad emotions, while participant 9 is a classic example of women reporting more intense emotional experiences.

Subject#	Gender	Felt emotion	App Detecting
1	M	2	2
2	M	2	2
3	M	1	1
4	M	1	5
5	M	2	2
6	M	1	1
7	F	1	1
8	M	1	1
9	F	2	1
10	F	1	1
11	F	5	5
12	M	2	2
13	M	1	5
14	M	2	2
15	F	1	5
16	F	1	1
17	F	2	2
18	F	2	2
19	F	1	1
20	F	1	3

Table 9.1: Results per subject for a noncontrolled experiment without feedback; The emotions are (1) neutral, (2) happiness, (3) sadness, (4) love, and (5) fear.

9.4.1.2 Controlled Experiment

We studied 120 cases for a total of 20 subjects. Figure 9.3 shows the results of the analysis using the McNemar test. Table 9.2 and Figure 9.4 shows the results of the analysis using the post-study questionnaire. A 2×2 matrix with factors of gender (male and female) and reporting tool (system and self-report) was developed to report the results for the different emotions (happy, sad, and afraid).

For the first case scenario without using the SLBF, there were 12 discordant pairs for emotion detection. There were 10 (83.333%) pairs for which the system was able to detect emotion correctly but the self-report was not and 2 (16.667%) pairs for which the self-report was correct but the system was not. The two-tailed P value was 0.0433, the Chi-squared statistic was 4.083 with 1 degree of freedom, and the odds ratio was 0.2000 with a 95% confidence interval of 0.021 to 0.939.

For the second case using the SLBF, there were six discordant pairs for emotion detection. There were 5 (83.333%) pairs for which the system was able to detect emotion correctly but the self-report was not and 1 (16.667%) pair for which the self-report was correct but the system was not. The two-tailed P-value was 0.2207, the Chi-squared statistic was 1.500 with 1 degree of freedom, and the odds ratio was 0.200 with a 95% confidence interval of 0.004 to 1.787.

By conventional criteria, in the first scenario, since the P-value is smaller than our assumed significance threshold ($\alpha = 0.05$), we reject the null hypothesis and assume that there is a significant difference between the two predictive models. In the second scenario, since the P-value is larger than our assumed significance threshold ($\alpha = 0.05$), we cannot reject the null hypothesis and assume that there is no significant difference between the two predictive models.

Moreover, McNemar's test provided further insight regarding model selection. We are interested in the two cases for which the results from both the system and the self-report were in agreement. Figure 9.3 graphically summarizes the analysis of the gender factor and effect type of different emotions. Studying the single emotions revealed that there was a noticeable 3.333% reduction in the discordant pairs resulting from the self-report bias.

Analysis of the gender factor revealed a result similar to [172] that some female participants report more intense emotional experiences for happiness. In addition, some males exhibit restrictive emotionality, particularly for feelings of sadness and fear, similar to [170]. Overall, iAware SLBF can potentially increase awareness related to gender-stereotypical expression, reducing the predictive error by 3.333% for female participants and by 16.673% for male participants. Analysis of the emotion signals showed similar findings to those reported in [170]. Female participants showed no significant difference with and without SLBF for the sad and afraid emotions, while male participants showed no significant difference for happy feelings. In general, participants performed substantially better using iAware than in the model without SLBF.

In the poststudy questionnaire regarding participants' satisfaction, the first six

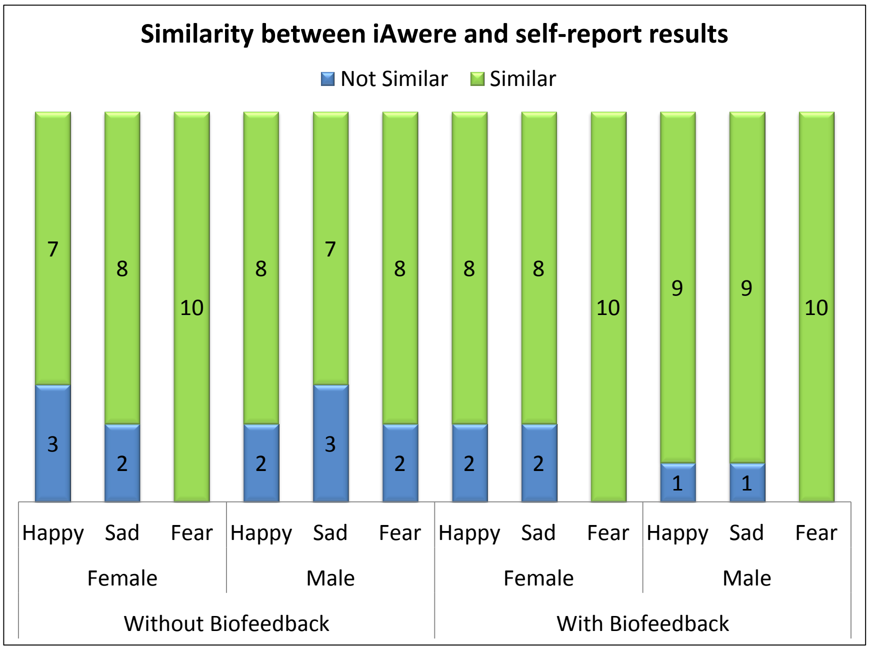


Figure 9.3: Similarity between the system and self-report results

Question #	Question	Average
1	The mobile feedback feature is effective	4.20
2	The app was able to detect my current feelings most of the time	4.05
3	The app feedback is helping me become self-aware of my feelings	4.10
4	Knowing my emotions is helping me to focus on regulating my emotions	4.20
5	I would use this app on a daily basis	3.90
6	I would recommend this app to my friends and family members	4.10

Table 9.2: Average results for the postexperiment questionnaire

questions were based on a Likert scale and prompted participants to specify their opinions on various aspects of the interactions by selecting one of five options. Table 9.2 presents the questions and the average results for the post-experiment questionnaire. The average plausibility score for all questions was 4.05 of 5. Figure 9.4 summarizes the results of the post-experiment questionnaires. Interestingly, 18 participants reported that they would use the system every day, and 17 participants said that they would recommend it to friends and family members.

For the open-ended questions, we asked users to describe in their own words whether they would like to know their current emotion. Interestingly, when asked this question,

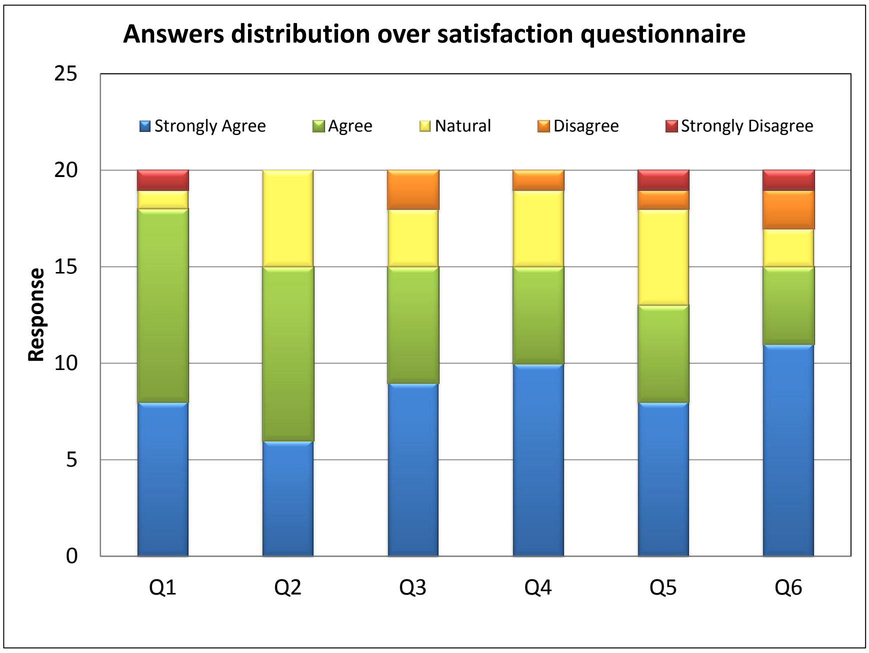


Figure 9.4: Distribution of answers to the satisfaction questionnaire

some users associated the question with increased awareness, for example, "*I would like to know how to control my emotions so that I don't overreact*", "*Sometimes my emotions are unclear, and using this system can help*", "*Yes, to control myself and to know I have to make a decision*", "*It helps better regulate my feelings*", and "*It helps me know more about myself*". One particular participant stated that "*Sometimes, I would like to know how exactly I feel, but sometimes, the system can make me stressed*". The results indicate that SLBF resulting from the system appears to positively affect self-awareness.

We also asked users whether they would like other people to know about their current emotion. Participants reported mixed feelings about this question. Some users answered "yes" in order to show a self-awareness connection with others to seek their support; for example, participants reported "*Only for asking for help*", "*Only if they are going to help*", and "*Yes, so that others can know when something makes me happy or sad and so that they will respond accordingly"*". The participants also highlighted the importance of context, especially social context, for example, "*It depends on who is going to see my emotions*", "*Not all the time; sometimes, I want to keep my emotions* *private*", and "*Only close family members so that they can understand my reaction*". The users' answers to this question suggested that FLBF may help increase emotional awareness for oneself and others; however, the issue of privacy is an essential element that must be considered, given responses such as "*It depends, I guess mostly No, as I prefer to keep my emotions personal*", and "*No, it is private*". People may vary in their ability to perceive and understand other people's emotions, which may affect their ability to recognize and manage social situations. The system is potentially useful in easing interpersonal interactions using FLBF, which is driven by sending and receiving social cues to make it easy to infer these cues. FLBF can also help amplify social cues and increase people's sensitivity towards such cues.

9.4.2 Randomized Controlled Trial (RCT)

Table 9.3 provides a global overview of the demographics and characteristics of the participants.

The results after the use of the application show no significant differences between group x time effects regarding the improvement of positive emotion. The time effect is always highly significant, indicating improvement over time for all subjects from a group who stayed in the study. rMANOVA approaches the level of 0.05 (P=.019). There is also no difference regarding the reduction of negative emotion over time (P=.033). Our observation may be explained by low power, small sample size, and small effect size due to either large variation in the sample or effects that are not substantial. However, figure 9.5, and 9.6 show that group three has more improvement in positive emotion. Moreover, figure 9.7 and 9.8 reduction effect in negative emotion. Perhaps the groups overlap too much, or there are not enough people in each group to establish a significant difference.

