Towards Physiologically-Responsive Interactive Garments with Machine Learning Techniques

by

Victoria Armstrong

A thesis submitted to the
School of Computing
in conformity with the requirements for
the degree of Master of Science

Queen's University
Kingston, Ontario, Canada
May 2022

Abstract

Emotional experiences shape our lives every day. Negative emotions can impact not only our mood but also our biological signals, overall health, and wellness, especially if they are not addressed. Emotion-regulation and self-care techniques, such as meditation and exercise, can help to alleviate these emotions, but we have to remember to actively engage in them. Compression applied to the body, called Deep Pressure Stimulation, has also been shown to help suppress reactions from our nervous system under stress. In this work, we address the challenges of emotion regulation when experiencing negative emotions while doing desk work. To accomplish this, we custom-built two interactive jackets that have a removable, embedded microcontroller, sensors, and airbags. The airbags are used to apply compression to the sides when a user presses a button. 12 participants interacted with the jackets during a user study and were interviewed after. Data collected from 8 of these 12 participants during the user study was used to train 3 machine learning models: Logistic Regression, Support Vector Machine and XGBoost. Over 4 different testing conditions, XGBoost proved to be an efficient and effective predictor of when users choose to turn on the pump. Coupled with interview data from participants exploring desires for slow interfaces and automatic actuation, we have established a foundation for individualized interactive garment analytics using customized models for affect prediction.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The experience of negative emotions is an inescapable part of the human experience. Despite this, prolonged experiences of stress, anxiety, or distraction can have a significant impact on our lives. There are negative health implications from sustained stress and anxiety [10, 16], as well as impacts on daily habits and social lives. Recent years have shown an emergence in self-care techniques, including movement reminders, exercise, breathing, meditation and weighted blankets. Mobile apps that encourage and promote emotion regulation have been developed to assist individuals in managing emotions, as well as wearables, like smartwatches. While some find these beneficial, others don't always enjoy having to interact with screens regularly. Whether it be a phone screen or a wrist-worn device, wearers have shown a desire to interact with slower interfaces, such as an e-textile [19].

Emotion regulation techniques require individuals to play an active role in their self-care practices. Individuals must first recognize the emotion they are experiencing, then decide which type of emotion-regulation technique they would like to engage in. Lastly, they must actively choose to begin this activity. This poses a number of challenges. First, it can be difficult to name feelings that we are experiencing, making it a challenge to decide when to engage in regulation practices. This impacts the efficacy of technological solutions to help manage emotions. Here, the application of machine learning can help to identify emotions that may not be perceived by the wearer. There is also expenditure of mental effort required to choose to engage in a task, regardless of whether the task may be beneficial in the long run.

One method known to help regulate emotions is Deep Pressure Stimulation (DPS) [51]. DPS involves the application of pressure to the body, which helps suppress the nervous system and can help relieve feelings of stress and anxiety [51]. This is the theory behind why weighted blankets or a firm hug can be helpful. In both these cases, an individual needs to actively decide to use pressure to help with their negative emotions. Additionally, it can require the use of another person. The employment of machine learning poses a solution for deciding when to apply pressure.

It has been shown that physiological data, like heart rate, skin temperature, and acceleration are indicators of affect. With mobile sensors becoming prevalent, it is now easy to collect this data directly from a wearer. Machine learning algorithms can be applied to this data to predict affect, including stress [27, 17]. While current techniques are able to achieve high accuracy [27], it can be a challenge generalizing training across participants, since physiological data can be user-specific [57]. Successful methods can also be high resource, including a large number of trainable parameters [12, 56], leading to increased memory and training time.

1.2 Problem

To address the challenges above, this work leverages the use of DPS and machine learning techniques. Two outer jackets were created with sensors that can be snapped in, and airbags attached to the sides. At the press of a button, these airbags inflate to apply pressure to the sides of the body. A wrist-worn heart rate sensor, and built-in microcontroller sensors collect data from the wearers. A user study followed by interviews explored the use of the garments while completing laptop tasks. After user experiments, the biological data is processed and used to train machine learning models. This work is done to address the problems of emotion identification and leverages the use of DPS and slow interfaces to explore how participants respond to emotions. Additionally, we explore the training of machine learning models that are low-resource and participant-specific to facilitate future work for the on-board prediction of emotions.

1.3 Research Questions

We strive to answer the following three questions.

1. How can interactive garments support wearers in managing daily stress?
2. Can we train machine learning models to predict when wearers choose to activate garment compression?
3. If so, is this training and inference able to work in a low-resource setting?

1.4 Contributions

In this work, two jackets with embedded sensors and airbags for compression were created through an iterative design process. Using these two garments, a user study of 12 participants was conducted. Participants wore a garment that applied pressure to the sides at the press of a button. Participants explored when they would like to apply or remove pressure, and reflected on their experiences. From these reflections, 6 themes were created. These themes validate existing work and demonstrate a desire for slow interfaces that include smart automation.

Additionally, data from 8 of 12 participants was used to train 3 predictive models: Logistic Regression, Support Vector Machine, and XGBoost. Different training and testing splits were explored leading to rich insight on the training data requirements, as well as current label and next step predictions. We observed that XGBoost performed the best across all 4 testing conditions. Additionally, XGBoost performs well when restricted to small portions of the training data. We observed minimal differences in accuracy when predicting current time steps and the next time step.

1.5 Organization of Thesis

Chapter 2 provides background information on related works, including their strengths and limitations. Furthermore, Chapter 2 presents general information on e-textiles, statistical measures, model performance metrics, and the models used in this research. Chapter 3 is broken into three distinct sections. First, the crafting of the physical prototypes is discussed. Next, the user study is outlined. Finally, the process for the application of predictive models is presented. Chapter 4 presents the results of the user study, followed by the results from training different models using data collected during the user study. Lastly, Chapter 5 provides a reflection on the research, including answers to the research questions listed above, limitations, and areas for future work. The appendix contains interview questions.

Chapter 2

Background

2.1 Preliminaries & Notation

2.1.1 Interactive Garments

The field of Human Computer Interaction (HCI) is a large and multidisciplinary field that explores how users interact with technology. A wide range of sub fields are encompassed within HCI, such as UX design, gaming, educational computing and interactive garments. Interactive garments refer to clothing items that wearer's can interface with. In their 2012 paper, Buechley and Perner-Wilson explored the use of carving, sewing, and painting as new methods for constructing electronics. Since then, there has been a great exploration into the field interactive garments and e-textiles [47]. Interactive garment actuation has been shown to include folding and crinkling [42], flower blossoming movement patterns [6], thermochromic change [61], and pneumatic actuation [33, 45, 22]. Commercial manufacturers have started to explore the field of interactive garments, such as Google ATP's Project Jacquard, where interactive clothing, bags, and shoes were created in collaboration with popular brands, like Levi's, Adidas and Samsonite [48].Beyond academic publications, there is a wealth of online resources posted publicly where creators can collaborate. E-textile microcontroller manufacturers, like Adafruit[1], provide tutorials, and KobaKant[2] provides resources for creating crafted sensors and actuators with e-textiles.

To control actuation, as well as sensing, in interactive garments and e-textiles, specific microcontrollers have been created that can be sewn into fabrics through large, sewable connection holes using conductive thread. Arduino and Adafruit are two major manufacturers of such boards, like the Flora, Gemma, and Circuit Playground. Relevant to this research, the Bluefruit Circuit Playground [31] contains embedded sensors for acceleration and temperature, as well as two onboard buttons, and a Bluetooth module [31]. This data can be accessed using built-in variables. Additionally, the Bluefruit Circuit Playground supports the use of TensorFlow Lite [1], which is a framework that allows for trained machine learning models to be converted to C++ and used for on-device prediction.

Other electrical components are also required in the construction of interactive garments, including transistors, resistors, and diodes. These components are small in weight and size, and contain metal leads that can be soldered together, or bent and sewn onto. Transistors are used to control the flow of electrical current. Resistors limit the flow of electrical current and come in different resistances. Diodes restrict the flow of electrical current.

2.1.2 Thematic Analysis

Presented by Virginia Braun and Victoria Clarke in 2012, Thematic Analysis formalizes the extraction of themes from qualitative data [7]. Thematic Analysis, sometimes

[1] `https://learn.adafruit.com/groups`
[2] `http://kobakant.at/DIY`

referred to as TA, emphasizes the active role played by a researcher to create themes from observations in the data, rather than the commonly used sentiment *themes emerged*. The process of TA is flexible and involves first creating codes for the data. This is an iterative process, whereby code names evolve and change as more passes of the data are executed. Following the coding of the data, the codes are grouped by commonalities. Based on these commonalities, themes are constructed. TA can be performed using two approaches: latent and semantic [7]. In the semantic approach, we move from surface-level to construct deeper meaning. In the latent approach, we enter the analysis with underlying ideas and assumptions and work backwards. To perform a qualitative analysis of our work, a thematic analysis was conducted following a latent approach.

2.1.3 Quantitative Analysis

The following outlines the relevant information required to understand our method for predicting user actions and evaluating their accuracy.

Classification Techniques

In this work, three different predictive models are trained using data generated from e-textile sensors. The models used are outlined below.

Logistic Regression Logistic regression is a classification model that predicts the probabilities of a dependent variable using the logistic function [44]. Logistic regression is used when output values are discrete, rather than continuous. To perform a binary classification, as is the case for this research, a cost function is optimized according to the log-likelihood of the probability of predicting the correct class. In practice, this model is implemented in Python using the Scikit-learn linear model class LogisticRegression [44].

Support Vector Machine Support Vector Machine (SVM) is a machine learning technique that can perform both linear and multi-dimensional classifications. SVM finds a hyper-plane that creates a boundary between the data. In 2-dimensional space, this hyperplane is a line, and there are various techniques to use for multi-class problems. A wide variety of applications exist for SVM, including but not limited to, bioinformatics, facial recognition, and credit card fraud detection [11]. This model can be implemented in Python using the Scikit-learn SVM SVC class [44].

XGBoost Extreme Gradient Boosting, or XGBoost, is a supervised learning algorithm that uses boosted trees to predict an output given a series of input values [13]. Based on traditional Tree Boosting, this method is a lower resource alternative to other machine learning algorithms [14]. The authors measure effectiveness on a rank task, insurance claim data, high energy physics data, and a click log data set [14]. XGBoost has also been applied to time series forecasting. Abbasi et al. use XGBoost to extract features and predict electrical load for the next time step [2]. This model can be implemented in Python using the XGBoost library [13].

Statistical Measures

To explore and analyze the data collected from the jackets, it is imperative that we use statistical measures to give insights into the data. The following metrics were used when exploring the data collected in this study, which will be discussed further in Sections 3 and 4.

Mean, median and mode are measures used to calculate the average of a dataset.

The arithmetic mean is the sum of all elements divided by the total number of elements. The median is the middlemost element in the data when organized from lowest to the highest value. The mode is also a measure of average which counts the number of occurrences of each unique element and reports the element with the highest frequency. These three measures can be automatically calculated using methods in the Pandas DataFrame class [43].

There are a number of metrics that we can use to understand the shape of our data distribution. First, we can observe our highest (max) and lowest (min) values in the dataset. Subtracting the two gives the range of the dataset values. This is part of how we can understand the spread of data. We also look at the standard deviation and variance to better understand the spread of our data.

Standard deviation represents how far away data values are from the mean. A lower standard deviation indicates that the data is closer to the mean, whereas a large standard deviation indicates a greater spread in the data. This can be calculated automatically in Python using built-in Pandas libraries [43]. Variance is calculated by squaring the standard deviation and represents the average dispersion of the data with respect to the mean. Similarly to standard deviation, this can be calculated directly from a Pandas DataFrame [43].

There are further statistics that give more insight into how data is distributed, including skew and kurtosis which give indicators on what the peak and tails of our data look like. Skew represents how asymmetrical data is. It can be positive or negative, corresponding to a shift in data to the left or right respectively. A perfectly symmetrical data set would have a skew coefficient of 0, however, the converse is not necessarily true. Skew is calculated using the formula below and can be automatically calculated using the Pandas library [43]. Kurtosis indicates whether data is more concentrated at the tail end or the peak. Kurtosis values range from 1 to positive infinity, and a kurtosis value greater than 3 indicates that data is taller and has thicker tails than a normal distribution. In general, the higher above 3 the kurtosis is, the taller the peak and fatter the tails. Similar to other methods mentioned, kurtosis can be automatically extracted using the Pandas library [43].

2.1.4 Measuring Model Performance

When evaluating a predictive model, it is important to observe a variety of measures to understand the model's performance. Using built-in metrics from the Scikit-learn metrics class [44], we can build a stronger understanding of model performance beyond accuracy.

Overall accuracy is calculated using the true value and predicted value, calculating the number that were predicted correctly. Going one step further, we can use a confusion matrix to observe the types of predictions made. In the case of binary classification, we can have a true positive, true negative, false positive, or false negative predictions. These values can be calculated using the true labels and predicted labels with the built-in Scikit-learn metrics method, confusion_matrix [44].

Using the values from the confusion matrix, we can calculate precision and recall. Precision reports the number of true positive values out of all predicted positive instances. Recall reports the number of true positive values out of all positive examples. Both of these are shown in Formulae 2.1 and 2.2 where TP represents the number of true positives, FP represents the number of false positives, and FN represents the number of false negatives. In the case of a perfect model, false positive and false

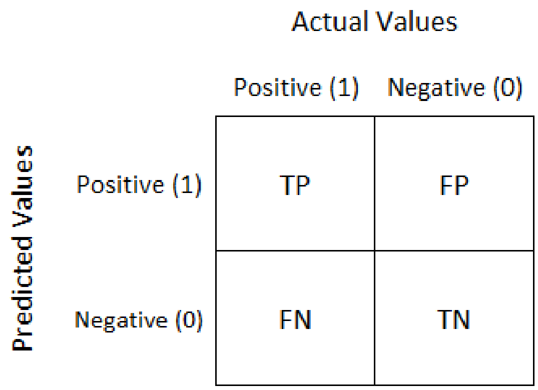

Figure 2.1: Confusion matrix for a binary classification task

negative values would be 0, giving both a precision and recall value of 1.

$$Precision = \frac{TP}{TP + FP} \tag{2.1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2.2}$$

The F1-Score is a way of combining the precision and recall values for a trained model. It is a weighted average of precision and recall, with a perfect score of 1 and a worst-case score of 0. The Scikit-learn metrics method f1_score outputs a value for each label in the classification [44] and is calculated as shown in Formula 2.3.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \tag{2.3}$$

2.2 Related Work

2.2.1 Emotions

Emotions can be difficult to describe and formally classify. In their 1979 publication *Affective Space is Bipolar*, Russell defines emotions on two axes: valence and arousal [52]. Valence refers to attraction (positive) or aversion (negative) towards a stimulus [52]. Arousal also ranges from negative to positive and refers to perceived intensity. In a subsequent paper, Russell classifies 28 emotions using the bipolar model [53]. Figure 2.2 shows different emotions classified within Russel's model. Of particular interest are the emotions in quadrants 2 and 3, where a low valence indicates negative emotions. Stress, although not listed in 2.2, exists in the upper left quadrant. Saini et al. argue that using this 2D model of affect facilitates the use of machine learning methods for affect prediction because it gives a target mapping of the emotions, rather than a qualitative expression that is difficult to capture [54].

In general, the definition of stress in vague is most literature and can be experienced physiologically and psychologically. Kemeny avoids using the term stress entirely in their paper *The Psychobiology of Stress* and instead distinguishes between stressors and distress [35]. The former being events that threaten a goal, and the latter being negative psychological feelings. Stressful periods can cause changes in levels of hormones present in the body and impact the nervous system [36, 35]. These physiological symptoms can have lasting effects; prolonged periods of stress have been shown to increase the risk of severe illnesses, like coronary heart disease [10]. Long working hours can also have an impact on the musculoskeletal system causing an increase in neck, shoulder, arm and back pain [16] which may interact with feelings

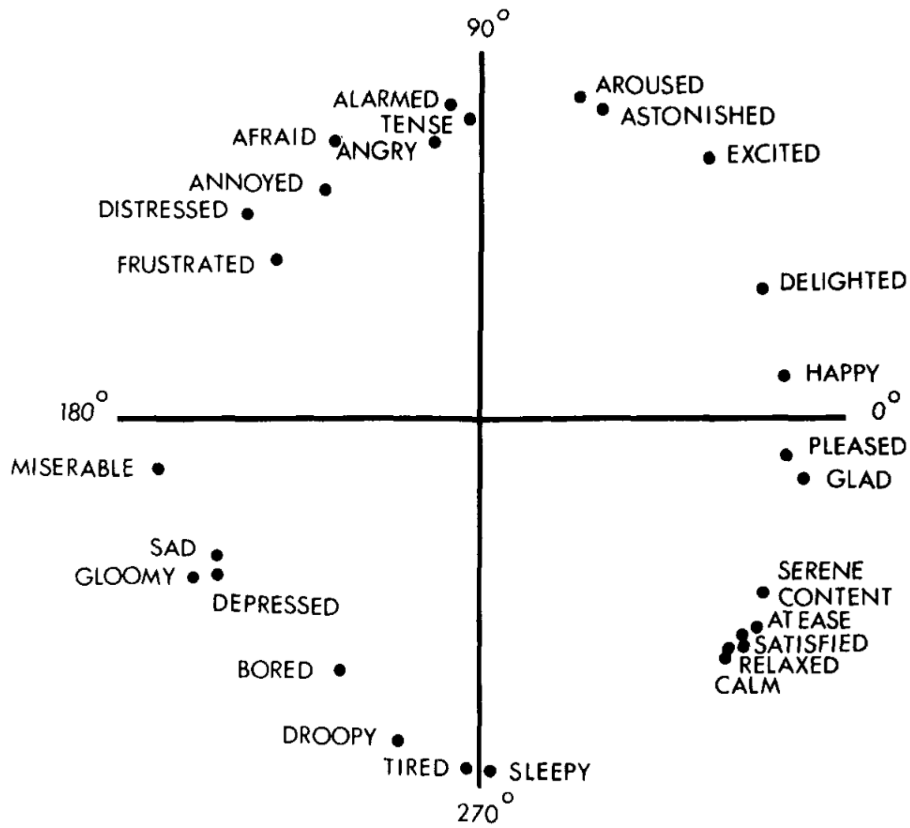

Figure 2.2: Russell's 1980 Circumplex Model of Affect

of stress. While having a physical effect on the body, stress can also have an impact on social interactions, work and study performance, and the completion of daily tasks. Heightened stress and anxiety can also be associated with different mental disorders like Generalized Anxiety Disorder, Attention Deficit Hyperactivity Disorder, Attention Deficit Disorder, Autism Spectrum Disorder and Depression [10].

Given the effects of stress on the body, biological data can indicate when an individual is experiencing stress. One accurate indicator of stress is heart rate [10]. From this, heart rate variability can be extracted [10]. Heart rate can be measured in a variety of ways from users, such as by measuring the distances between peaks in a blood volume pressure (BVP) signal collect through photopletysmography (PPG). Other physiological indicators of stress include electrodermal activity [46], skin temperature [8], acceleration [64] and electrocardiogram [55].

PPG is a non-invasive technique for measuring an individual's heart rate. A PPG signal can be collected continuously and on the go, without the use of a reference signal. Because of the aforementioned factors, PPG has developed popularity in wrist-worn health monitoring devices in the last decade [39]. A light-emitting diode (LED) coupled with a photo-receiver measures the amount of reflected infrared light when the LED is pressed against the skin. Plotting the amount of light measured by the photo-receiver produces a wave, wherein the peaks can be extracted using a peak-detection algorithm to extract the heart rate.

2.2.2 Sensing Heart Rates Using Wearables

There are a number of commercially available smart devices that are capable of detecting heart rate, and other physiological data. Table 2.1 summarizes different sensors identified in the works relating to wearable stress detection and e-textiles. As is evident in the table, there is a range of different sensor readings and a wide variety of price points. Some work relies on data collected from smartwatches that is then relayed to the e-textile [20, 64, 63] , while others wore embedded sensors [60, 49, 29]. One study crafted an ECG sensor themselves, rather than relying on preexisting sensors [66]. Lastly, one study demonstrated the feasibility of using a phone camera to extract PPG data [15]. Data collected from devices listed above that are not already embedded in e-textiles, could be used to collect data and passed to the e-textile over Bluetooth connectivity.

Product	Measurement	Price
Shimmer GSR+ [30]	GSR, PPG	unlisted
Empatica E4 [21]	PPG, EDA, Acceleration	$1690.00
Pulse Oximeter [3]	BPM	$49.95
FitBit [25]	HR, HRV, BPM, Skin Temperature	$169.95
Garmin [28]	Pulse Ox, HRV	$479.99
Apple Watch [4]	Blood Oxygen, ECG, HR, HRV	$529.00

Table 2.1: Commercially Available Wearable Data Sensors

Beyond smartwatches, there is an overlap in the use of biological signals in interaction design. A number of works highlight the desire to explore how stress impacts our bodies: Yu et al. used a finger-clasp heart rate monitor connected to an interactive art display to allow onlookers to see a visual representation of their pulse and observe changes in their heart rate [65]. Similarly, the creators of Heart Waves used a crafted bracelet to extract a wearer's heart rate and use this to control the flow of water in a miniature sculpture [23]. As the heart rate increases, the flow of water becomes more rapid [23]. These works serve as methods for participants to observe their heart rate in a passive and slow fashion, providing sensory indicators of their heart rate, rather than simply the interpretation of a number. There is no deliberate emotional regulation technique applied here, however, awareness and sensory focus can have a meditative effect.

Mauriello et al. created three athletic shirt prototypes with flexible displays to show data on runs, including heart rate. Participants noted that this helped them stay on track and they didn't find the display obtrusive while running [38]. Ganti et al. created a rudimentary microcontroller circuit for heart rate data collection [27], and Dabby et al. created three fully-integrated heart rate monitoring garments in the form of a compression shirt, compression shorts and a sports bra that are also washable [18]. These works demonstrate interest and applications in the area of heart rate monitoring, including the use of e-textiles to collect heart rate data [38, 18].

2.2.3 Compression in Wearables

Muthukumarana et al. developed and validated a method to attach small tiles with embedded shape-memory alloy wire that actuate. As current is applied, the tiles move closer or further together. Although not yet explored, this could see applications in compression. Pohl et al. explore the use of pneumatic actuation on the wrist as a tactile notification, given that vibrotactile notifications can be easily ignored [45]. In this case, a user study was not completed to validate the method. Endow et al. developed a toolkit for prototyping with compression to either apply force to the body or sense when an external stimulus has compressed the interface [22]. Amongst other contributions in the toolkit, the authors propose a silicone air bladder that can be fabricated in under 30 minutes [22]. This bladder can be inflated with an automated air pump and applies force directly to the desired body part. Their method of fabrication uses 3D printing, allowing designers to create customized bladder shapes that are tailored to their design [22]. Given silicone's waterproof qualities, this is desirable for a wearable garment that will be undergoing general wear and tear.

2.2.4 Emotion Regulation Through Actuation

Numerous emotion regulation techniques exist, both formally defined and self-discovered methods. In terms of negative emotions that may be experienced during school or work life, such as stress, anxiety or distraction, techniques can include controlled breathing, movement reminders and the use of weighted blankets. Researchers have explored triggers and solutions to workplace stressors, and broadly classify solutions as problem-oriented or emotion-oriented [24]. Problem-oriented solutions focus avoidance or addressing of stressors [24], whereas emotion-oriented solutions focus on mindfulness and relaxation [24].