9.4.2.1 Perceived Usefulness and Effectiveness

After using inHarmony for a week, participants provided feedback for what they liked most and least about the application and recommendations for improving the application.

	Group 1		Group2		Group3		Over all	
Category	Res. (#)	Res. (%)	Res. (#)	Res. (%)	Res. (#)	Res. (%)	Res. (#)	Res. (%)
Gender								
Male	5	100.0	1	20.0	2	40.0	8.0	53.3
Female	0	0.0	4	80.0	3	60.0	7.0	46.7
Age (years)								
18 - 29	0	0.0	0	0.0	1	20.0	1.0	6.7
30 - 44	2	40.0	0	0.0	0	0.0	2.0	13.3
45 - 59	3	60.0	3	60.0	2	40.0	8.0	53.3
60+	0	0.0	2	40.0	2	40.0	4.0	26.7
Race/Ethnicity								
Asian	1	20.0	1	20.0	1	20.0	3.0	20.0
African American	0	0.0	1	20.0	0	0.0	1.0	6.7
Middle Eastern	2	40.0	1	20.0	2	40.0	5.0	33.3
White / Caucasian	0	0.0	2	40.0	1	20.0	3.0	20.0
European	0	0.0	0	0.0	1	20.0	1.0	6.7
Multiple ethnicity	2	40.0	0	0.0	0	0.0	2.0	13.3
Relationship status								
Married	3	60.0	2	40.0	3	60.0	8.0	53.3
Divorced	0	0.0	1	20.0	1	20.0	2.0	13.3
Cohabiting others	0	0.0	1	20.0	0	0.0	1.0	6.7
Single	2	40.0	1	20.0	1	20.0	4.0	26.7
Education level								
High school	0	0.0	2	40.0	2	40.0	4.0	26.7
Bachelor	4	80.0	2	40.0	1	20.0	7.0	46.7
Master's	1	20.0	1	20.0	1	20.0	3.0	20.0
Doctorate	0	0.0	0	0.0	1	20.0	1.0	6.7
Experience								
<=1 year	2	40.0	1	20.0	1	20.0	4.0	26.7
1 - 2 years	1	20.0	0	0.0	0	0.0	1.0	6.7
3 - 5 years	1	20.0	1	20.0	1	20.0	3.0	20.0
6 - 9 years	1	20.0	0	0.0	0	0.0	1.0	6.7
>= 10 years	0	0.0	3	60.0	3	60.0	6.0	40.0
Weekly work (h)								
25-30	1	20.0	1	20.0	1	20.0	3.0	20.0
31-34	1	20.0	1	20.0	1	20.0	3.0	20.0
35-40	1	20.0	2	40.0	3	60.0	6.0	40.0
>=41	2	40.0	1	20.0	0	0.0	3.0	20.0
Job title								
Student	5	100.0	2	40.0	1	20.0	8.0	53.3
Staff	0	0.0	3	60.0	4	80.0	7.0	46.7

Table 9.3: Characteristics of the participants (n=15)

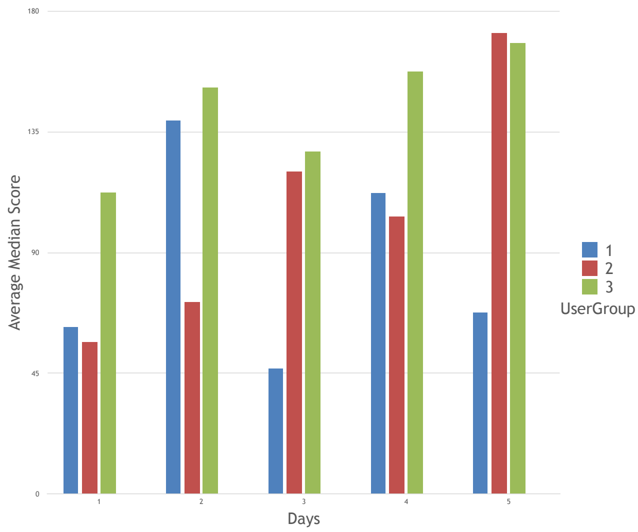


Figure 9.5: The comparative results between the intervention modes for positive emotion: happiness. The three groups use 1) baseline silent mode, 2) action card mode, and 3) stress feedback mode.

Categories	Average level of agreement		
	Group 1	Group 2	Group 3
Usability	4.6	4.3	5.0
Reliability	3.3	3.4	4.2
Trust	3.4	3.3	4.2
Loyalty	4.2	3.5	4.3
Appearance	4.6	4.0	4.5

Table 9.4: Participants' level of agreement for the satisfaction questionnaire using the Likert scale

The 5-point rating scale ranged from 1 (Strongly disagree) to 5 (Strongly agree). Agree ratings are the agree and strongly agree scores combined with mean agreement ratings of $>$ 4.0 considered as the user agrees on the measure quality. We test the usability study for three groups. Group one uses mode one: biofeedback alone mode; group two uses mode two: biofeedback and intervention; and group three uses mode three: biofeedback and intervention taking into account the user feedback. Table 9.4 shows the participants' level of agreement results for the Likert-scale questions.

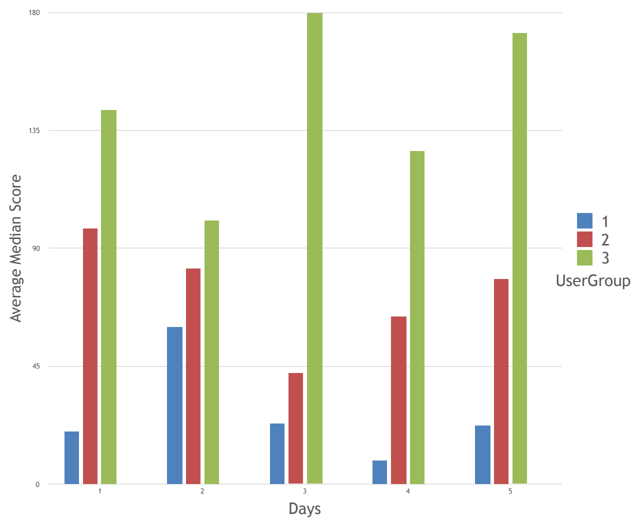


Figure 9.6: The comparative results between the intervention modes for positive emotion: love. The three groups use 1) baseline silent mode, 2) action card mode, and 3) stress feedback mode.

Usability All participant scores for usability measures through all three modes were very high in agreement. All participants agreed that the application was easy to use and easy to navigate within (mean agreement rating = 4.6, 4.3, and 5.0 for modes one, two, and three, respectively).

Reliability For the emotional feedback, all participants found the feedback credible and trustworthy (mean agreement rating = 4.2) for mode three, while participants scored lower agreement ratings for modes one and two (mean agreement rating = 3.3 and 3.4 for modes one and two, respectively).

Trust For the intervention feedback and information, all participants found the feedback credible and trustworthy (mean agreement rating = 4.2) for mode three, while participants scored lower agreement ratings for modes one and two (mean agreement rating = 3.4 and 3.3 for modes one and two, respectively).

Loyalty For the biofeedback alone mode, mean agreement rating = 4.2; for the biofeedback and intervention mode, mean agreement rating = 3.5; and for the biofeed-

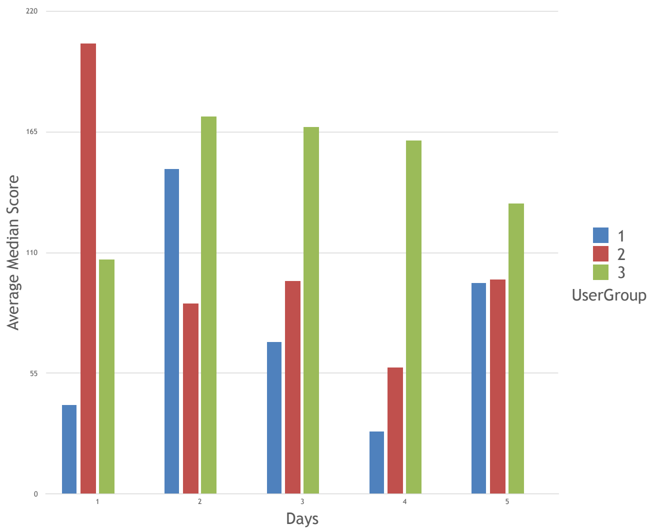


Figure 9.7: The comparative results between the intervention modes for negative emotion: sadness. The three groups use 1) baseline silent mode, 2) action card mode, and 3) stress feedback mode.

back and intervention after user feedback mode, mean agreement rating = 4.3. In addition, 100% of the participants in mode three were likely to recommend the application to a friend (mean agreement rating = 4.25) compared to 80% using the first mode (mean agreement rating = 4.25). However, only 20% of the participants using the second mode were likely to recommend the application to a friend (mean agreement rating = 3.25). This difference is due to the fact that most of the participants using the second mode remained neutral. Moreover, 100% in modes one and three agreed that they would use the app in the future compared to 80% in mode two (mean agreement ratings = 4.25, 3.75., and 4.0 for modes one, two, and three, respectively).

Appearance All participants agreed that the application had an attractive, clean and simple presentation for the biofeedback alone mode (mean agreement rating = 4.6), for the biofeedback and intervention mode (mean agreement rating = 4.0), and for the biofeedback and intervention after user feedback mode (mean agreement rating = 4.5).

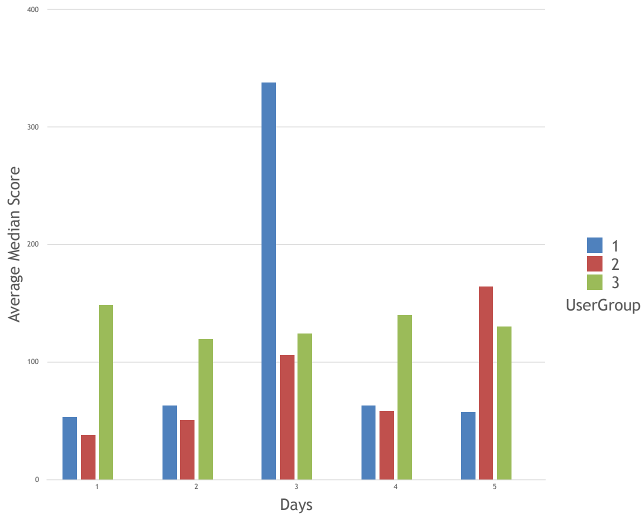


Figure 9.8: The comparative results between the intervention modes for negative emotion: fear. The three groups use 1) baseline silent mode, 2) action card mode, and 3) stress feedback mode.