A number of works have focused on presenting notifications and reminders in a calm fashion without the use of screens. Devendorf et al. demonstrated the shortcomings of screens and why individuals may prefer a tactile user interface over a screen [19]. Nabil et al. created an actuating table runner and found that a richer and unique spatial experience was presented through the use of fabric. Current literature has explored the use of tangible shape-changing break-reminders and notifications [58, 32]. While the effectiveness of these reminders was demonstrated, these reminders rely solely on visual input, which is already stimulated when performing tasks on a computer or looking at a phone screen. Other works have explored visual notifications that have a tactile experience as a side-effect. Speer et al. used a combination of app-based and visual shape changes to explore emotion regulation, like breathing or exercise, in a group of children [59]. This technique also relies on visual input and required significant training and instruction for teachers and users [59]. Fransén Waldhör et al. explored colour change and light emission as visual notifications in fabric [26]. They also explored thermo (heat) feedback as a byproduct of thermochromic notifications (heat triggered colour change) but did not conduct a user study [26]. Spikey Starfish is a cigarette handbag that morphs by inflating when someone tries to take a pack of cigarettes. This shape change does apply some tactile feedback to the wearer, however, it is the visual change in the bag that communicates feelings of anger [37].

Deep pressure stimulation has been shown to impact physiological changes in the body as a result of stress and involves applying pressure to the body [51]. In 2021, Jung et al. published work on bringing awareness of breathing through deep pressure stimulation [33]. They found that pressure applied to the torso allowed them to become more conscious of their breathing. Participants also noted that when the breathing pattern applied to their torso did not match their current breathing pattern, it caused a noticeable discomfort that prompted a change in their breathing to match the applied pressure [33]. This indicates the use of deep pressure stimulation to regulate or change breathing patterns that may become altered by periods of stress.

In order to discover how customized haptic vibration plays a role in anxiety management, Umair et al. engaged participants in an anxiety-inducing task and allowed participants to create their own calming vibrotactile and temperature notifications [62]. Participants could choose different beat per minute (bpm) ranges for vibration notifications on their wrist. A number of participants picked 30 beats per minute, despite it being significantly below the average, adult resting heart rate, as they felt it forced them to slow their breathing, and consequently, their heart rate [62]. This work differs from previous works exploring the visual sense and focuses on biological processes rather than notifications or break reminders.

2.2.5 Machine Learning for Emotion Classification

Recent human emotion recognition has been able to classify emotional states using a variety of different types of data, including speech/text analysis, facial analysis, movement and body position, and biometric data [27]. A number of different affective states have been predicted with varying accuracies including joy, sadness, anger, calm, scared, and valence and arousal [17]. Classifiers that have been used include

SVM, Random Forest and more intricate deep learning techniques [17]. At Shibaura Institute of Technology, facial skin temperature was used in conjunction with heart rate variability to predict fear, joy and sadness with 88.75% accuracy using a simple neural network implementation [34].

Simple models have been used to classify emotions. Rakshit et al. use features derived from PPG sensors to accurately classify sad, happy and neutral affect with an SVM [50]. The creators of WESAD use a number of classifiers to predict amusement, stress and baseline affect for participants completing different pressure and time-sensitive tasks [57]. They collected data on heart rate, skin temperature, electrodermal activity, and acceleration and trained models to serve as a baseline for those who use their data set [57]. They observed that within-participant classification outperforms predictions made on models trained with more than one participant, due to unique differences in biological signals across participants [57]. The authors of *Activity Recognition and Stress Detection via Wristband* were successfully able to detect stress from acceleration and a physiological signal from a device worn on the wrist. They were also able to distinguish between events containing high physical activity and heightened stress levels [64].

More sophisticated deep learning techniques have also been employed for emotion classification from biological signals. Chao et al. experimented with loss functions for dimensional emotion detection using a Long Short Term Memory Recurrent Neural Network on a data set containing ECG and EDA data [12]. Sarkar and Etemad take a self-supervised approach using a convolutional neural network to classify emotions from the SWELL and AMIGOS data sets, both of which contain physiological data. While effective, these models contain a large number of trainable parameters, given the nature of the deep learning techniques used. The authors of *Emotions on the Go* create user-independent emotion classifiers using the facial recognition toolkit OpenFace [5] and were able to accurately detect a variety of emotions using a mobile phone. In this case, the models trained are accurate, but are user-dependent and were not tested for generalizability.

Chapter 3

Methodology

Our research is divided into two distinct phases. First, two interactive garment jackets were created for users to wear and explore instances where they feel heightened negative emotions and have a desire for compression, while also collecting physiological data. Second, predictive modelling was used to determine when users decided to engage in an emotion regulation technique. This section breaks down the design process and final prototypes, the user study set up, and the models explored for predictive purposes.

3.1 Garment Prototyping

Two physical prototypes were created. These prototypes are in the form of jackets with airbags on the sides which apply compression at the press of a button, and heart rate monitors on the left wrist. The first jacket, Jacket A, is a women's small and is hand-sewn from black fabric with a faux leather finish. The second jacket, Jacket B, is a men's large and is commercially produced in a taupe cotton blend. This section outlines the design process and culminates in the presentation, as well as the technical specifications, of the two jackets.

3.1.1 Design Decisions

A number of important decisions guided the design of our jackets. Key decisions regarding hardware greatly influenced aspects of the prototype's physical appearance and interactivity. Additionally, a desire for discretion rather than a tech-savvy statement piece also influenced how the final products were designed.

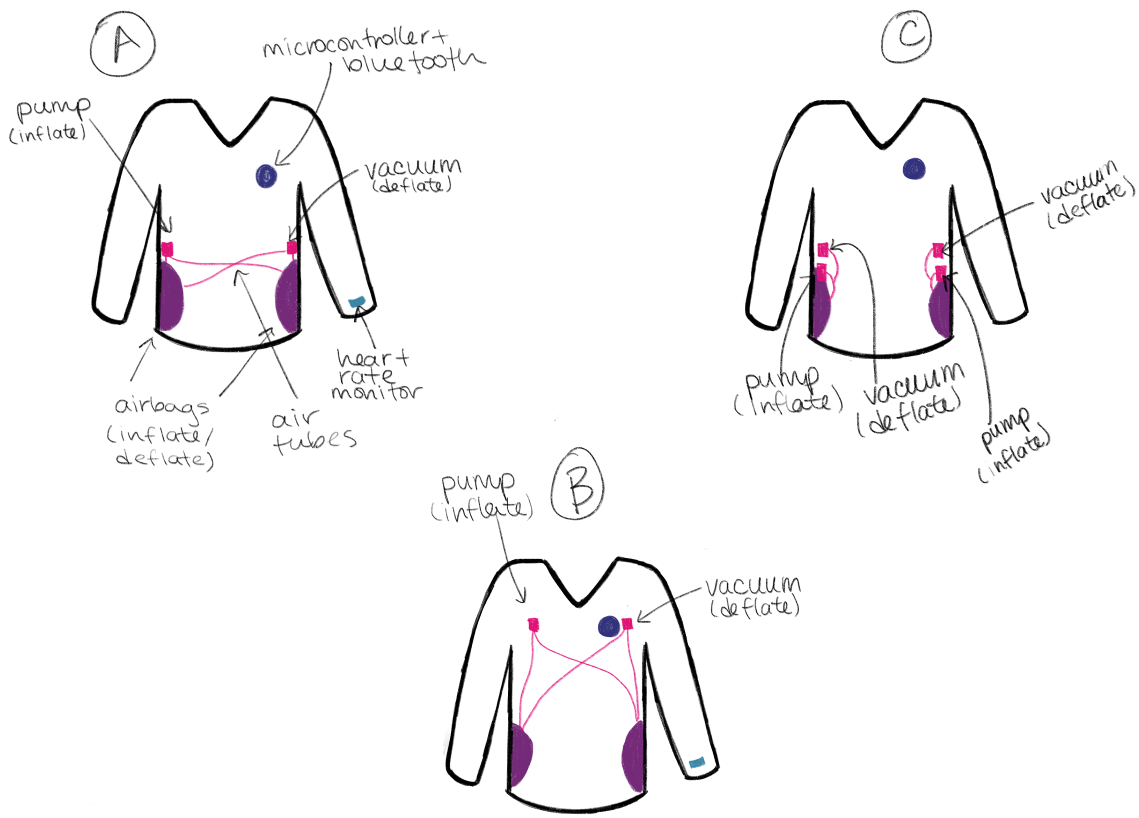

Figure 3.1: This sketch shows initial considerations for air pump placement inside a shirt, created during a brainstorming session early in the design process.

The choice of using inflation of airbags, discussed further on in Section 3.1.1, as the source for pressure actuation led to design decisions made to effectively hide the hardware for the air pumps. These pumps are used to inflate airbags tied to the waist.

The garment design needed to accommodate two 27x45mm air pumps. Additionally, the pump's weight of approximately 100g needed to be supported so that it did not weigh down the fabric and cause unintended changes in the shape and drape of the garment. To combat this, fashion trends were explored. Large, structured belts, would allow the air pumps to be supported around the waist or hips in a position that is familiar, but caused limitations when used in conjunction with airbags also present on the waist. The presence of both could reduce the performance of each other, or result in too much bulk around the waist. Cowl and ruched necklines also presented opportunities to hide the shape of the air pumps, while supporting their weight on the shoulders, but may not have enough fabric to conceal the size of the pumps. Lastly, structured shoulders allowed the shape and weight of the air pump to be neatly concealed on the shoulder, which can be technically more difficult to situate. They are further away from the actuators by the waist, but we deemed this to be the most comfortable and least problematic solution for users. Figure 3.1 highlights different locations on the body where air pumps were initially considered.

In addition to the pumps, the design was influenced by the positioning of a sensor to accurately measure the user's heart rate, given the use of a PPG sensor. As discussed in the previous section, this is a common method for monitoring heart rates in commercially available wearables, like the Apple Watch. In order to get an accurate reading from the PPG sensor, the sensor must be pressed firmly against the skin. This guided the decision to use a long sleeve and some form of cuff to allow the sensor to be squarely placed on the wrist.

Lastly, individuals may not always want to disclose their feelings. Indeed in some situations, it could increase the levels of discomfort that they are experiencing. Additionally, participants may not want to show off that they are wearing a technology enable garment. Therefore, the design of the jackets placed emphasis on ensuring that the jackets mimicked something that fit into everyday fashion. In order to be more inclusive of size and personal preference, we decided to develop two jackets in different styles.

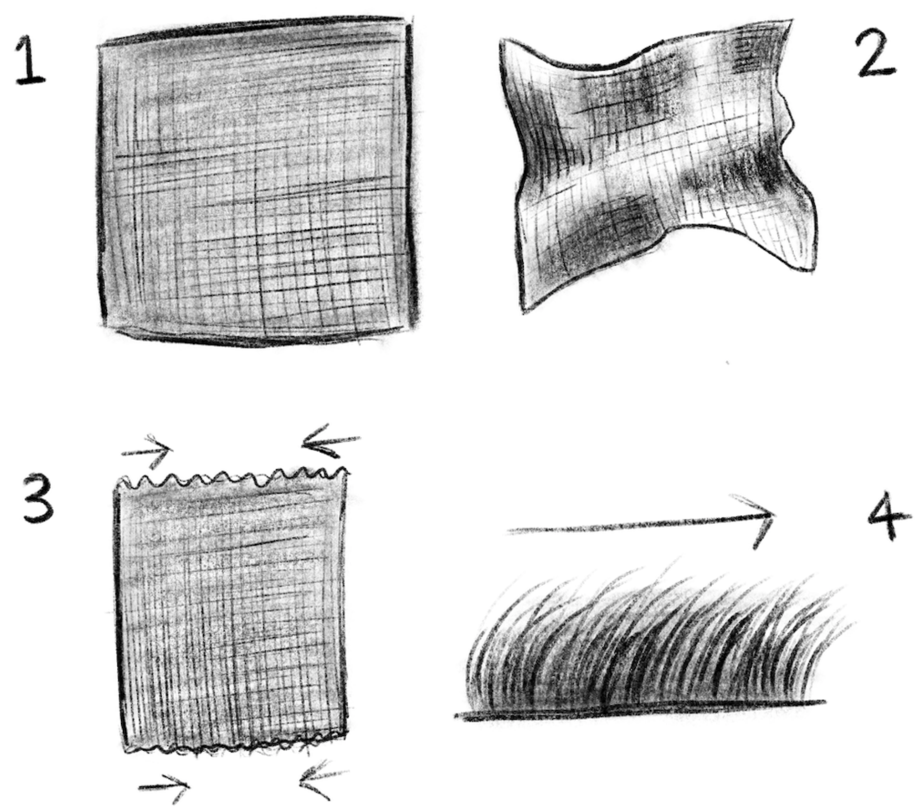

Figure 3.2: Different actuations that were brainstormed early in the creation process: (1) Neutral, (2) Crumpling, (3) Compression, and (4) Brushing.

3.1.2 Initial Iterations

The process of crafting the final jackets was an iterative process, involving sketches, basic fabric prototypes, and the final physical product. The initial sketches focused on different types of actuation that could be envisioned. The types of actuation envisioned served the purpose of calming the feelings of anxiety a participant might be feeling. These possible actuations included compression, scrunching, and brushing, shown in Figure 3.2. Other ideas considered were vibrotactile feedback and light emissions to breathe alongside. Wearers may not always want to broadcast their emotions, so light emissions were eliminated as a possible actuation because of their lack of discretion.

Figure 3.3: Image depicting the airbags inflated (top) state and in the neural state (bottom).

Drawing on inspiration from Deep Pressure Stimulation [51], compression was chosen as the form of actuation. This led to the development of different ways to induce compression, without being restrictive to the wearer. Ideas included Shape Memory Alloy constriction, raised beads pressing on the skin, and airbags. Airbag compression presented the most benefits [33]. Inspired by previous work in [33], we chose to explore pressure applied to the sides. Thick, but flexible, plastic sheets connected to pumps through silicone tubing were chosen as the materials to implement garment compression. Figure 3.3 shows an airbag filled and empty.

One issue that arose when initially crafting the prototype was the fragility of wires when combined with the weight of the air pumps used. The air pumps have more weight to them than other electronic components, at approximately 100 grams. Once sewn in, if they shifted at all, the wires attached to the pump started to fatigue. This led to the wires disconnecting the pump from the circuit. This was solved by cutting out a portion of a foam shoulder pad and placing the pump inside, leaving room for the tubes to exit, and then sewing it shut. This reduce the movement of the pumps while supporting the weight on the shoulders. Figure 3.4 shows the configuration of the pump and shoulder pads.

The air pumps were also difficult to attach to the circuit with conductive thread. The pumps have two posts (one positive, one negative) equally spaced in the middle of the base. This lead to difficultly using conductive thread to connect to the posts. Therefore, wires were used to connect the components. Again, we ran into the issue of metal fatigue. Longer wires which accommodated more room for movement were first tried, however, this still lead to metal fatigue. This lead to multiple iterations of the configuration of the electrical components to ensure that the circuit functioned

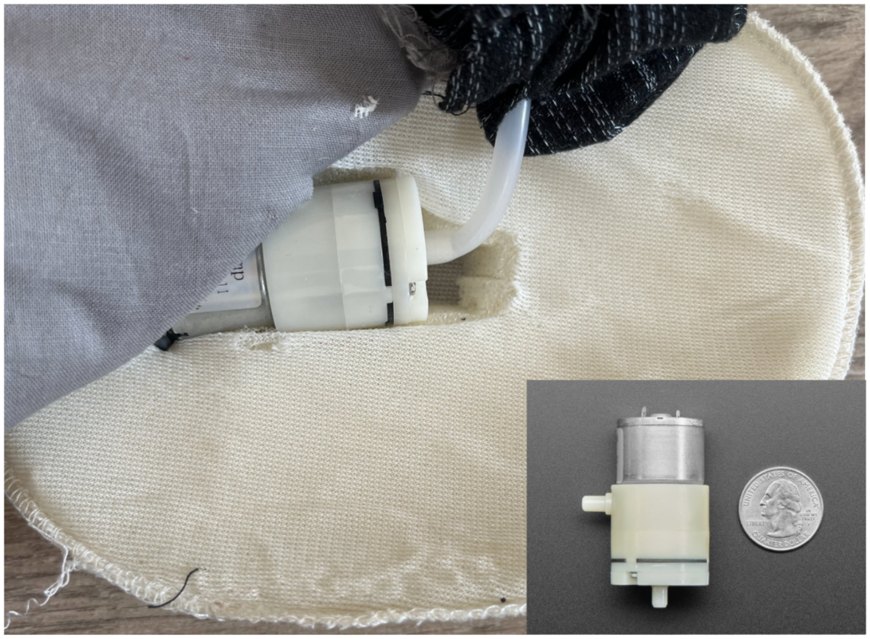

Figure 3.4: A hole was cut in the shoulder pad for the pump to fit inside and was then secured with fabric stitched to covering it. The inset image shows the pump and its size in relation to a quarter.

as it should. Finally, a combination of conductive thread and metal wires were used. Thread was also used to secure the resistors, transitors, and diodes with bent leads down, and then short wires were used to make connections across wires. Using only conductive thread presented issues with reliability within the circuit. Time constraints required a shorter term fix, which is why shorter wires were used rather than exploring conductive thread further. Figure 3.5 shows the initial configurations and the final position.

Initially, each prototype would have electronic components stitched in. However, given the challenges outlined above and the associated time constraints, a single fabric skeleton containing the board, pumps, airbags and sensors was created and snaps were added to the inside of both prototypes. This allowed the skeleton to be moved

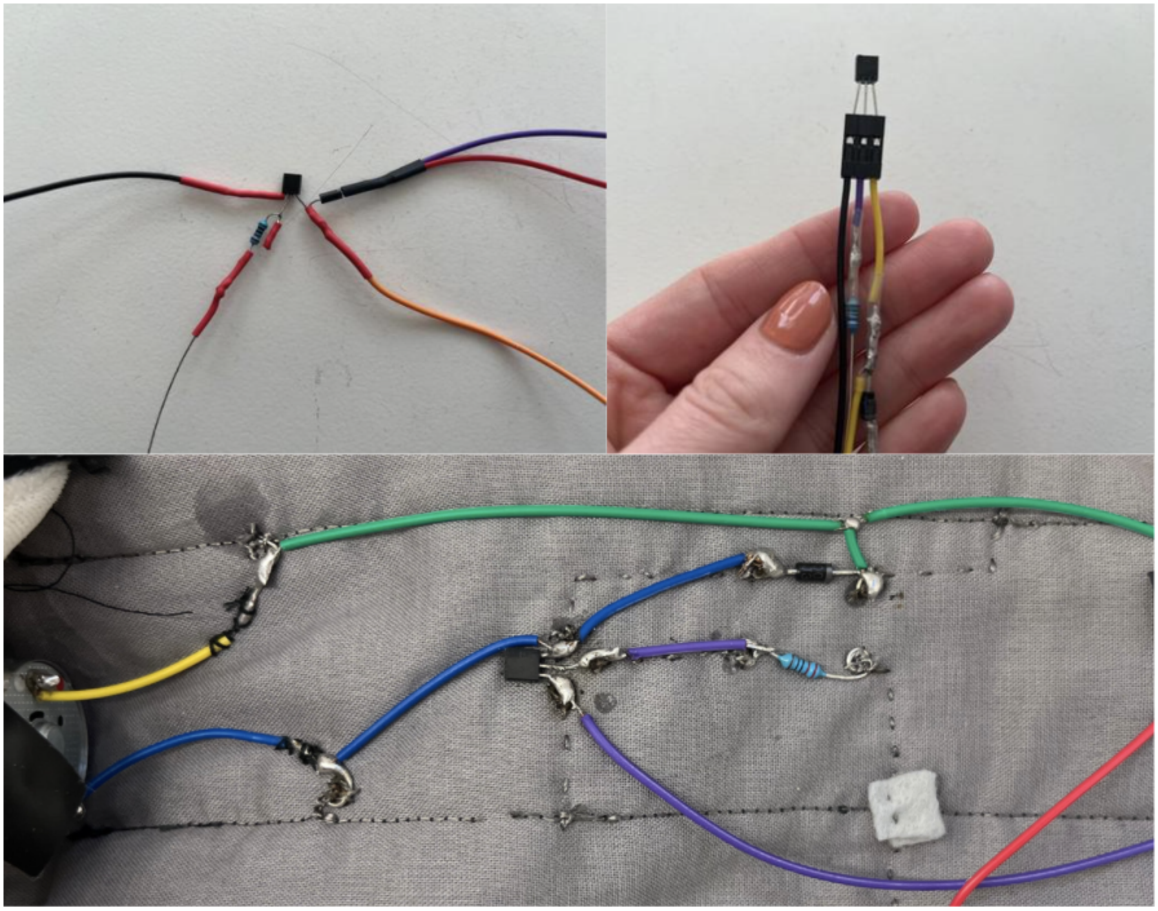

Figure 3.5: Initial configurations that led to metal fatigue (top left and right) and final configuration (bottom).

seamlessly between the two jackets and has the added benefit that other garments could be easily outfitted to house the actuators and sensors by the simple addition of snaps. Additionally, this allowed participants to pick the size and fit they liked best. Before the final construction of the prototypes, two drawings of the physical appearance of the jackets were created. These are shown in Figure 3.6.

Figure 3.6: Sketches of Jacket A (left) and Jacket B (right) before final construction.

3.1.3 Final Prototype

While the physical appearance of the garment has changed significantly through the iterative process, the main theme remains the same. Two jackets were produced that provide compression on the torso whenever the user desires. This pressure is meant to bring awareness to and potentially relieve negative emotions experienced by the wearer. Given challenges with metal fatigue in components and time constraints, the two jackets created are outer shells without any electronic components. Therefore, the jackets themselves do not hold any sensing or actuating capabilities. Instead, standard, metal sewing snaps are affixed inside both jackets. A skeleton containing all electronic components was constructed and matching snaps were sewn on. This allowed the electronic components to be seamlessly passed between the jacket shells. Wearers can then choose the shell that they prefer and fits their body best. Additionally, if the need ever arose for additional jackets, standard clothing could be easily outfitted with the addition of more snaps within a couple of hours.

Figure 3.7: Jacket A (left) and Jacket B (right) displayed on a mannequin and individual respectively.

The final prototypes for Jacket A and Jacket B are shown in Figure 3.7, with the internal structure shown in Figure 3.8. Figure 3.9 shows the inner skeleton snapped into Jacket B. Airbags are visible at the sides, and the pumps and should pads are hidden under the jacket fabric. The left wrist contains a heart rate sensor. This sensor is attached to fabric that runs the length of the left sleeve to meet the microcontroller on the chest. Conductive thread was used to connect the sensor to the microcontroller, and a Velcro strap allows the heart rate monitor to be placed flush with the wrist, while accommodating different wearer wrist sizes. A vacuum and a pump are located on the left and right shoulder respectively. The pump and vacuum are connected to two airbags located on either side of the torso. The pump is primarily used and the addition of the vacuum allows for quicker deflation if needed by the wearer. A microcontroller is placed on the left chest with buttons facing outwards for the wearer to turn on and off the pump and vacuum.