9.4.2.2 Participant Likes, Dislikes, and Recommendations

Upon completion of the tasks, participants provided feedback for what they liked most and least about the application and recommendations for improvements.

The following comments capture what the participants liked most:

1. **For the interface:** Easy to use, easy to set up, quick to connect and start, user-friendly interface and the graphics.

2. **For emotional biofeedback:** It is fun to use, displaying my emotions through an emoji, giving me some information about the different emotions that I am experiencing at the same time. I like the novelty of a machine telling me about my emotions and measuring my emotion in real time.

3. **For interventions:** Makes me realize that I was too long in front of the screen, or too inactive. The app is very easy to navigate, and the notifications were 90% right on time and when needed and showed the way to make me relax.

The following comments capture what the participants liked the least:

1. Requires network connectivity to keep tracking,

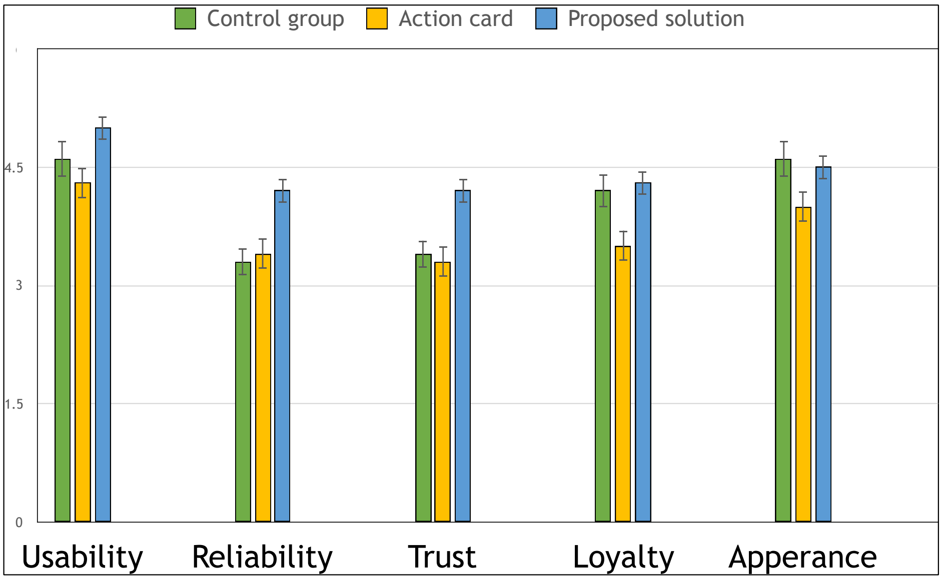


Figure 9.9: Likert-Scale and the level of agreement result after using inHarmony

2. Sometimes it shows that emotions are varying very quickly, which may not be the case (could be a margin of error).

3. I don't agree with the predictions sent. The main category names: the name selected for some range of emotions (like scared) did not reflect the whole range and had me questioning my own feelings.

There are two main dislike features: first, that the application relies on the internet and second, emotional feedback. The first feature is a design requirement to reduce the offload, and the second relies on the participants "emotional vocabulary." The application provides the user with five primary emotions: neutral, happy, sad, love, and fear. The participant then can discover the potential root cause of emotions that can vary based on the different levels of expression of the inner emotion. For example, rejection and helplessness are subsets of feeling afraid, while tired and bored are subsets of feeling sad.

It is also worth mentioning the negative effect of the biofeedback. One participant said, *"I think the instant feedback is not careful about states of mind. The home screen,* *when I look at that screen, I was very self-conscious, and that stressed me quite a bit."* Thus, selecting the feedback mode is crucial. We recommend using cumulative emotions instead of the seconds-based view to avoid the negative impact.

Change	Justification	Severity
Add emotional subsets	Participant comments also included the need for emotional subset	Medium
Used only the biofeedback, Intervention mode after the user feedback	Participants across both tests using modes one and two rated reliability and trust at 3-3.75 (out of 5).	

Table 9.5: Recommended changes and justifications

Table 9.5 provides the recommended changes and justifications driven by the participant success rate, behaviors, and comments.

The analysis of the RCT shows that most of the participants found inHarmony to be reliable, comprehensive, trustworthy and very useful, and easy to use. Having a just-in-time intervention with the right information content is key to many if not all of the participants. Implementing the recommendations and continuing to work with users will ensure continued user loyalty. In general, the participants had a positive perception of inHarmony. No major issues were discovered, and the results are generally promising and encouraging.

9.4.3 Limitations and Threats to Validity

Although the study results seem to be promising and encouraging, there are a number of limitations and threats to validity that are worth mentioning. The first is that the number of participants is small. It is always difficult to conduct usability studies with a high number of participants over the period of one week. Hence, the study had 15 participants with relevant and diverse gender, background, and education level. However, fifteen participants may not be sufficient to derive statistically significant conclusions. Indeed, more participants and experts are needed to improve the confidence in the results.

A second limitation is that the study was designed and conducted mainly by the thesis author. To mitigate the perception of bias, in designing the study, especially the survey questions, the SUPR-Q tool was used, and the questions were reviewed by usability experts. Regarding how the study was conducted, both the author's supervisor and a colleague assisted in the process to ensure that the usability sessions and the presented material were as neutral as possible in order to avoid influencing the participants. Participants' answers to open-ended questions were reported as is to mitigate biases that could have been introduced while reporting on the results.

A third limitation was participant inclusion criteria due to physiological signal characteristics. We assumed that all the users were healthy. None of the participants had experienced symptoms of excessive sweating (hyperhidrosis) or hypokalemia (bradycardia or tachycardia) or had a known history of heart disease. Moreover, physiological signals are vulnerable to drug and other substance use. Thus, none of the participants were experiencing mental health problems or taking antianxiety or antidepressant medications. In addition, coffee can affect the quality of physiological signals; therefore, we asked users to limit their caffeine intake during the experiment.

9.5 Concluding Remarks

In this chapter, the usability study design and components were presented. This usability study was conducted at University of Ottawa.

In our evaluation for biofeedback, we show that the system helps increase emotional self-awareness by reducing the predictive error by 3.333% for female participants and by 16.673% for male participants. The primary results suggest the usefulness of and need for the system to provide users with real-time biofeedback based on physiological signals.

Overall, the participants had a positive perception of inHarmony. In general, the results of the evaluation reveal an overall positive impact of the system on self-awareness and emotional well-being. No major issues were discovered, and the results are generally promising and encouraging. Thus, this system can provide real-time monitoring of emotions at any time and in any place, providing the user with more mobility and freedom. The solution therefore has the potential to help users increase productivity and efficiency, boost staff morale and reduce absenteeism.

Chapter 10

Conclusion and Future Work

This chapter repeats the thesis contributions and answers the research questions. In addition, it discusses some research opportunities and future directions.

10.1 Conclusion

In this thesis, we formalize emotion-aware systems through the notion of digital twins and LBFs. In contrast to the existing systems, in this thesis, we develop an emotion-aware biofeedback system. The development of the system and its main component are based on the live biofeedback model. The methods all have a similar scope and goal in relation to the problem of enabling a higher level of emotion well-being. Our interest is in the development of an emotion recognition algorithm using physiological signals that can then enable the development of LBF systems. Opportune moment detection methods are used to provide the dynamic intervention system and provide personalized recommendations for coping mechanisms. Moreover, the usability study provides insight into the perceived usefulness and effectiveness of the digital twin system in the workplace.

Most importantly, we discuss the process involved in the design and implementation of emotion recognition systems using physiological information, including physiological signal data acquisition, preprocessing, feature extraction, feature selection, and classification. We also discuss several existing LBF domains.

One of the main challenges in emotion detection is individual differences. The key point

to understanding individual differences in human emotions and affective phenomena is how to reveal and understand the exponential and dynamic nature of emotion, including how specific emotional components occur and fluctuate, and how the outlook of the component architecture is dependent on the person, the situation, and time. Time is the main element that explains the divergent dynamics of emotions across individuals, which highlights the importance of segmentation in the implementation process. We are interested in evaluating the relationship between a stimulus and emotion, focusing mainly on five emotions: neutral, happy, sad, erotic, and horror.
In contrast to the existing systems, in this thesis, we develop emotion-aware biofeedback systems that possess the following properties:

Real-time emotion recognition using physiological signals was used to access the recommendations. According to our research, no previous studies have reported the use of emotion recognition. Most studies have considered emotions based on the strength of various physiological signals. Moreover, to cluster users in real-time applications, sufficient and reliable data are needed. Such clustering may not be possible when using physiological signals because of the complexity of sensor placement, data analysis, and accuracy [69].

We define customizable procedures to limit the divergent dynamics of emotions across individuals. The aforementioned studies support the importance of the segmentation process in selecting the required features. However, to the best of our knowledge, minimal work has been conducted to simultaneously optimize the window frame size, delay time, and feature selection. Thus, we studied the contribution of each factor to the predictions. The relationships among frame size, delay time, feature selection and recognition accuracy were examined to optimize the window frame size and delay time to obtain the most important information.

This thesis addresses two main research questions related to the use of digital twins in emotional well-being:
RQ1. For emotional well-being, to what extent does the system help
a) illustrate the proposed goals, requirements, methods and constraints?
b) improve emotional well-being?

Answer: The case studies and the study results demonstrate that the system can be used to improve self-awareness and emotional well-being.

The primary results suggest the usefulness of and need for the system in providing users with real-time biofeedback based on physiological signals. In general, the results of the evaluation reveal an overall positive impact of the system on self-awareness and emotional well-being.

In addition, the system addresses the five principles of emotional intelligence by providing the user with emotional biofeedback for self-awareness and stress management techniques for self-regulation as well as motivation and social skills. Recognizing and sensing how other people feel will improve empathy skills and power relationships. Moreover, in addition to reading a group's emotions to discern the cause of the feelings, the system can help an organization address the needs of individuals and improve its ability to gain trust and influence.