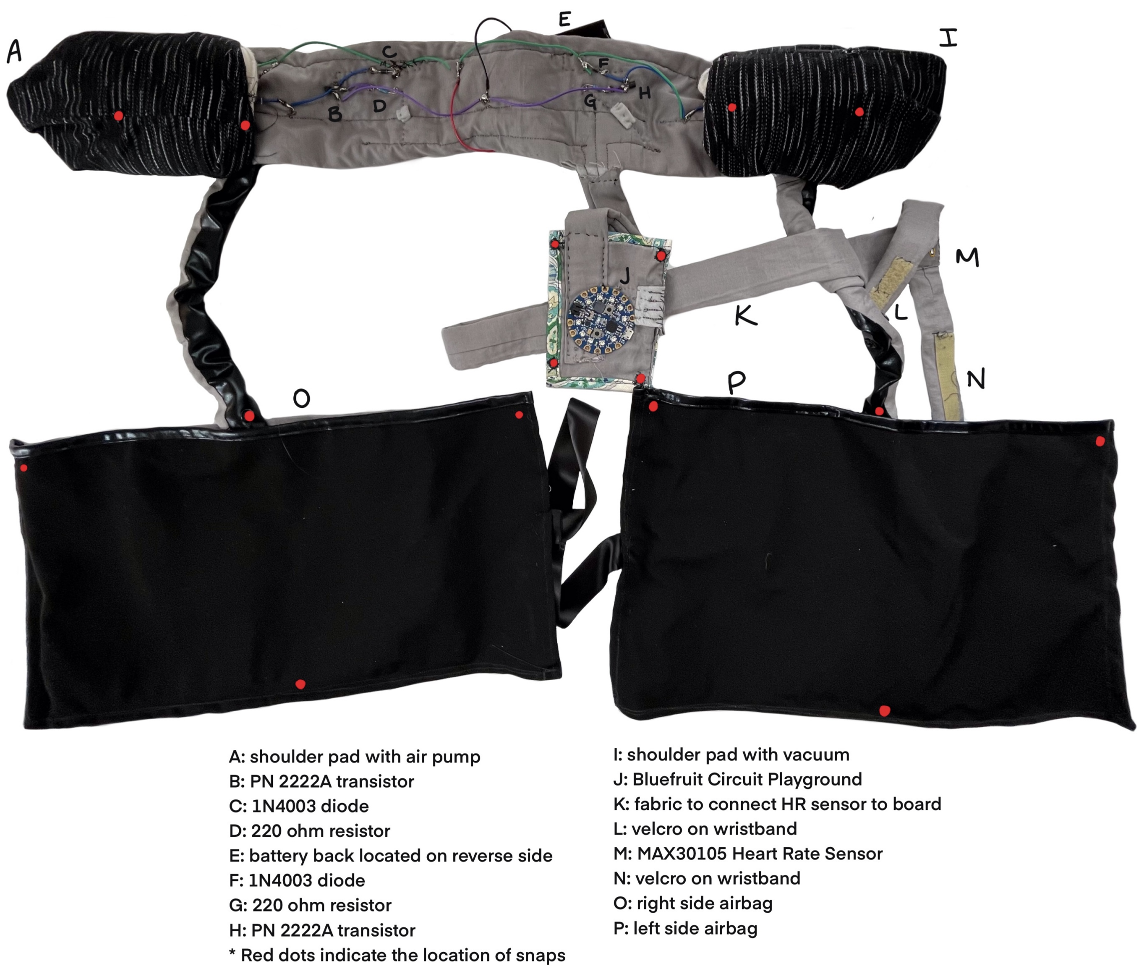

Figure 3.8: The inner skeleton that attaches to either jacket with the use of snaps.

The microcontroller used to control the functionality of the interactive garment was the Bluefruit Circuit Playground by Adafruit. Two identical 4.5V Air Pump and Vacuum DC Motors were used. The pumps can be used interchangeably as a vacuum or pump by changing which post the silicone tubing is attached to, as one

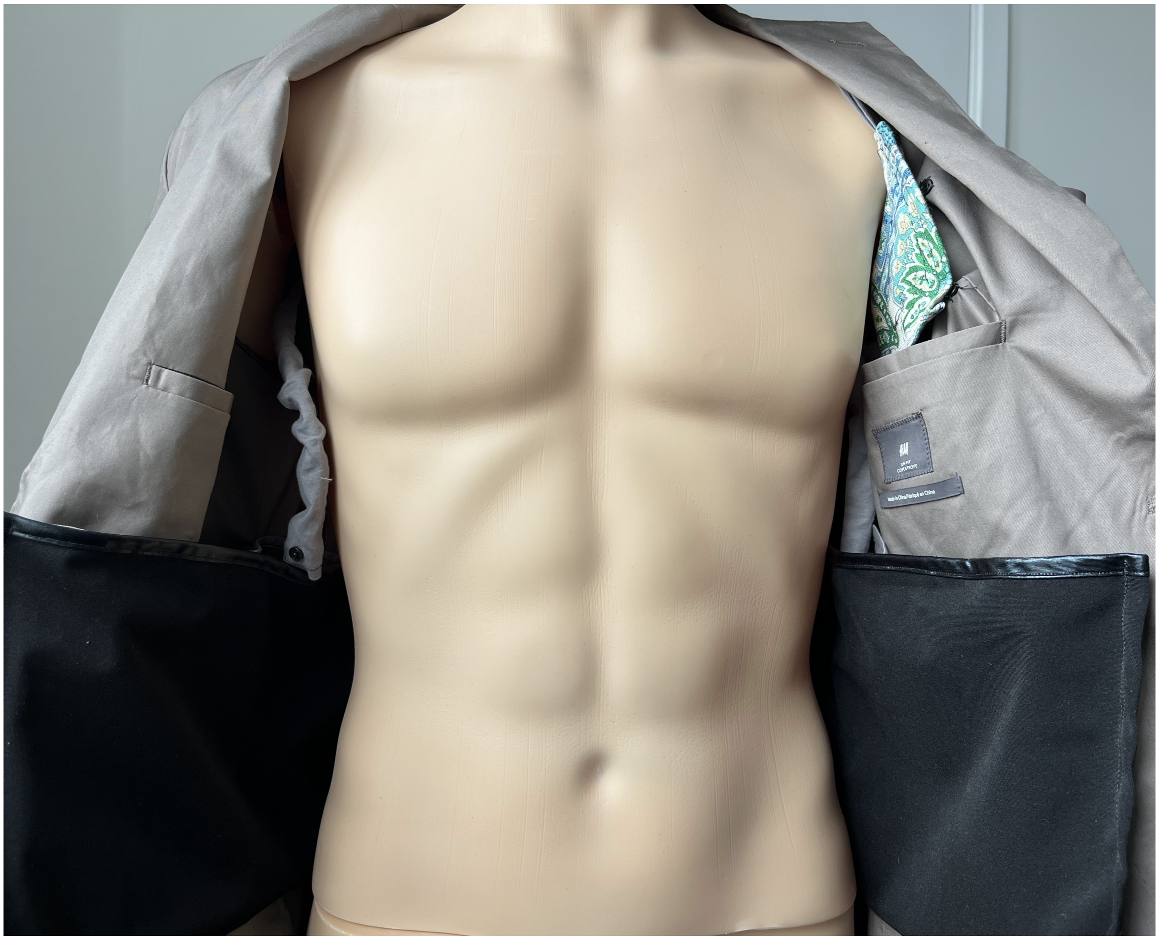

Figure 3.9: The inner skeleton has been snapped into Jacket B.

sucks in air and the other blows out air. A 6V battery pack containing 4 double-A batteries was used to supplement power to the pumps. Although rated for 4.5V, motor specifications indicate that higher voltage can be sustained by the pumps. Batteries lasted for approximately half of the participants. After the first instance of the batteries dying, batteries were changed between every other participant to ensure no data was lost due to power failure. The heart rate monitor used is the MAX30105 Particle Sensor by SparkFun. Additionally, PN 2222A transistors, 220 Ohm resistors and 1N4003 diodes were used to allow the pvm pins on the Adafruit board to turn on and off the pumps using the built-in buttons. Figure 3.10 shows the technical specifications for wiring the circuit.

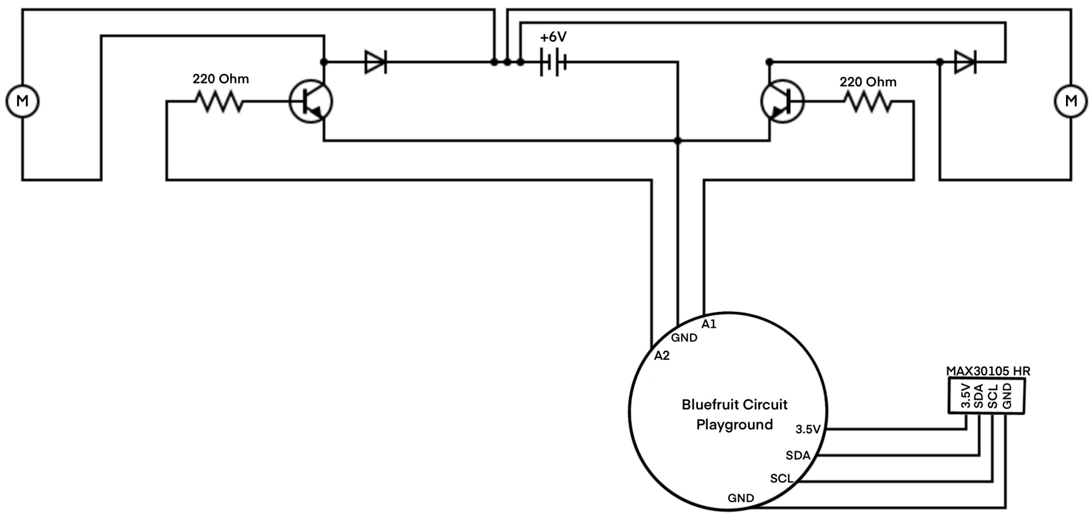

Figure 3.10: The technical specifications of the circuit wiring.

In the case of both jackets, once the wearer dons the garment, they interact with it by pressing fabric buttons sewn on the left-hand side of the chest that correspond to built-in buttons on the board. The green button corresponds to the air pump. By pressing the pump, the air is gradually pumped into the bags at the side over the course of a few minutes. Leaving the air pumps on for extended periods of time does not pose a risk to popping the bags. Wearers can turn off the pump by pressing the green button a second time. There is no set time the pump turns off, and wearers must always choose to turn it off, i.e. it will not stop automatically. The red button turns on the vacuum, which sucks air from the airbags to deflate them. While air is naturally lost from the bags when the air pump is off, the vacuum allows for more rapid deflation of the bags if the wearer is uncomfortable. Pressing the red button a second time turns off the vacuum. Similar to the pump, there is not an automatic stop after the vacuum has been on for a period of time. On the wrist using an adjustable Velcro strap, the heart rate sensor is affixed close to the skin for best results. The built-in sensors in the board collect data on movement and temperature while the wearer interacts with the prototype.

3.2 User Study

In order to gain insights into users' experiences of the jackets, we performed a user study. In this user study, participants were asked to wear either of the jackets for a period of time while they were doing school-related tasks. They decided when to apply and remove compression from their sides and then reported the impact on their emotions while doing so. Additionally, data was collected during the interaction that is used for post-hoc classification to determine when the wearer decided to apply compression.

Participants were recruited from the Queen's University School of Computing, in both the undergraduate and graduate programs. Ethical approval was received from the ethics board for Queen's University. Participants were asked to come in-person to a lab space in Goodwin Hall on Queen's Campus. Figure 3.11 illustrates the lab and setup where participants completed the study. Participants completed work at the lab meeting table. The lab was primarily unoccupied, with 2-3 other students completing work silently at their desks beyond the dividers (K, L). Occasionally others entered through the main door (A) to use the coffee setup (B). The study lasted for a two-hour period and was broken into two blocks.

In the first block, they were asked to wear pick and wear either Jacket A or B for

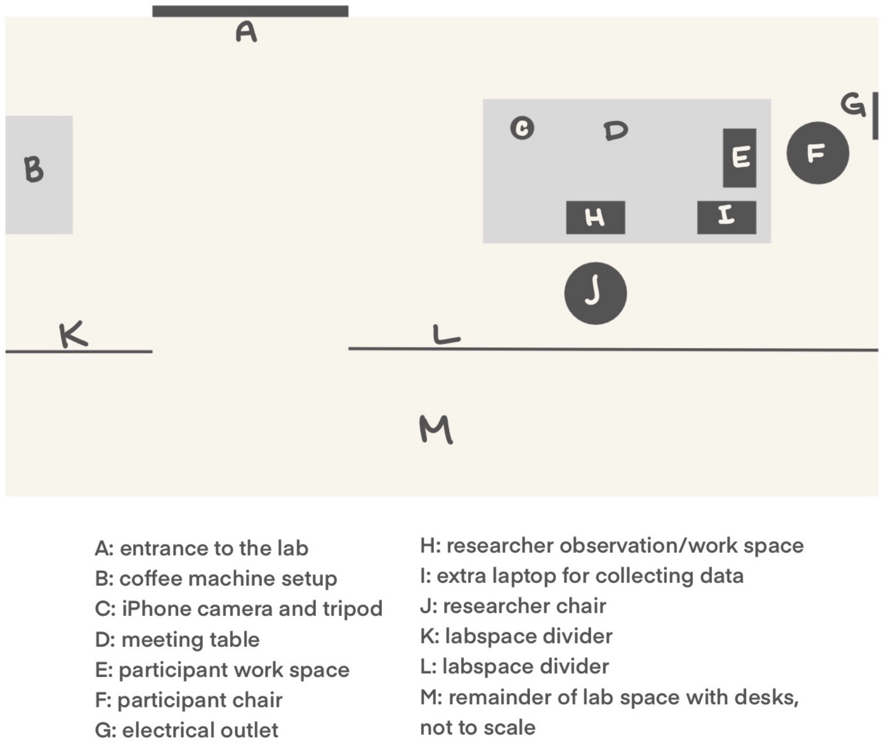

Figure 3.11: A floor plan of where the experiment was conducted (not to scale).

a period of 50 minutes while performing school-related tasks. As the students were all recruited from the School of Computing, and the study was conducted during the Winter 2022 term, all students had work they could complete. Schoolwork was defined as assignments, readings, or lecture review. Participants were asked to bring all materials they needed to complete work, but a table, chair, and outlet would be available to them. Participants were video and audio recorded on an iPhone 13 mini while completing their work to facilitate analysis after testing.

Participants were shown the jackets and given a demonstration of how the inflation works, as well as the buttons to control the pump and vacuum. Next, participants donned each jacket without the components snapped in to determine the best fitting one. The component's skeleton was affixed, and the textile was plugged into a laptop

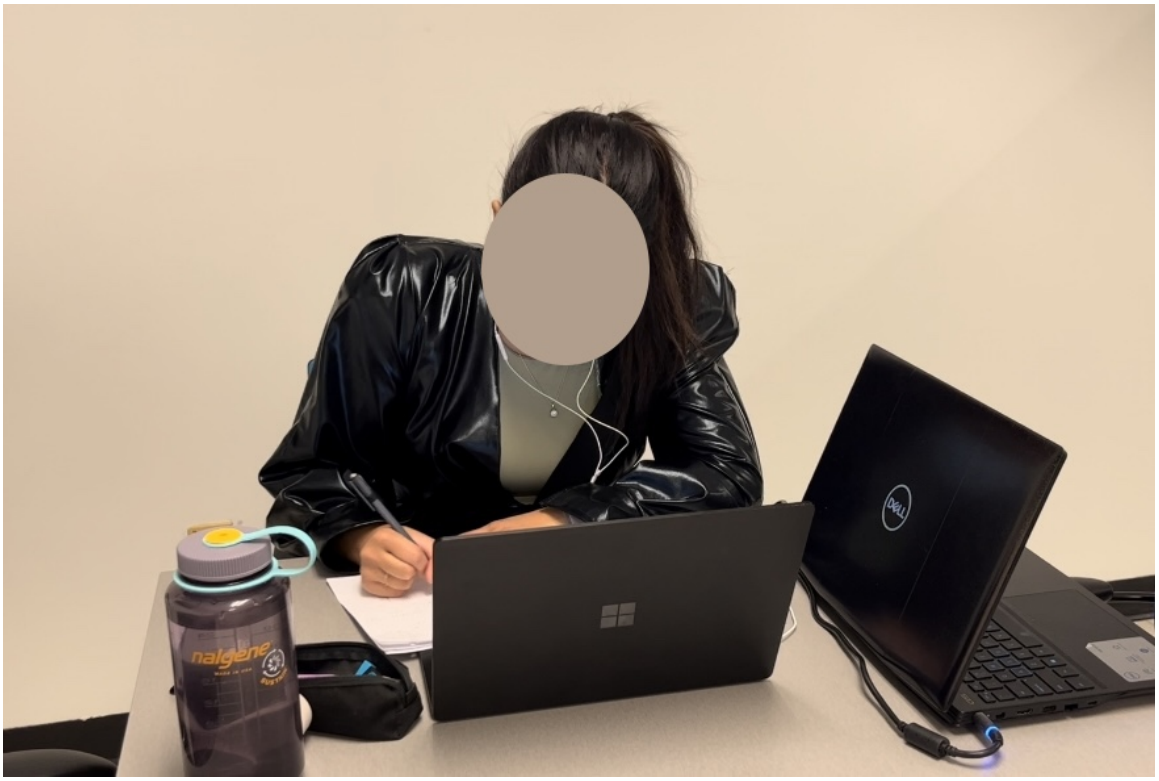

Figure 3.12: Image showing a participant completing school work while wearing Jacket A.

to collect the physiological data. Although they were required to have the jacket's closed to ensure they felt compression on their sides, participants tied (Jacket A) or buttoned (Jacket B) it to their level of comfort. Participants worked on laptops or paper at a desk. They were instructed to experiment with compression when they were feeling emotions, such as stress, anxiety, or distraction by pressing the wearable button to trigger the actuation. They were not given any specific number of interactions to complete, as there was a desire for these interactions to happen naturally to mimic everyday interactions as closely as possible.

In the second block, participants completed a series of semi-structured interview questions that were audio-recorded. These interview questions, in Appendix A, asked about both experiences with negative emotions in general life, and their experiences with the interactive jacket.

Following the study, interviews were transcribed using Microsoft Word's automatic transcription tool. Participant interviews were given a randomized ID in the range [1,12]. Following the method outlined for thematic analysis [7] in Section 2.1.2, the interviews were manually coded following a latent approach. These codes were then extracted and categorized into themes, and are presented in Chapter 5.

3.3 Predictive Modelling

After collecting data from the participants as they completed the study, predictive models were applied to classify when the users had the airbags inflated. In this section, details on the dataset created, the testing conditions, and models are given.

3.3.1 Data Description

Data was collected from the jackets using the Bluefruit Circuit Playground and the Max30105 Heart Rate monitor. Using these sensors, temperature, x-axis acceleration, y-axis acceleration, z-axis acceleration, infrared radiation, beats per minute, and pump-on were all collected. Infrared radiation is the raw reflected light signal collected by the heart rate sensor, as it uses PPG. More details on PPG can be found in 2. While heart rate alone has been shown to be indicative of stress [10], a larger number of features were extracted to get a well-rounded overview of emotions and how movement patterns may influence or indicate emotions like distraction. Table 3.3 summarizes the data collected. Figure 3.13 shows the directions of acceleration that are measured from the board. The pump-on label was collected by lifting the variable that contains whether or not the on-board button has been pressed directly from the Bluefruit Circuit Playground. A sampling rate of 100Hz was used. Data was read from the serial port connected to the board using Processing. Every 500 samples the data would automatically save to a comma-separated value file to minimize data loss in case of technological failure. After 24,000 samples were taken, approximately 50 minutes, data collection ceased.

Data	Type	Range
blood volume pressure (bvp)	float	(163796.0, 262729.0)
beats per minutes (bpm)	float	(0.0, 48.0)
temperature (temp)	float	(23.0, 30.0)
x-acceleration	float	(-14.0, 2.0)
y-accelerations	float	(-13.0, -4.0)
z-acceleration	float	(-17.0, 11.0)
pump on	int	0, 1

Table 3.1: Data collected from the interactive garment used as features to train the models.

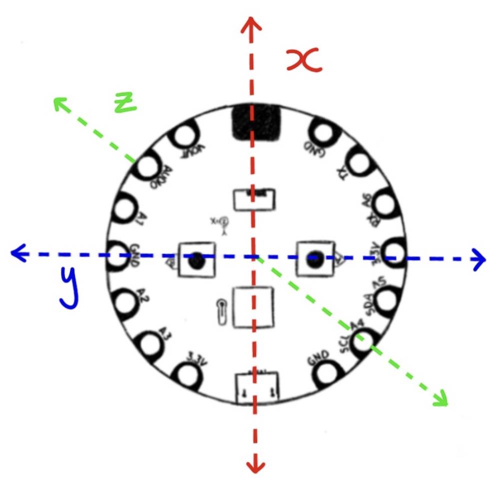

Figure 3.13: The x-, y-, and z-axis that acceleration is measured on on the Bluefruit Circuit Playground.

3.3.2 Preprocessing and Statistics

For each participant, the smaller files containing 500 samples were concatenated into a single Pandas DataFrame containing all the data for the respective participant [43]. The data was then indexed and sorted by time. A 75Hz high pass filter was then applied to the IR signal to convert it to BVP. Due to some fluctuation in button pressing on the board, the pump on the column was smoothed by passing a moving average window, specifically mode. Lastly, the first five percent and last five percent of the data was dropped as outliers. Ultimately, the DataFrames for each participant contained six features that were used to train predictive models in a binary classification task.

Following prepossessing, a report was generated for each participant to extract statistical measures. For each continuous variable, the mean, median, mode, minimum, maximum, 25-quantile, 50-quantile, 75-quantile, kurtosis, and skew were generated. For the categorical column, the count, frequency, and unique occurrences were generated. In addition, methods to count the number of pump interactions and the total amount of time the pump was on were created. Visualizations of the data were also generated and visually scanned to ensure the data behaved as expected. Hexplots plots were created to visualize relationships between pairs of continuous features. To visualize the relationship between continuous features and categorical features, bar graphs were used.

3.3.3 Experiments

Three different models were created to perform a binary classification: Logistic Regression, SVM, and XGBoost classifier. Deep learning models weren't explored as we had a limited amount of data and did not want to overfit. These models were implemented using the industry standard for parameters. These models were trained independently to perform two types of classification tasks. The first classification task is predicting whether the pump should be on or off at a given time step. The second classification task is predicting whether the pump is on or off for the next time step given 45 previous time steps, which corresponds to approximately 5 seconds of data. For the second classification type, a 2.5-second sliding window was applied to the data to create 5-second windows to pass to the classifier.

Traditional		**Backward Increasing**	
Train	**Test**	**Train**	**Test**
10%	90%	10%	20%
20%	80%	20%	20%
30%	70%	30%	20%
40%	60%	40%	20%
50%	50%	50%	20%
60%	40%	60%	20%
70%	30%	70%	20%
80%	20%	80%	20%
90%	10%		

Table 3.2: All training and testing splits used to train three different models.

Two different methods of splitting the data into training and testing sets were explored. Within these methods, a number of different train-test splits were applied. This allowed us to explore how different amounts of pretraining can influence model performance. The first train-test split kept data sorted temporally and started with a testing size of 10% and 90% training. The testing size increased by 10% each time until it reached a maximum size of 90% testing with 10% training. The second train-test split keeps the data sorted temporally and fixes the last 20% of the data

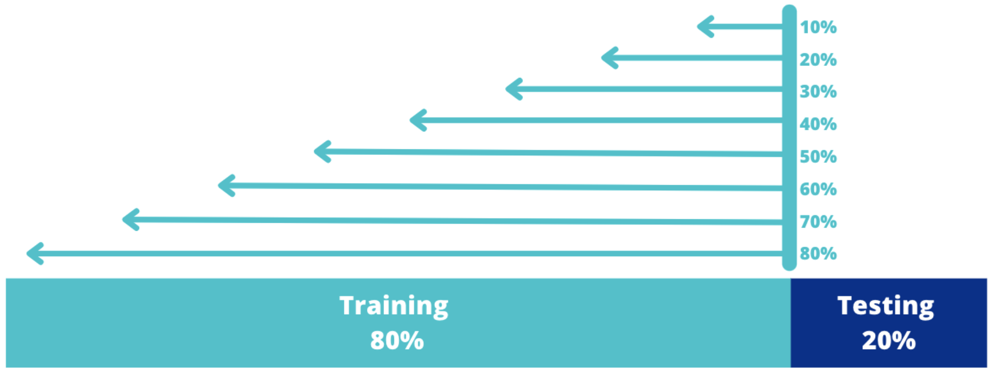

Figure 3.14: Splitting the training and testing sets by moving backwards in time from a fixed testing set.

as testing in all cases. This is done so we can explore the affects of offering different training amounts on the fixed testing set. The training set starts as 10% of the data immediately before the testing set. The training data increases in 10% increments moving backwards in time from the start of the testing set, stopping when the training set reaches 80%. We refer to this as the backward increases and Figure 3.14 illustrates the concept. Furthermore, Table 3.2 summarizes all training and testing sets. All three models were trained for both classification types, using both train-test splits. Table 3.3 summarizes the different conditions evaluated. For each participant, all tests were run. Data from participants was kept separate, as we hypothesized that there are unique differences between participant biological signals and use cases.