Overall, discrimination among physiological patterns was improved using the ensemble process. This superiority is attributed to three reasons. First, the optimization methods help select the best parameters for the signals to have the most significant information possible. Second, the WMD-DTW learned the dynamic nature of the emotion and the individual differences in the variability of emotional intensity over time without the need for group-based models. Third, the use of the ensemble approach allows for multiple emotion models and learning algorithms to be jointly embedded within a metalearning model. This result is because the stacking result captures the individual differences in both emotional perception and the component architecture.

RQ2. What are the usefulness and usability of the proposed inHarmony system for emotional well-being?

Answer: To answer this question, a usability study was carried out with 35 practitioners. The results of the study were positive. The participants saw potential on an everyday basis. In addition, the participants agreed on the usefulness and effectiveness of the system in the workplace environment.

The usability study showed that most of the participants found inHarmony to be reliable, comprehensive, trustworthy and useful, and easy to use. The use of the just-in-time intervention was key to many if not all of the participants. In general, the participants had a positive perception of inHarmony. No major issues were discovered, and the results are generally promising and encouraging.

10.2 Future Directions

There are many opportunities for following up on the thesis work, including these directions.

Further evaluation of inHarmony: Another usability study could be conducted with more participants. In addition, we plan to study the dynamic effect of biofeedback and the intervention over a long period (e.g., increasing productivity and efficiency, improving staff morale and reducing absenteeism).

Additionnal case studies: Conducting additional case studies would help improve the generalizability of the system, especially in situations not yet covered by this thesis, with collaboration between researchers, practitioners, and proficients. For example, the study in this thesis was concerned mainly with individual subjects. In the future, we are interested in studying collective groups of people in the workplace or in particular communities or cities at a particular time of day to analyze the reasons for their emotions in order to achieve a better smart city.

Extend emotion recognition to emotional subsets: As some participants in the usability study suggested, subsets of emotions could be added. It is encouraging to investigate this idea by taking advantage of contextual information.

Reference

[1] R. Tschäppeler and M. Krogerus, *The decision book: fifty models for strategic thinking*, Profile (2011).

[2] D. Holman, S. Johnson, and E. O'Connor. *Stress management interventions: Improving subjective psychological well-being in the workplace.* In *Handbook of wellbeing. Salt Lake City, UT: DEF Publishers. DOI: nobascholar. com*, (2018).

[3] F. Al Machot, A. Elmachot, M. Ali, E. Al Machot, and K. Kyamakya. *A Deep-Learning Model for Subject-Independent Human Emotion Recognition Using Electrodermal Activity Sensors.* volume 19, page 1659. Multidisciplinary Digital Publishing Institute, (2019).

[4] R. Dodge, A. P. Daly, J. Huyton, and L. D. Sanders. *The challenge of defining wellbeing.* In *International journal of wellbeing*, volume 2, (2012).

[5] E. Langeland. *Emotional Well-Being.* In *In Encyclopedia of Quality of Life and Well-Being Research*, pages 1874–1876. Springer, (2014).

[6] G. E. Coverdale and A. F. Long. *Emotional wellbeing and mental health: an exploration into health promotion in young people and families.* In *Perspectives in public health*, volume 135, pages 27–36. Sage Publications Sage UK: London, England, (2015).

[7] A. S. Drigas and C. Papoutsi. *A New Layered Model on Emotional Intelligence.* In *Behavioral Sciences*, volume 8, page 45. Multidisciplinary Digital Publishing Institute, (2018).

[8] D. Goleman, *Working with emotional intelligence*, Bantam (1998).

[9] J. J. Gross and O. P. John. *Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being.* In *Journal of personality and social psychology.* American Psychological Association, (2003).

[10] G. Gonzalez, J. L. De La Rosa, M. Montaner, and S. Delfin. *Embedding emotional context in recommender systems.* In *Data Engineering Workshop, IEEE 23rd Inter. Conf. on*, (2007).

[11] P. Salovey and J. D. Mayer. *Emotional intelligence.* In *Imagination, cognition and personality.* Sage Publications Sage CA: Los Angeles, CA, (1990).

[12] V. Dulewicz and M. Higgs. *Can emotional intelligence be developed?* In *The International Journal of Human Resource Management*, volume 15, pages 95–111. Taylor & Francis, (2004).

[13] C. E. Izard and S. Buechler. *Aspects of consciousness and personality in terms of differential emotions theory.* In *Theories of emotion.* Elsevier, (1980).

[14] A. Thiruchelvi and M. Supriya. *Emotional intelligence and job satisfaction.* In *Asia Pacific Business Review*, volume 5, pages 109–115. SAGE Publications Sage India: New Delhi, India, (2009).

[15] J. D. Mayer and Y. N. Gaschke. *The experience and meta-experience of mood.* In *Journal of personality and social psychology.* US: American Psychological Association, (1988).

[16] P. Salovey, J. D. Mayer, D. Caruso, and P. N. Lopes. *Measuring emotional intelligence as a set of abilities with the Mayer-Salovey-Caruso Emotional Intelligence Test.* In *Handbook of positive psychology assessment.* American Psychological Association, (2003).

[17] Z. Jiang. *Emotional intelligence and career decision-making self-efficacy: Mediating roles of goal commitment and professional commitment.* In *journal of employment counseling*, volume 53, pages 30–47. Wiley Online Library, (2016).

[18] T. R. Schneider, J. B. Lyons, and S. Khazon. *Personality and Individual Differences.* In *Emotional intelligence and resilience*, volume 55, pages 909–914. Elsevier, (2013).

[19] N. S. Schutte, J. M. Malouff, C. Bobik, T. D. Coston, C. Greeson, C. Jedlicka, E. Rhodes, and G. Wendorf. *Emotional intelligence and interpersonal relations.* In *The Journal of social psychology*, volume 141, pages 523–536. Taylor & Francis, (2001).

[20] N. Nouri, B. Moeini, A. Karimi-Shahanjarini, J. Faradmal, A. Ghaleiha, and M. Asnaashari. *Relationship between emotional intelligence and communication skills among high school students in Hamadan based on the theory of social support.* In *Journal of Education and Community Health*, volume 1, pages 38–46. Directory of Open Access Journals, (2014).

[21] C. E. Izard. *Emotion theory and research: Highlights, unanswered questions, and emerging issues.* In *Annual review of psychology*, volume 60, pages 1–25. Annual Reviews, (2009).

[22] B. Kitchenham. *Procedures for performing systematic reviews.* In *Keele, UK, Keele University*, volume 33, pages 1–26, (2004).

[23] P. Ekman, R. W. Levenson, and W. V. Friesen. *Autonomic nervous system activity distinguishes among emotions.* American Association for the Advancement of Science, (1983).

[24] L. F. Barrett. *Are emotions natural kinds?* In *Perspectives on psychological science*, volume 1, pages 28–58. SAGE Publications Sage CA: Los Angeles, CA, (2006).

[25] C. A. Smith and L. D. Kirby. *Putting appraisal in context: Toward a relational model of appraisal and emotion.* In *Cognition and Emotion*, volume 23, pages 1352–1372. Taylor & Francis, (2009).

[26] C. A. Smith, R. S. Lazarus, et al. *Emotion and adaptation.* In *Handbook of personality: Theory and research*, pages 609–637, (1990).

[27] C. E. Izard. *Organizational and motivational functions of discrete emotions.* Guilford Press, (1993).

[28] P. Kuppens, J. Stouten, and B. Mesquita. *Individual differences in emotion components and dynamics: Introduction to the special issue.* In *Cognition and Emotion*, volume 23, pages 1249–1258. Taylor & Francis, (2009).

[29] J. A. Russell and G. Pratt. *A description of the affective quality attributed to environments.* In *Journal of personality and social psychology*, volume 38, page 311. American Psychological Association, (1980).

[30] J. A. Russell and A. Mehrabian. *Distinguishing anger and anxiety in terms of emotional response factors.* In *Journal of consulting and clinical psychology*, volume 42, page 79. American Psychological Association, (1974).

[31] M. F. King and G. C. Bruner. *Social desirability bias: A neglected aspect of validity testing.* In *Psychology and Marketing*, volume 17, pages 79–103, (2000).

[32] P. Totterdell, B. Parkinson, R. B. Briner, and S. Reynolds. *Forecasting feelings: The accuracy and effects of self-predictions of mood.* In *Journal of Social Behavior and Personality*, volume 12, page 631. Select Press, (1997).

[33] K. Poels and S. Dewitte. *How to capture the heart? Reviewing 20 years of emotion measurement in advertising.* In *Journal of Advertising Research*, volume 46, pages 18–37. Journal of Advertising Research, (2006).

[34] B. T. Engel. *Stimulus-response and individual-response specificity.* In *AMA Archives of General Psychiatry*, volume 2, pages 305–313. American Medical Association, (1960).

[35] G. Stemmler and J. Wacker. *Personality, emotion, and individual differences in physiological responses.* In *Biological psychology*, volume 84, pages 541–551. Elsevier, (2010).

[36] P. Lang and M. M. Bradley. *The International Affective Picture System (IAPS)*

in the study of emotion and attention. In *Handbook of emotion elicitation and assessment*, volume 29. Oxford University Press USA New-York, NY, USA, (2007).

[37] J. Kim and E. Andre. *Emotion recognition based on physiological changes in music listening.* In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 30, page 12, (2008).

[38] K. H. Kim, S. W. Bang, and S. R. Kim. *Emotion recognition system using short-term monitoring of physiological signals.* In *Medical and biological engineering and computing*, volume 42, pages 419–427. Springer, (2004).

[39] S. Carvalho, J. Leite, S. Galdo-Álvarez, and O. F. Gonçalves. *The emotional movie database (EMDB): A self-report and psychophysiological study.* In *Applied psychophysiology and biofeedback*, volume 37, pages 279–294. Springer, (2012).

[40] P. Ekman. *Are there basic emotions?* In *Psychological review*, number 99, pages 550–3. American Psychological Association, (1992).

[41] J. Sørensen. *Measuring emotions in a consumer decision-making contextapproaching or avoiding.* In *Aalborg University, Denmark.* Citeseer, (2008).

[42] R. Plutchik and H. Kellerman, *Emotion Profile Index*, Western Psychological Services (1974).

[43] J. D. Morris. *Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response.* In *Journal of advertising research*, volume 35, pages 63–68, (1995).

[44] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert. *Discriminating stress from cognitive load using a wearable EDA device.* In *IEEE Transactions on information technology in biomedicine*, volume 14, pages 410–417. IEEE, (2010).