	Current Step	**Next Step**
Traditional	Condition A	Condition B
Backwards Increases	Condition C	Condition D

Table 3.3: A breakdown of the four conditions used in experiments.

After completing these experiments, the top-performing model was selected for each train-test split under each different condition. From here, the most frequently occurring model was reported across participants. Additionally, the top performing model was selected and averaged across participants only if better than baseline accuracy was achieved. This aggregation of data facilitated the analysis of trends reported in the results section.

Chapter 4

Results

The results chapter is split into two parts. The first section, Understanding the User Experience, reports the qualitative data collected from participants during interviews. The second section, Quantitative Results, reports on interaction statistics, as well as results from the application of predictive models on collected data.

4.1 Understanding the User Experience

In this study, 12 participants were asked to interact with one of the two prototypes while doing independent desk work. Afterwards, they were interviewed about their experiences. All participants were anonymized as P1 - P12 and were recruited by email from the Queen's School of Computing. Of the 12 participants in the study, 4 identify as female and 8 identify as male. Participant ages range from 18 years old to 30 years old. Half of the participants are undergraduate students, while the other half of participants are currently enrolled in graduate studies, either at the masters or doctoral level. 6 participants reported their cultural background as Canadian, 2 as Middle Eastern, 2 participants as South East Asian, 1 as Indian, and 1 as Chinese American.

Participants were asked to reflect on their emotions during everyday school life, as well as their experiences interacting with the prototype. Interviews were conducted in English and lasted 20 - 40 minutes depending on the participant. Interviews were automatically transcribed using Microsoft Word's automatic transcription tool and manually checked for transcription errors. Transcripts ranged from 3027 to 5051 words in length.

From these interviews, six themes were created using thematic analysis [7] from a total of 31 distinct codes. The six themes are: Stress and Distraction Shape our Daily Lives, Everyone Copes Differently, Emotion Regulation Requires Human Initiative, Physical Interactions Ground Individuals, Emotional Garments Are Not One Size Fits All, and Automation is Encouraged to a Point. Table 4.1 summarizes the themes and associated codes, and these themes are outlined in further detail below.

4.1.1 Theme 1: Stress and Distraction Shape our Daily Lives

All participants mentioned ways in which stress or distraction shapes their lives, from schoolwork to basic needs like food and sleep. While frequency and intensity varied between participants, commonalities between elements in their lives that were affected were observed.

Frequency Most participants expressed feeling stressed or distracted at school multiple times a week, such as P10, who experiences notable stress *"[on] average, I would say two times a week"*. Many others experience stress more frequently. For example, P3 expressed that they experience stress and distraction, *"every time, [I do schoolwork] like I don't know how people can not get distracted"*. One participant, P2, reflected on stress experienced *"a couple years ago"* when they were still completing

Theme	Code
Theme 1: Stress and Distraction Shape our Daily Lives	frequency tipping point impacts on time unhealthy relationships with food impacts on sleep emotional interactions
Theme 2 Everyone Copes Differently	separation acceptance exercise breathing scheduling
Theme 3: Emotion Regulation Requires Human Initiative	forgetting effort to engage in tasks time take out of the day inconvenient times habits
Theme 4: Physical Interactions Ground Individuals	stimulating the senses screenless interactions familiar patterns grounding personification
Theme 5: Emotional Garments Are Not One Size Fits All	experimental use cases love it or hate it design qualities pressure points potential use cases
Theme 6: Automation is Encouraged to a Point	enthusiastic interest a manual override is a must relieving the mental load negative feedback loop explanations

Table 4.1: Summary of themes and codes created from interview data.

their undergraduate degree.

Tipping point: When stressed, two participants felt that they had levels of stress that boosted their academic performance, but hit a point where stress becomes a hindrance. P5 felt that when they became too stressed, they *"freeze completely"* and can't work. Other times, they *"turn into Autopilot mode in your mind and start* *getting things done"*. P3 thinks that *"a little bit of stress is good, but too much stress, it's just counterproductive"*. P1 sees stress as *"sort of like a necessary evil. I know it has to be there"* but then also finds it impacts their enjoyment of personal time.

Impacts on time: Distraction while completing work caused participants to take longer to complete tasks than they would like, bleeding into their personal time. When asked about how stress and distraction impacted their day, P8 observed that their *"days end up being longer .... [I] could work for like ten hours straight but get nothing done somehow"*. P7 does not enjoy working on tasks in the evening, but *"if I get distracted for an hour, then I lose that hour later 'cause I still have to do it and I always catch up ... it's just during times like I don't want to be working on school work"*. P8 also said *"if you had condensed what little work you did into like an hour, you would have had nine hours for yourself"*. P12 feels that procrastinating tasks when stressed *"makes it so that I have less time to do things that I enjoy and hobbies and stuff"*.

Unhealthy relationships with food: Eating habits are impacted by feelings of stress and distraction. P12 *"didn't eat all [their] meals"* when experiencing school stress. Instead of missing meals, P5 said eating sweets was an emotion management technique when stressed, saying, *"whenever I'm stressed I just go right to the fridge, get something to eat, and then I'm back again"* and when asked if it helped, P5 said *"it just made me gain more weight"*. When asked about stress and distraction impacting their daily lives, P9 said *"sometimes I forget to eat ... I do care a lot about my diet, it's just sometimes I like skip meals and stuff"* and they binge later in the evening to make up for it. When describing a longer lunch with friends in the cafeteria, P1 expressed guilt since they could have been working instead: *"you really* *shouldn't have spent that much time there in the first place"*. In total 4 participants mentioned forgetting to eat meals or having unhealthy meals because they were faster to prepare.

Impacts on sleep: With deadlines approaching, participants expressed that they would regularly sacrifice sleep in order to complete work. Coupled with tasks taking longer than they should because of stress and distractions, this becomes a difficult predicament. P12 makes a plan for each day and tries *"to complete it before I go to sleep, but sometimes that means that I literally won't go to sleep until I complete it, so it might be like 2 in the morning by the time I get around to getting ready for bed"*. Not only is it a late night, but this takes away from feelings of productivity that P12 experiences in the morning. P6 noted that, *"sometimes when you're really stressed about something it's like difficult to sleep"*. When asked if there were certain times that stress or distractions felt more intense, a majority of participants said in the evenings before bed.

Emotional interactions: Participants were asked to discuss negative emotions, such as stress, anxiety or distraction. Participants discussed these feelings independently from each other as well as their interactions with each other. Stress and distraction can play off of each other, such as P6 feeling *"like the more work I have to do, like the more you want to procrastinate all of it"*. P5 mentions other factors that compound on schoolwork stress: *"there is also like some personal stresses that come along the way"*. P1 mentioned a similar sentiment *"I get too distracted and then it goes over the line, and then you get into this stress and anxiety kind of"*. When stressed, P12 skips meals, but also notes that missing meals *"makes me a lot more moody ... it feels like everything is more elevated ... if I'm stressed then I'll feel* *even more stressed"* creating a vicious cycle.

4.1.2 Theme 2: Everyone Copes Differently

Throughout these interviews, each participant presented their own routine for dealing with the negative emotions that stemmed from school life. While unique to each individual, commonalities exist between the activities. Furthermore, a number of techniques were not given when directly asked, but rather emerged through more discussion. It appears that in a number of instances, participants had come up with ways to cope with their emotions without being conscious of it.

Separation: When dealing with stressful events in their lives, some participants turned to separate themselves from the activity as a way to manage their emotions. For example, P2 mentions watching *"Youtube, or playing videos games, or do something else or hangout with my friends, and trying to avoid it"* when dealing with difficult coding problems. P11 enjoys *"talking to people"* during breaks in their work. P9 enjoys giving in to some of their distractions and playing video games with classmates: *"that's my little thing where it's like a social thing where I also get to play a game I like"*.

Acceptance: For many participants, they have accepted feeling stressed as part of their daily life, with a number of participants echoing the sentiment that work must get done no matter their emotions. P5 finds they hit a point close to a deadline and *"somehow turn into Autopilot mode in your mind and start getting things done"*. P10 feels that, especially for graduate students *"stress is something part of the life"* and you must accept this. Additionally, P4 sometimes finds that just tackling stressful tasks *"kind of gets rid of"* the stress.

Exercise: Physical activity helps some participants manage their emotions. P10 has started *"running early morning...it's really helped me"*. P4 find that *"going on walks is good, or skating at [the park] where I live"*. P9 has made changes to their life to deal with stress and emotions, out of the changes *"going to the gym is one of them. That's a big de-stressor for me"*. When asked about movement times as a way to manage emotions, participants largely found that they either (a) came at inconvenient times, which will be discussed in Theme 3, or (b) were already getting enough movement in their day. P9 doesn't see the use in devices that prompt you to move because *"I get my steps in and I'm walking to campus. I go to the gym. I assume that's enough activity"*. Similarly, P3 and P2 find that small tasks that regularly occur throughout the day, like bathroom breaks and getting coffee, have them up and moving with the same frequency of a movement reminder.

Breathing: Some had limited experience working with controlling their breath as a way to manage their stress, typically only trying it during workshops or exercises classes, such as yoga. P6 uses an app that guides meditation and includes controlled breathing exercises. P9 uses breathing techniques learned in Taekwondo classes as a child when they feel stress. P12 has a breathing routine that they have used since high school for test-taking. They *"pick a point like somewhere in front of me, and I would just like breath in for a few counts ... just until my lungs are filled and then I breath out again in a slow and controlled manner"*. They attribute this technique as the reason for an increase in their course marks from the high 80s to mid-90s.

Scheduling: A few participants found that having a plan or schedule helps to lessen feelings of stress. When P2 is feeling stressed with school work, they list the reasons why they are feeling stressed then breaks down solutions to these stressors and begins to address them one by one. P12, P7 and P5 all find the use of daily schedules to be helpful. P12 enjoys the use of their planner because it makes them *"feel like there's nothing that can really creep up on me since like I see it all coming in advance"*. P4 uses an online calendar and likes the colouring coding because it *"makes fun stuff differentiated from school stuff"*. P4 also writes down tasks in their phone's reminders and has them displayed on their phone's lock screen.

4.1.3 Theme 3: Emotion Regulation Requires Human Initiative

When discussing emotion regulation techniques, participants noted that there is active decision-making involved. Whether it be personal strategies, guided meditation or movement reminders, a person must choose to do the activity.

Forgetting: Participants mentioned that they would simply forget to engage in an activity, despite finding it helpful. When discussing emotion management, P4 said they have a stress ball they bounce around to fidget with while working, however *"sometimes I forget and then just don't remember it till later"*. P7 remembers downloading self-care apps and forgetting to use them *"I'll just use it for like 2 days and never open it again"*. Despite finding them useful, when asked about using controlled breathing to help with stress, P4 says they *"just forget about it as an option a lot"*. P5 found that notifications, like movement or taking a breath, became easy to ignore because *"the novelty wore off"* and they would forget to engage in these activities.

Effort to engage in tasks: While phones offer a sense of convenience, participants must choose to engage in digital emotion regulation techniques. P5 said *"for a phone or an app or whatever, I have to do some interaction, do some work, for that thing* *to start acting the way or do the functionality I want"*. P12 highlights that you need to actively chose to help yourself: *"you have to do something to help yourself like all [notifications] can do is just remind you or be a tool, but you have to be like oh I'm going to I'm going to do this thing that my laptop's reminding me to do, but you could just choose to not do it"*. P7 feels that there is some effort exerted when choosing to open an app, saying *"you actually have to check the phone ... so if I don't feel motivated to like open an app, then I' never going to use that app"*. P11 discusses the steps taken to open an app *"if I have to use a phone I have to like open the phone, unlock the phone, remove the [face] mask [for face recognition]"* which leaves the participant not wanting to use a phone regularly. P6 felt there was a drop in effectiveness in break/breath/movement reminders if you don't engage in the tasks: *"I feel like those notifications, they don't really work unless you actually do it every time"*.

Time taken out of the day: Participants had positive feelings towards emotion regulation but didn't feel they could take the time out of their day to do them. P1 stopped practicing meditation because *"I feel like I'm taking a chunk of time out of my day, which is kind of weird 'cause I feel like I should take chunk out of my day to do like schoolwork, so I should really split equally, but I just don't do that"*. P2 stopped meditating on the subway before work because the internship was over and they no longer had an hour-long block of time that they could dedicate to meditation.

Mental load: At times, remembering to engage in these tasks can become mentally tiresome. P12 says they *"have to think about like opening the app and then I have to think about using it"*. P12 discussed how choice is involved in self-care reminders: *"you have to do something to help yourself, like all it can do is just remind you or be a* *tool, but you have to be like, 'oh, I'm going to do this thing that my laptop's reminding me to do..., but you could just like choose to not do it"*. This suggests that there is a level of choice to use the tools and strategies put before you, and this choice is a decision one must make. When asked if they would like to make emotion regulation techniques a habit, P3 responded: *"oh yeah, yeah, because doing [it] consciously is exhausting"* indicating that there is some mental exertion required when choosing to do these activities.

Inconvenient times: Notifications from apps can come at times that aren't beneficial for users, despite being geared towards self-care. P6, when discussing movement reminders, said *"I feel like sometimes when you're like in the middle of it like you want to get up every hour like something you're doing like I don't know, I feel like I would just lose it or like you know, forget about it"*. P1 doesn't like reminders because *"I'd find if I were to get up then I would probably just be more distracted really"*. P5 experimented with the Pomodoro Technique, where you work for 25 minutes and then take a 5-minute break. They didn't find it effective because *"sometimes I feel like I want to work for an hour sometimes I feel like I want to work for five minutes, so just sticking to the 25-minute rule is kind of annoying and it doesn't work for me"*. P12 spoke about getting movement reminders in class, saying *"if I am in class then I obviously won't"* listen to the notification.

Habits: A number of participants expressed the desire to engage in emotion regulation techniques as a habit. When discussing the use of the garment for emotion regulation, P7 felt that they would just be making one small change to their daily habits as opposed to using an app: *"with this [garment] it's like you sit at your desk, you put it on, you make it a part of your routine and it's always going to be there* *and you can't just like not do it".* P12 was able to help manage stress by making controlled breathing a part of their routine saying they were *"programming myself to have this response in a stressful situation".* P9 also found they were able to make a habit of reducing time on social media apps by using a reminder to get off the app: *"I did find that after starting to do it, even when I dismiss [the reminder to get off the app], I still would spend less time than when I just didn't have the feature altogether".*

4.1.4 Theme 4: Physical Interactions Ground Individuals

For many participants, the exploration of the prototype became a personal journey, stimulating different senses, evoking memories, and providing comfort.

Stimulating the senses: Compression, vibration and sound provided participants with a sensory experience while interacting with the prototype. P8 liked vibration and sound as a distraction notification over laptop notifications because *"if you're doing computer work, you're already stimulating your visual sense ... so if like your reminder is also through that sense and you're already consuming so much information through the same sense ... it's ineffective".* P6 thought the stimulation of their senses through sound and vibration was distracting in a positive way: *"if you're caught up in an emotion and then you have pressure applied or the vibration and sound, I feel like it kind of distracted you from that".* P11 enjoyed the gradual reduction of pressure when the pumps were turned off more than the application of pressure, stating that *"it felt that things were like being lighter or something ... something heavy is getting down, melting off".*

Screenless Interaction: Participants generally enjoyed the lower fidelity nature of the jacket, enjoying it's physical presence and the fact that it didn't have as many distractors as a screen-based device. P9 uses their phone a lot, *"I will go on Instagram or Tik Tok, or any sort of dopamine applied thing"*. Similarly, if in front of a screen, P8 finds themselves procrastinating *"I'll check email or something like that, especially if something is loading then I'm like oh I can pop open Reddit and scroll down, it's like an hour later..."*. P7 gets distracted by their phone so has to put it on the other side of the room, but they *"don't like putting stuff in like do not disturb mode or silent mode 'cause what if I miss something that I need to look at"*. When comparing the prototype to a phone screen, P3 noted that *"I like tactile, I don't like screens everywhere. I think if you compare to the phone, I'd rather have the fabric itself"*, although P3 does note the convenience of Apple Watches for payment and not needing to use Face Recognition. P4 differentiates the prototype and screens, saying the prototype *"is on you all the time and like you can sort of leave and forget about it I guess. I feel like I can't really forget about like a laptop in front of you in the same way like I could forget what t-shirt I threw on this morning"*. P12 describes the jacket as *"a physical thing that helps you physiologically"*. P1 felt that they could ignore a screen, whereas with something physically on the body *"it takes a lot more actual effort for me to take off the whole [garment] e- and ... I'm not going to take off the whole thing just to ignore like that notification"*. P11 finds that they don't like using a phone a lot, and would rather use the physical prototype: *"we already use the phone a lot so it gets too tiring sometimes...I'm not a fan of like overuse of phones"*. Instead, P11 would prefer to use buttons on the garment to control actuation, but *"maybe it's like better to not to keep too many buttons"*, lending itself to an idea of a lower-fidelity garment.

Familiar patterns: Garment actuation through vibration is a tactile sensation participants are used to. P7 noted that *"my watch, it vibrates sometimes"* when discussing vibrating notifications and P3 related it to *"those vibrating chairs"*, i.e. massage chairs. P1 felt the vibration *"brought back some memories. I used to have at my house, from there was a neck massage thing"*. Like a number of other participants, this participant likened the haptic sensation from the garment to the relaxing experience of a massage. Interactions with the garment should be intuitive and follow the same movement patterns that participants use to interact with non-electronic textiles. P4 mentioned having the on and off buttons in the pocket rather than on the chest: *"you could maybe stick it in a pocket and then it's a little bit more subtle when you click it"*. They related this to putting their hand in their pocket, which is a subtle action they complete daily. P3 likes the idea of having buttons on the opposite wrist as touching a wristwatch is a common movement they already do daily.

Grounding: P8 said *"the word grounding is really effective"* when discussing how the use of sound made them feel. P8 also related this to Buddhism and sound meditation, reminding oneself that you have a *"corporeal form"*. P12 said that pressure from a hug "makes me feel a lot more grounded" when feeling overwhelmed by emotions. P9 thought the pressure *"felt like a little hug around my ribcage"* and expressed their enjoyment of it, although the hug could have been more intense. P5 and P4 also felt that the garment pulled them in for an embrace.

Personification: In a few instances, participants began attributing human-like qualities to the prototype, viewing it more as a living entity than a material creation. P5 spoke about the garment as if it were offering comfort, *"[the prototype] was like massaging me and it's like relax; here's a massage for you and here's a hug for you"*. P10 spoke of the garment as if it could reach out and touch them with a hand, saying it would be helpful if the garment could "poke me" or *"punch my hand"* when they became distracted. When discussing feelings of distraction, P8 would turn on the pumps and said *"my brain would be like 'you should distract yourself, you should distract yourself, you should distract yourself' and then the little Angel on my shoulder was like 'uh uh uhh'"*. Personification of the garments intensified when the idea of garment automation was introduced. P5 expressed that if the prototype turned on automatically *"maybe it wants to tell me it's okay or something? I don't know, it would have a life of its own"*.

4.1.5 Theme 5: Emotional Garments Are Not One Size Fits All

Participants were asked to pick one of two prototypes, outlined in Methodology. They interacted with this prototype and decided when to turn it on and off. Afterwards, they were asked to report on how they interacted with the garment, how they felt and how it impacted their emotions. Similar to Theme 2, no two participants interacted with the garment in the same way, and likes and dislikes varied greatly.

Experimental use cases: Some participants chose to activate the air pumps that aligned with the initial imagined use case of stress or distraction and discussed these cases during the interview. P5 received a message about a deadline and activated the pump when stressed and P7 decided to activate the pumps when they felt that they had lost focus. P12 chose to inflate the pump when they realized how long a reading was: *"I was like I'll just turn it on now since I'm feeling a little like overwhelmed"*. Use cases also emerged beyond the intended stress or distraction emerged. P9 chose to initiate the actuation when they changed tasks, *"whenever I was done something and starting something new, I wound tend to inflate"*. P11 inflated the pumps because they were curious. P3 initially activated the pumps *"because I wanna play around* *with it"* but over time, began pressing the button to inflate the pump because they *"want to, you know, play around with some stuff, I'll put my hand somewhere and just click the buttons"*. In two instances, participants could not pinpoint explanations for their interactions with the garment. P9 stated they chose to inflate the airbags at one point for *"no real explicable reason"*. P1 observed that *"there was a weird part where I finally got my code to work and I was like 'yes!' and then I immediately went to inflate the pressure and I didn't really know why"* lending a quality of spontaneity to the interactions. P2 was so focused on their work, that they entirely forgot to interact with the garment.

Love it or hate it: Some elements of the prototype created a stark divide amongst participants. For instance, P8 found the noise given off by the motors in the air pump akin to sound meditation: *"you don't try to decipher it or figure out what it is or anything like that ... I could tune into the sound"*. P10 observed *"it was a smooth continuous sound, I like that"*. However, P11 repeatedly noted that the sound was *"irritating"*, and P3 expressed a strong dislike for any sounds produced by the prototype. For some, the colour and style of the structured blazer was pleasing, such as P10: *"I like your style ... I think it is designed very well"*. For others, the blazer was, emphatically, *"not my shade of brown"* (P9).

Design Qualities As an expression of self, the visual appeal of clothes played a role in participants wanting to wear an interactive garment in public. When asked if how they are perceived by others is important when choosing what to wear, P9 says *"I like to look good and I like other people to think I look good"*; a sentiment that was repeated by other participants. P3 doesn't want to wear anything that makes them look *"too obvious"* that they're wearing an interactive garment. However, aesthetics matter less in private, as P12 mentioned that *"in private, I would wear anything"*. Aesthetics must also balance with functionality in daily life, as P7 brings up that *"I don't think I can use a backpack with this at the moment"* given the position of the air pumps on the shoulders.

Pressure points: The sensation of pressure on the body was enjoyed by participants on the torso, but when asked, participants offered a variety of different points on the body that they might also want pressure. Figure 4.1 shows different places where participants indicated they may want to feel pressure. The darker a region is, the more times it was mentioned by participants. Participants also had different experiences with the feeling of pressure. Some felt the pressure on their sides, such as P12 and P7, others did not, such as P8. Some felt it was not enough pressure, like P9 who *"wouldn't mind if it was a bit tighter"*. Some only were aware of the pressure on one side of the body, such as P11 who felt it *"only there on one side"*. P11 also enjoyed the feeling of the release of pressure by deflating the airbags, rather than the application of pressure to the sides. P5 discussed a similar idea saying *"I felt like [the pressure] was lingering on for too long, so I just wanted to deflate this so the next time I feel the pressure again so that I'm not use to it"*.

Potential use cases: When asked how they could see integrating this into their everyday school lives, participants came up with creative use cases, some closely relating to emotions and others reinventing the garment's use in innovative ways. P9 envisioned using the device if it was late and they needed to stay awake, imagining *"if I'm doing something a 1 a.m. I might, just tuck it on and see if that helps me focus"*. P9 also thought colour changing notifications for stress could be fun envisioning *"an all-black [out]fit and then my arm goes red randomly ... I think*