[45] S. C. Jacobs, R. Friedman, J. D. Parker, G. H. Tofler, A. H. Jimenez, J. E. Muller, H. Benson, and P. H. Stone. *Use of skin conductance changes during mental stress*

testing as an index of autonomic arousal in cardiovascular research. In *American heart journal*, volume 128, pages 1170–1177. Elsevier, (1994).

[46] M. L. Loggia, M. Juneau, and M. C. Bushnell. *Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity.* In *PAIN®*, volume 152, pages 592–598. Elsevier, (2011).

[47] D. Goleman, *Emotional intelligence*, Bantam (2006).

[48] C. H. Peterson. *The individual regulation component of group emotional intelligence: Measure development and validation.* In *The Journal for Specialists in Group Work*, volume 37, pages 232–251. Taylor & Francis, (2012).

[49] E. S. Owens, F. W. McPharlin, N. Brooks, and K. Fritzon. *The Effects of Empathy, Emotional Intelligence and Psychopathy on Interpersonal Interactions.* In *Psychiatry, Psychology and Law*, volume 25, pages 1–18. Taylor & Francis, (2018).

[50] S. Stanley and G. M. Bhuvaneswari. *Reflective ability, empathy, and emotional intelligence in undergraduate social work students: A cross-sectional study from India.* In *Social Work Education*, volume 35, pages 560–575. Taylor & Francis, (2016).

[51] F. Akhtar. *Empathy: What it is and Why it Matters.* In *Journal of Social Work Practice*, volume 27, pages 474–476. Taylor & Francis, (2013).

[52] H. H. Newman, F. N. Freeman, and K. J. Holzinger. *Twins: A study of heredity and environment.* Univ. Chicago Press, (1937).

[53] B. Gogarty. *What exactly is an exact copy? And why it matters when trying to ban human reproductive cloning in Australia.* In *Journal of medical ethics*, volume 29, pages 84–89. Institute of Medical Ethics, (2003).

[54] A. E. Saddik. *Digital Twins: The Convergence of Multimedia Technologies.* In *IEEE MultiMedia*, volume 25, pages 87–92, (2018).

[55] S. S. David W. Cearley, Brian Burke and M. J. Walker. *Top 10 Strategic Technology Trends for 2018.* In *Gartner*, (2017).

[56] A. Gaggioli and G. Riva. *From mobile mental health to mobile wellbeing: opportunities and challenges.* In *MMVR*, pages 141–147, (2013).

[57] V. Mishra, T. Hao, S. Sun, K. N. Walter, M. J. Ball, C.-H. Chen, and X. Zhu. *Investigating the Role of Context in Perceived Stress Detection in the Wild.* In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 1708–1716. ACM, (2018).

[58] H. Al Osman, M. Eid, and A. El Saddik. *U-biofeedback: a multimedia-based reference model for ubiquitous biofeedback systems.* In *Multimedia tools and applications*, volume 72, pages 3143–3168. Springer, (2014).

[59] H. Al Osman, H. Dong, and A. El Saddik. *Ubiquitous biofeedback serious game for stress management.* In *IEEE Access*. IEEE, (2016).

[60] K. Lansing, W. Yu, and B. Samanta. *A Non-intrusive Wearable Bio-sensor Based Assistive Robotic System for Human Mental and Physical Intervention.* In *ASME 2017 International Mechanical Engineering Congress and Exposition*, pages V04BT05A025–V04BT05A025. American Society of Mechanical Engineers, (2017).

[61] V. Rajanna, F. Alamudun, D. Goldberg, and T. Hammond. *Let me relax: toward automated sedentary state recognition and ubiquitous mental wellness solutions.* In *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare*, pages 28–33. ICST (Institute for Computer Sciences, Social-Informatics and . . . , (2015).

[62] A. Millings, J. Morris, A. Rowe, S. Easton, J. K. Martin, D. Majoe, and C. Mohr. *Can the effectiveness of an online stress management program be augmented by wearable sensor technology?* In *Internet Interventions*. Elsevier, (2015).

[63] P. Jercic, P. J. Astor, M. T. P. Adam, O. Hilborn, K. Schaaff, C. Lindley, C. Sennersten, and J. Eriksson. *A Serious Game using Physiological Interfaces for*

Emotion regulation Training in the Context of Financial Decision-Making. In *ECIS*, (2012).

[64] S. Serino, P. Cipresso, A. Gaggioli, F. Pallavicini, S. Cipresso, D. Campanaro, and G. Riva. *Smartphone for self-management of psychological stress: a preliminary evaluation of positive technology app.* In *Revista de Psicopatología y Psicología Clínica*, volume 19, pages 253–260. AEPCP, (2014).

[65] B. Zhu, A. Hedman, S. Feng, H. Li, and W. Osika. *Designing, prototyping and evaluating digital mindfulness applications: a case study of mindful breathing for stress reduction.* In *Journal of medical Internet research*, volume 19. JMIR Publications Inc., (2017).

[66] P. E. Paredes, Y. Zhou, N. A.-H. Hamdan, S. Balters, E. Murnane, W. Ju, and J. A. Landay. *Just breathe: In-car interventions for guided slow breathing.* In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, volume 2, page 28. ACM, (2018).

[67] P. Paredes and M. Chan. *CalmMeNow: exploratory research and design of stress mitigating mobile interventions.* In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 1699–1704. ACM, (2011).

[68] B. Cvetković, M. Gjoreski, J. Šorn, P. Maslov, M. Kosiedowski, M. Bogdański, A. Stroiński, and M. Luštrek. *Real-time physical activity and mental stress management with a wristband and a smartphone.* In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 225–228. ACM, (2017).

[69] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, and X. Yang. *A Review of Emotion Recognition Using Physiological Signals.* In *Sensors (Basel, Switzerland)*, (2018).

[70] M. Munster-Segev, O. Fuerst, S. A. Kaplan, and A. Cahn. *Incorporation of a stress reducing mobile app in the care of patients with type 2 diabetes: a prospec-*

tive study. In *JMIR mHealth and uHealth*, volume 5. JMIR Publications Inc., (2017).

[71] B. K. Wiederhold, C. Boyd, C. Sulea, A. Gaggioli, and G. Riva. *Marketing analysis of a positive technology app for the self-management of psychological stress.* In *Stud Health Technol Inform*, volume 199, pages 83–87, (2014).

[72] S. Clarke, L. G. Jaimes, and M. A. Labrador. *mstress: A mobile recommender system for just-in-time interventions for stress.* In *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–5. IEEE, (2017).

[73] A. Lentferink, L. Polstra, M. de Groot, H. Oldenhuis, H. Velthuijsen, and L. van Gemert-Pijnen. *The Values of Self-tracking and Persuasive eCoaching According to Employees and Human Resource Advisors for a Workplace Stress Management Application: A Qualitative Study.* In *International Conference on Persuasive Technology*, pages 160–171. Springer, (2018).

[74] L. G. Jaimes, M. Llofriu, and A. Raij. *A stress-free life: just-in-time interventions for stress via real-time forecasting and intervention adaptation.* In *Proceedings of the 9th International Conference on Body Area Networks*, pages 197–203. ICST (Institute for Computer Sciences, Social-Informatics and . . . , (2014).

[75] M. Matthews, G. Doherty, D. Coyle, and J. Sharry. *Designing mobile applications to support mental health interventions.* In *Handbook of research on user interface design and evaluation for mobile technology*, pages 635–656. IGI Global, (2008).

[76] M. Chen, P. Zhou, and G. Fortino. *Emotion Communication System.* In *IEEE Access*, volume 5, pages 326–337, (2017).

[77] C. Katsis, N. Katertsidis, G. Ganiatsas, and D. Fotiadis. *Toward emotion recognition in car-racing drivers: a biosignal processing approach.* In *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on Systems, Man and Cybernetics*, volume 38, pages 502–512, (2008).

[78] G. Keren, T. Kirschstein, E. Marchi, F. Ringeval, and B. Schuller. *END-TO-END LEARNING FOR DIMENSIONAL EMOTION RECOGNITION FROM PHYSIOLOGICAL SIGNALS.* In *IEEE International Conference on Multimedia and Expo (ICME)*, (2017).

[79] A. Greco, A. Lanata, L. Citi, N. Vanello, G. Valenza, and E. P. Scilingo. *Skin admittance measurement for emotion recognition: A study over frequency sweep.* In *Electronics*, volume 5, page 46. Multidisciplinary Digital Publishing Institute, (2016).

[80] F. Ringeval, F. Eyben, E. Kroupi, A. Yuce, J.-P. Thiran, T. Ebrahimi, D. Lalanne, and B. Schuller. *Prediction of asynchronous dimensional emotion ratings from audiovisual and physiological data.* In *Pattern Recognition Letters*, volume 66, pages 22–30. Elsevier, (2015).

[81] G. Valenza, A. Lanatá, and E. P. Scilingo. *Improving emotion recognition systems by embedding cardiorespiratory coupling.* In *Physiological measurement*, volume 34, page 449. IOP Publishing, (2013).

[82] J. A. Russell. *Core affect and the psychological construction of emotion.* In *Psychological review*, volume 110, page 145. American Psychological Association, (2003).

[83] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic. *A multimodal database for affect recognition and implicit tagging.* In *IEEE Transactions on Affective Computing*, volume 3, pages 42–55. IEEE, (2012).

[84] M. B. H. Wiem and Z. Lachiri. *Emotion Classification in Arousal Valence Model using MAHNOB-HCI Database.* In *Int. J. Adv. Comput. Sci. Appl. IJACSA*, volume 8, (2017).

[85] C. Li, C. Xu, and Z. Feng. *Analysis of physiological for emotion recognition with the IRS model.* In *Neurocomputing*, volume 178, pages 103–111. Elsevier, (2016).

[86] M. Zhao, F. Adib, and D. Katabi. *Emotion recognition using wireless signals.* In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, pages 95–108. ACM, (2016).

[87] Z. Khalili and M. H. Moradi. *Emotion detection using brain and peripheral signals.* In *Biomedical Engineering Conference CIBEC*, pages 1–4, (2008).

[88] C. Maaoui and A. Pruski. *Emotion recognition through physiological signals for human-machine communication.* In *Cutting Edge Robotics 2010*. InTech, (2010).

[89] Y. Gu, S.-L. Tan, K.-J. Wong, M.-H. R. Ho, and L. Qu. *A biometric signature based system for improved emotion recognition using physiological responses from multiple subjects.* In *Industrial Informatics (INDIN), 2010 8th IEEE International Conference on*, pages 61–66. IEEE, (2010).

[90] S. Hamann and T. Canli. *Individual differences in emotion processing.* In *Current opinion in neurobiology*, volume 14, pages 233–238. Elsevier, (2004).

[91] M. D. Lewis. *Bridging emotion theory and neurobiology through dynamic systems modeling.* In *Behavioral and brain sciences*, volume 28, pages 169–194. Cambridge University Press, (2005).

[92] D. Sander, D. Grandjean, and K. R. Scherer. *A systems approach to appraisal mechanisms in emotion.* In *Neural networks*, volume 18, pages 317–352. Elsevier, (2005).

[93] P. Verduyn, E. Delvaux, H. Van Coillie, F. Tuerlinckx, and I. Van Mechelen. *Predicting the duration of emotional experience: two experience sampling studies.* In *Emotion*, volume 9, page 83. American Psychological Association, (2009).

[94] P. Kuppens, I. Van Mechelen, J. B. Nezlek, D. Dossche, and T. Timmermans. *Individual differences in core affect variability and their relationship to personality and psychological adjustment.* In *Emotion*, volume 7, page 262. American Psychological Association, (2007).

[95] I. B. Mauss, R. W. Levenson, L. McCarter, F. H. Wilhelm, and J. J. Gross. *The tie that binds? Coherence among emotion experience, behavior, and physiology.* In *Emotion*, volume 5, page 175. American Psychological Association, (2005).

[96] D. Sankoff and J. B. Kruskal. *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison.* In *Addison-Wesley Publication*, (1983).

[97] G. A. ten Holt, M. J. Reinders, and E. A. Hendriks. *Multi-Dimensional Dynamic Time Warping for Gesture Recognition.* In *Thirteenth annual conference of the Advanced School for Computing and Imaging*, (2007).

[98] J. Wang, A. Samal, J. R. Green, and F. Rudzicz. *Sentence recognition from articulatory movements for silent speech interfaces.* In *Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, (2012).

[99] F. Petitjean, J. Inglada, and P. Gançarski. *Satellite image time series analysis under time warping.* In *IEEE transactions on geoscience and remote sensing*, volume 50, pages 3081–3095. IEEE, (2012).

[100] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. *Deap: A database for emotion analysis; using physiological signals.* In *IEEE Transactions on Affective Computing*, volume 3, pages 18–31. IEEE, (2012).

[101] J. A. Healey and R. W. Picard. *Detecting stress during real-world driving tasks using physiological sensors.* In *IEEE Transactions on intelligent transportation systems*, volume 6, pages 156–166. IEEE, (2005).

[102] E. Douglas-Cowie, R. Cowie, I. Sneddon, C. Cox, O. Lowry, M. Mcrorie, J.-C. Martin, L. Devillers, S. Abrilian, A. Batliner, et al. *The HUMAINE database: addressing the collection and annotation of naturalistic and induced emotional data.* In *Affective computing and intelligent interaction*, pages 488–500. Springer, (2007).

[103] C. Godin, F. Prost-Boucle, A. Campagne, S. Charbonnier, S. Bonnet, and A. Vidal. *Selection of the most relevant physiological features for classifying emotion.* In *Emotion*, volume 40, page 20, (2015).

[104] P. Kuppens, F. Tuerlinckx, M. Yik, P. Koval, J. Coosemans, K. J. Zeng, and J. A. Russell. *The relation between valence and arousal in subjective experience varies with personality and culture.* In *Journal of personality*, volume 85, pages 530–542. Wiley Online Library, (2017).

[105] A. P. Happonen, E. Mattila, M.-L. Kinnunen, V. Ikonen, T. Myllymaki, K. Kaipainen, H. Rusko, R. Lappalainen, and I. Korhonen. *P4Well concept to empower self-management of psychophysiological wellbeing and load recovery.* In *2009 3rd International Conference on Pervasive Computing Technologies for Healthcare*, pages 1–8. IEEE, (2009).

[106] A. El Saddik. *Digital twins: the convergence of multimedia technologies.* In *IEEE MultiMedia*, volume 25, pages 87–92. IEEE, (2018).

[107] P. Sanches, K. Höök, E. Vaara, C. Weymann, M. Bylund, P. Ferreira, N. Peira, and M. Sjölinder. *Mind the body!: designing a mobile stress management application encouraging personal reflection.* In *Proceedings of the 8th ACM conference on designing interactive systems*, pages 47–56. ACM, (2010).

[108] R. Ilies, S. S. Aw, and V. K. Lim. *A naturalistic multilevel framework for studying transient and chronic effects of psychosocial work stressors on employee health and well-being.* In *Applied Psychology*, volume 65, pages 223–258. Wiley Online Library, (2016).

[109] P. Paredes, R. Gilad-Bachrach, M. Czerwinski, A. Roseway, K. Rowan, and J. Hernandez. *PopTherapy: Coping with stress through pop-culture.* In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pages 109–117. ICST (Institute for Computer Sciences, Social-Informatics and . . . , (2014).

[110] J. Löwgren. *Pliability as an experiential quality: Exploring the aesthetics of interaction design.* In *Artifact: Journal of Design Practice*, volume 1, pages 85–95. Intellect, (2007).

[111] D. V. Dimitrov. *Medical internet of things and big data in healthcare.* In *Healthcare informatics research*, volume 22, pages 156–163, (2016).

[112] U. Lee, K. Han, H. Cho, K.-M. Chung, H. Hong, S.-J. Lee, Y. Noh, S. Park, and J. M. Carroll. *Intelligent positive computing with mobile, wearable, and IoT devices: Literature review and research directions.* In *Ad Hoc Networks*, volume 83, pages 8–24. Elsevier, (2019).

[113] D. C. Mohr, S. M. Schueller, E. Montague, M. N. Burns, and P. Rashidi. *The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions.* In *Journal of medical Internet research*, volume 16. JMIR Publications Inc., (2014).

[114] B. L. Seaward, *Managing stress*, Jones & Bartlett Learning (2017).

[115] M. Borritz, U. Bültmann, R. Rugulies, K. B. Christensen, E. Villadsen, and T. S. Kristensen. *Psychosocial work characteristics as predictors for burnout: findings from 3-year follow up of the PUMA Study.* In *Journal of occupational and environmental medicine*, volume 47, pages 1015–1025. LWW, (2005).

[116] I. Nahum-Shani, S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, and S. A. Murphy. *Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support.* In *Annals of Behavioral Medicine*, volume 52, pages 446–462. Oxford University Press US, (2017).

[117] G. Mark, D. Gudith, and U. Klocke. *The cost of interrupted work: more speed and stress.* In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 107–110. ACM, (2008).

[118] R. Plutchik. *The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice.* In *American scientist.* JSTOR, (2001).

[119] K. S. LaBar and R. Cabeza. *Cognitive neuroscience of emotional memory.* In *Nature Reviews Neuroscience*, volume 7, pages 54–64. Nature Publishing Group, (2006).

[120] M. Garbarino, M. Lai, D. Bender, R. W. Picard, and S. Tognetti. *Empatica E3—A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition.* In *Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on*, pages 39–42. IEEE, (2014).

[121] J. O. Gomez, Carles and J. Paradells. *Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology.* In *Sensorss*, volume 12, pages 11734–11753, (2012).

[122] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino. *Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges.* In *Information Fusion*, volume 35, pages 68–80. Elsevier, (2017).

[123] A. Sonderegger, K. Heyden, A. Chavaillaz, and J. Sauer. *AniSAM & AniAvatar: Animated Visualizations of Affective States.* In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4828–4837. ACM, (2016).

[124] M. Yik, J. A. Russell, and J. H. Steiger. *A 12-point circumplex structure of core affect.* In *Emotion*, volume 11, page 705. American Psychological Association, (2011).

[125] Gallup and J. Clifton. *Gallup 2017 Global Emotions Report.* Gallup World Poll, (2017).

[126] F. Gino and G. Pisano. *Toward a theory of behavioral operations.* In *Manufacturing & Service Operations Management*, volume 10, pages 676–691. INFORMS, (2008).

[127] A. Greco, G. Valenza, A. Lanata, E. P. Scilingo, and L. Citi. *cvxEDA: A convex optimization approach to electrodermal activity processing.* In *IEEE Transactions on Biomedical Engineering*, pages 797–804. IEEE, (2016).

[128] S. M. LaValle, M. S. Branicky, and S. R. Lindemann. *On the relationship between classical grid search and probabilistic roadmaps.* In *The International Journal of Robotics Research*, volume 23, pages 673–692. SAGE Publications, (2004).

[129] D. McGlynn and M. G. Madden. *An ensemble dynamic time warping classifier with application to activity recognition.* In *Research and development in intelligent systems xxvii*, pages 339–352. Springer, (2011).

[130] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. Keogh. *Generalizing DTW to the multi-dimensional case requires an adaptive approach.* In *Data Mining and Knowledge Discovery*. Springer, (2017).

[131] Y.-S. Jeong, M. K. Jeong, and O. A. Omitaomu. *Weighted dynamic time warping for time series classification.* In *Pattern Recognition*. Elsevier, (2011).

[132] T. Arici, S. Celebi, A. S. Aydin, and T. T. Temiz. *Robust gesture recognition using feature pre-processing and weighted dynamic time warping.* In *Multimedia Tools and Applications*. Springer, (2014).

[133] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. *Algorithms for Hyper-parameter Optimization.* In *Proceedings of the 24th International Conference on Neural Information Processing Systems*, NIPS'11, pages 2546–2554, USA, (2011). Curran Associates Inc.

[134] S. M. LaValle, M. S. Branicky, and S. R. Lindemann. *On the relationship between classical grid search and probabilistic roadmaps.* In *The International Journal of Robotics Research*. SAGE Publications, (2004).

[135] J. Snoek, H. Larochelle, and R. P. Adams. *Practical bayesian optimization of machine learning algorithms.* In *Advances in neural information processing systems*, pages 2951–2959, (2012).

[136] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. *Taking the Human Out of the Loop: A Review of Bayesian Optimization.* In *Proceedings of the IEEE*, volume 104, pages 148–175, (2016).

[137] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. *Scikit-learn: Machine learning in Python.* In *Journal of Machine Learning Research*, volume 12, pages 2825–2830, (2011).

[138] D. H. Wolpert. *Stacked generalization.* In *Neural networks*, volume 5, pages 241–259. Elsevier, (1992).

[139] C. C. Aggarwal, *Data classification: algorithms and applications*, CRC Press (2014).

[140] B. Clarke. *Comparing Bayes model averaging and stacking when model approximation error cannot be ignored.* In *Journal of Machine Learning Research*, volume 4, pages 683–712, (2003).

[141] K. M. Ting and I. H. Witten. *Issues in stacked generalization.* In *J. Artif. Intell. Res.(JAIR)*, volume 10, pages 271–289, (1999).

[142] G. Seni and J. F. Elder. *Ensemble methods in data mining: improving accuracy through combining predictions.* In *Synthesis Lectures on Data Mining and Knowledge Discovery*, volume 2, pages 1–126. Morgan & Claypool Publishers, (2010).

[143] R. Kohavi et al. *A study of cross-validation and bootstrap for accuracy estimation and model selection.* In *Ijcai*. Stanford, CA, (1995).

[144] K. Sherine. *An Introduction to Emotive UI.* Online, (2016).

[145] J. P. Cummings and K. I. Pargament. *Medicine for the spirit: Religious coping in individuals with medical conditions.* In *Religions*, volume 1, pages 28–53. Molecular Diversity Preservation International, (2010).

[146] H. van Wietmarschen, B. Tjaden, M. van Vliet, M. Battjes-Fries, and M. Jong. *Effects of mindfulness training on perceived stress, self-compassion, and self-reflection of primary care physicians: a mixed-methods study.* In *BJGP open*, volume 2, page bjgpopen18X101621. Royal College of General Practitioners, (2018).

[147] J. Kabat-Zinn. *Mindfulness-based interventions in context: past, present, and future.* In *Clinical psychology: Science and practice*, volume 10, pages 144–156. Wiley Online Library, (2003).

[148] S. L. Shapiro, L. E. Carlson, J. A. Astin, and B. Freedman. *Mechanisms of mindfulness.* In *Journal of clinical psychology*, volume 62, pages 373–386. Wiley Online Library, (2006).

[149] S. A. Welsh. *Health Habits, Coping Behaviors, and Perceived Social Support in Primary Care Physicians as a Function of Level of Burnout.* (2017).

[150] H. Shin, Y. M. Park, J. Y. Ying, B. Kim, H. Noh, and S. M. Lee. *Relationships between coping strategies and burnout symptoms: A meta-analytic approach.* In *Professional Psychology: Research and Practice*, volume 45, page 44. American Psychological Association, (2014).

[151] A. Albraikan, B. Hafidh, and A. El Saddik. *iAware: A Real-Time Emotional Biofeedback System Based on Physiological Signals.* In *IEEE Access*, volume 6, pages 78780–78789. IEEE, (2018).

[152] M. P. Leiter and C. Maslach, *Banishing burnout: Six strategies for improving your relationship with work*, John Wiley & Sons (2005).

[153] J. Montero-Marín, P. Skapinakis, R. Araya, M. Gili, and J. García-Campayo. *Towards a brief definition of burnout syndrome by subtypes: development of the" burnout clinical subtypes questionnaire"(BCSQ-12).* In *Health and quality of life outcomes*, volume 9, page 74. BioMed Central, (2011).

[154] J. Montero-Marin, F. Zubiaga, M. Cereceda, M. M. P. Demarzo, P. Trenc, and J. Garcia-Campayo. *Burnout subtypes and absence of self-compassion in primary healthcare professionals: A cross-sectional study.* In *PLoS One*, volume 11, page e0157499. Public Library of Science, (2016).

[155] R. Tangri, *Stress costs, stress cures*, Trafford Publishing (2003).

[156] C. Williams and N. J. *Stress at work.* In *Canadian Social Trends*, volume 4, pages 7–13, (2003).

[157] J. Burton. *The business case for a healthy workplace.* In *Industrial Accident Prevention Association, Ontario*, (2008).

[158] M. Shain et al. *Mental Health and Substance Abuse at Work: Perspectives from Research and Implications for Leaders, a background paper prepared by the Scientific Advisory Committee to the Global Business and Economic Roundtable on Addiction and Mental Health*, (2002).

[159] S. E. Hobfoll and A. Shirom. *Conservation of resources theory: Applications to stress and management in the workplace.* In *Marcel Dekker*, volume 2, pages 57–81, (2001).

[160] C. Maslach, W. B. Schaufeli, and M. P. Leiter. *Job burnout.* In *Annual review of psychology*, volume 52, pages 397–422. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, (2001).

[161] J. Montero-Marín, J. García-Campayo, D. M. Mera, and Y. L. del Hoyo. *A new definition of burnout syndrome based on Farber's proposal.* In *Journal of Occupational Medicine and Toxicology*, volume 4, page 31. BioMed Central, (2009).

[162] J. Montero-Marín and J. García-Campayo. *A newer and broader definition of burnout: Validation of the" Burnout Clinical Subtype Questionnaire (BCSQ-36)".* In *BMC Public Health*, volume 10, page 302. BioMed Central, (2010).

[163] C. Maslach, S. E. Jackson, M. P. Leiter, W. B. Schaufeli, and R. L. Schwab, *Maslach burnout inventory*, Consulting Psychologists Press Palo Alto, CA (1986).

[164] W. B. Schaufeli, A. B. Bakker, and M. Salanova. *The measurement of work engagement with a short questionnaire: A cross-national study*. In *Educational and psychological measurement*, volume 66, pages 701–716. Sage Publications Sage CA: Thousand Oaks, CA, (2006).

[165] A. Albraikan, D. P. Tobón, and A. E. Saddik. *Toward User-Independent Emotion Recognition using Physiological Signals*. In *IEEE Sensors Journal*, volume 6, pages 1–1, (2019).

[166] R. E. Thomson and W. J. Emery, *Data analysis methods in physical oceanography*, Newnes (2014).

[167] W. Boucsein, *Electrodermal activity*, Springer Science & Business Media (2012).

[168] A. B. Diego Zandrino. *2018 FIFA World Cup Russia*. https://www.fifa.com. [Online]. Accessed: 29-July-2018.

[169] L. R. Brody and J. A. Hall. *Gender and emotion in context*. In *Handbook of emotions*, volume 2, (2008).

[170] J. Jansz et al. *Masculine identity and restrictive emotionality*. In *Gender and emotion: Social psychological perspectives*, volume 2, (2000).

[171] S. R. Wester, D. L. Vogel, P. K. Pressly, and M. Heesacker. *Sex differences in emotion: A critical review of the literature and implications for counseling psychology*. In *The Counseling Psychologist*. Sage Publications Sage CA, (2002).

[172] A. H. Fischer and A. S. Manstead. *The relation between gender and emotions in different cultures*. In *Gender and emotion: Social psychological perspectives*, (2000).

[173] J. Nielsen. *Quantitative studies: How many users to test?* In *Alertbox, June*, volume 3, (2006).

[174] Q. McNemar. *Note on the sampling error of the difference between correlated proportions or percentages*. In *Psychometrika*. Springer, (1947).

[175] A. L. Edwards. *Note on the "correction for continuity" in testing the significance of the difference between correlated proportions.* In *Psychometrika*. Springer, (1948).

[176] J. Kirakowski and B. Cierlik. *Measuring the usability of web sites.* In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, volume 42, pages 424–428. SAGE Publications Sage CA: Los Angeles, CA, (1998).

[177] J. Sauro. *SUPR-Q: a comprehensive measure of the quality of the website user experience.* In *Journal of usability studies*, volume 10.2, pages 68–86, (2015).

[178] J. Kirakowski and M. Corbett. *SUMI: The software usability measurement inventory.* In *British journal of educational technology*, volume 24.3, pages 210–212, (1993).

[179] J. R. Lewis. *IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use.* In *International Journal of Human-Computer Interaction*, volume 7, pages 57–78, (1995).

[180] J. Brooke. *SUS-A quick and dirty usability scale.* In *Usability evaluation in industry*, volume 189, pages 4–7, (1996).

[181] A. Williams. *How to ... Write and analyse a questionnaire.* In *Journal of orthodontics*, volume 30, pages 245–252, (2003).

[182] T. Park. *A comparison of the generalized estimating equation approach with the maximum likelihood approach for repeated measurements.* In *Statistics in Medicine*, volume 12, pages 1723–1732. Wiley Online Library, (1993).

Appendix A: Supportive Material

Self-report method

The participants were asked to report their feelings immediately using AniAvatar; an effective feedback tool for affective states [123]. Figure 1 illustrates the AniAvatar tool.

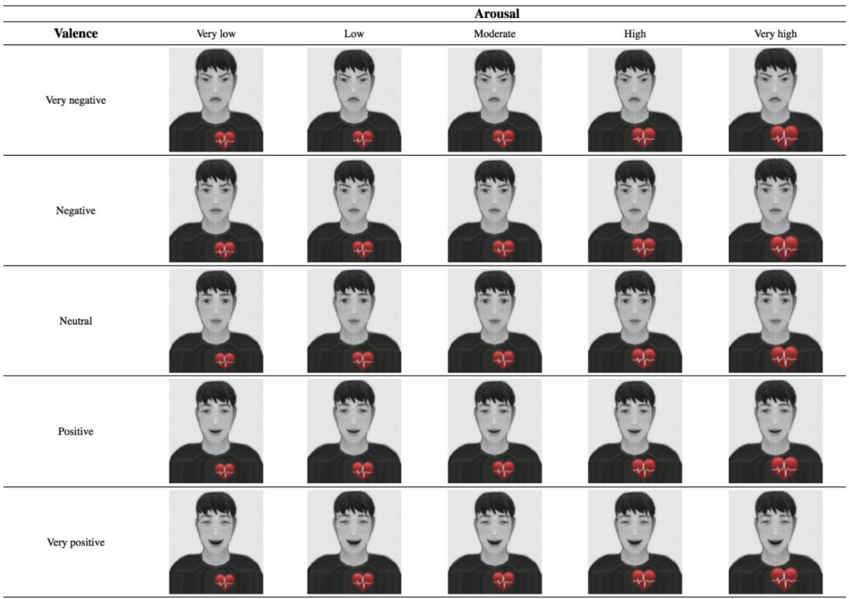


Figure 1: Self-report method

Appendix B: Evaluating the usability of inHarmony

Biofeedback Satisfaction Questionnaire

iAware Mobile Application Satisfaction Questionner

Name:________________________________ Date:_____________________

INSTRUCTION

It is listed below questions about iAware Mobile Application satisfaction level.

a- To what degree you agree with the following statements. Please, circle the letter that corresponds better to your answer to each question. The list of words offers the meaning of each number.

1= strongly Disagree/ 2= Disagree/ 3= Nether Agree nor Disagree/ 4= Agree/ 5= Strongly Agree

Remember that you answers have to show your opinions about the iAware Mobile Application you are using now.

ID	Questions	Answers				
1	iAware Mobile feedback feature are amazing	1	2	3	4	5
2	The App was able to detect my current feelings most of the time	1	2	3	4	5
3	The App feedback is helping me become self-aware of my feelings	1	2	3	4	5
4	Knowing my emotions is helping me to focused on regulate my emotions	1	2	3	4	5
5	I would use this App in a daily bases	1	2	3	4	5
6	I will recommend this App to my friends and family members	1	2	3	4	5

b- Do you like to know your current emotion, Yes or No and Why?

C- Do you like other people to know about your current emotion, Yes or No and Why?

D- Please provide us with any suggestions to improve the Application

Thank you for your participation

Digital Twin Satisfaction Questionnaire

iAware Mobile Application Satisfaction Questionner

Name:___________________________________ Date:____________________________

INSTRUCTION

It is listed below questions about iAware for workplace Mobile Application satisfaction level.

a- To what degree you agree with the following statements. Please, circle the letter that corresponds better to your answer to each question. The list of words offers the meaning of each number.

1= strongly Disagree/ 2= Disagree/ 3= Nether Agree nor Disagree/ 4= Agree/ 5= Strongly Agree

* Remember that you answers have to show your opinions about the iAware Mobile Application you are using now.

ID	Questions	Answers				
1	The application is easy to use	1	2	3	4	5
2	The App was able to detect my current feelings most of the time	1	2	3	4	5
3	The App feedback is helping me become self-aware of my feelings	1	2	3	4	5
4	It is easy to navigate within the application	1	2	3	4	5
5	The feedback from the application is credible	1	2	3	4	5
6	The information from the feedback is trustworthy	1	2	3	4	5
7	How likely are you to recommend this application to a friend or colleague?	1	2	3	4	5
	I will likely reuse the application in the future	1	2	3	4	5
9	I find the application to be attractive	1	2	3	4	5
10	The application interface has a clean and simple presentation	1	2	3	4	5

b- What are two things about the application that you really liked?

c- What are two things about the application that you didn't like?

Thank you for your participation

Appendix C: Burnout Self-report Questionnaire

Burnout Clinical Subtype Questionnaire(BCSQ-12)

"Burnout Clinical Subtype Questionnaire — English" (BCSQ-12)

The following is a series of statements indicating experiences that may occur at work. Read each statement carefully and mark with an X the option that best represents how you feel, what you do and what you think about your work. There are no right or wrong answers. Please **DO NOT LEAVE ANY STATEMENT UNANSWERED.**

	Totally disagree	Strongly disagree	Disagree	Unsure	Agree	Strongly agree	Totally agree
1. I think the dedication I invest in my work is more than what I should for my health	O	O	O	O	O	O	O
2. I would like to be doing another job that is more challenging for my abilities	O	O	O	O	O	O	O
3. When things at work don't turn out as well as they should, I stop trying	O	O	O	O	O	O	O
4. I neglect my personal life when I pursue important achievements in my work	O	O	O	O	O	O	O
5. I feel that my work is an obstacle to the development of my abilities	O	O	O	O	O	O	O
6. I give up in response to difficulties in my work	O	O	O	O	O	O	O
7. I risk my health when I pursue good results in my work	O	O	O	O	O	O	O
8. I would like to be doing another job where I can better develop my talents	O	O	O	O	O	O	O
9. I give up in the face of any difficulties in my work tasks	O	O	O	O	O	O	O
10. I overlook my own needs to fulfil work demands	O	O	O	O	O	O	O
11. My work doesn't offer me opportunities to develop my abilities	O	O	O	O	O	O	O
12. When the effort I invest in work is not enough, I give in	O	O	O	O	O	O	O

Correction algorithm:

The 'overload' dimension is made up of items 1, 4, 7, 10.

The 'lack of development' dimension is made up of items 2, 5, 8, 11.

The 'neglect' dimension is made up of items 3, 6, 9, 12.

The answers are scored between 1 (totally disagree) and 7 (totally agree).

“Cuestionario de Subtipos Clínicos de Burnout — French” (BCSQ-12)

A continuación se presentan una serie de enunciados que indican vivencias que puede experimentar en el trabajo. Lea cada frase con atención y señale con una X la opción que mejor represente cómo se siente, lo que hace o lo que piensa respecto a su actividad laboral. No existen respuestas correctas o incorrectas. Por favor, **NO DEJE NINGUNA RESPUESTA SIN CONTESTAR**.

	Totalmente en desacuerdo	Muy en desacuerdo	En desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	Totalmente de acuerdo
1. Creo que invierto más de lo saludable en mi dedicación al trabajo	O	O	O	O	O	O	O
2. Me gustaría dedicarme a otro trabajo que planteara mayores desafíos a mi capacidad	O	O	O	O	O	O	O
3. Cuando las cosas del trabajo no salen del todo bien dejo de esforzarme	O	O	O	O	O	O	O
4. Descuido mi vida personal al perseguir grandes objetivos en el trabajo	O	O	O	O	O	O	O
5. Siento que mi actividad laboral es un freno para el desarrollo de mis capacidades	O	O	O	O	O	O	O
6. Me rindo como respuesta a las dificultades en el trabajo	O	O	O	O	O	O	O
7. Arriesgo mi salud en la persecución de buenos resultados en el trabajo	O	O	O	O	O	O	O
8. Me gustaría desempeñar otro trabajo en el que pudiera desarrollar mejor mi talento	O	O	O	O	O	O	O
9. Abandono ante cualquier dificultad en las tareas de mi trabajo	O	O	O	O	O	O	O
10. Ignoro mis propias necesidades por cumplir con las demandas del trabajo	O	O	O	O	O	O	O
11. Mi trabajo no me ofrece oportunidades para el desarrollo de mis aptitudes	O	O	O	O	O	O	O
12. Cuando el esfuerzo invertido en el trabajo no es suficiente, me doy por vencido	O	O	O	O	O	O	O

Algoritmo de Corrección:

La dimensión ‘sobrecarga’ está constituida por los ítems nº: 1, 4, 7, 10.

La dimensión ‘falta de desarrollo’ está constituida por los ítems nº: 2, 5, 8, 11.

La dimensión ‘abandono’ está constituida por los ítems nº: 3, 6, 9, 12.

Las respuestas se valoran de 1 (totalmente en desacuerdo) a 7 (totalmente de acuerdo).

Appendix D: Ethics Approvals

02/11/2018

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number	H-10-18-1211
Titre du projet / Project Title	inHarmony: A Digital Twin Based Emotion-Aware Recommender System
Type de projet / Project Type	Projet indépendant d'étudiant / Independent student project
Statut du projet / Project Status	Approuvé / Approved
Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)	02/11/2018
Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)	01/11/2019

Équipe de recherche / Research Team

Chercheur / Researcher	Affiliation	Role
Amani ALBRAIKAN	École de science informatique et de génie électrique / School of Electrical Engineering and Computer Science	Chercheur Principal / Principal Investigator
Abdulmotaleb SADDIK	École de science informatique et de génie électrique / School of Electrical Engineering and Computer Science	Superviseur / Supervisor

Conditions spéciales ou commentaires / Special conditions or comments

550, rue Cumberland, pièce 154
Ottawa (Ontario) K1N 6N5 Canada

550 Cumberland Street, Room 154
Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

02/11/2018

Université d'Ottawa

Bureau d'éthique et d'intégrité de la recherche

University of Ottawa

Office of Research Ethics and Integrity

Le Comité d'éthique de la recherche (CÉR) de l'Université d'Ottawa, opérant conformément à l'*Énoncé de politique des Trois conseils* (2014) et toutes autres lois et tous règlements applicables, a examiné et approuvé la demande d'éthique du projet de recherche ci-nommé.

L'approbation est valide pour la durée indiquée plus haut et est sujette aux conditions énumérées dans la section intitulée "Conditions Spéciales ou Commentaires". Le formulaire « Renouvellement ou Fermeture de Projet » doit être complété quatre semaines avant la date d'échéance indiquée ci-haut afin de demander un renouvellement de cette approbation éthique ou afin de fermer le dossier.

Toutes modifications apportées au projet doivent être approuvées par le CÉR avant leur mise en place, sauf si le participant doit être retiré en raison d'un danger immédiat ou s'il s'agit d'un changement ayant trait à des éléments administratifs ou logistiques du projet. Les chercheurs doivent aviser le CÉR dans les plus brefs délais de tout changement pouvant augmenter le niveau de risque aux participants ou pouvant affecter considérablement le déroulement du projet, rapporter tout évènement imprévu ou indésirable et soumettre toute nouvelle information pouvant nuire à la conduite du projet ou à la sécurité des participants.

The University of Ottawa Research Ethics Board, which operates in accordance with the *Tri-Council Policy Statement* (2014) and other applicable laws and regulations, has examined and approved the ethics application for the above-named research project.

Ethics approval is valid for the period indicated above and is subject to the conditions listed in the section entitled "Special Conditions or Comments". The "Renewal/Project Closure" form must be completedfour weeks before the above-referenced expiry date to request a renewal of this ethics approval or closure of the file.

Any changes made to the project must be approved by the REB before being implemented, except when necessary to remove participants from immediate endangerment or when the modification(s) only pertain to administrative or logistical components of the project. Investigators must also promptly alert the REB of any changes that increase the risk to participant(s), any changes that considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project or the safety of the participant(s).

Riana MARCOTTE
Responsable d'éthique en recherche / Protocol Officer
Pour/For **Daniel LAGAREC** Président(e) du/ Chair of the **Comité d'éthique de la recherche en sciences sociales et humanités / Social Sciences and Humanities Research Ethics Board**

550, rue Cumberland, pièce 154 550 Cumberland Street, Room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics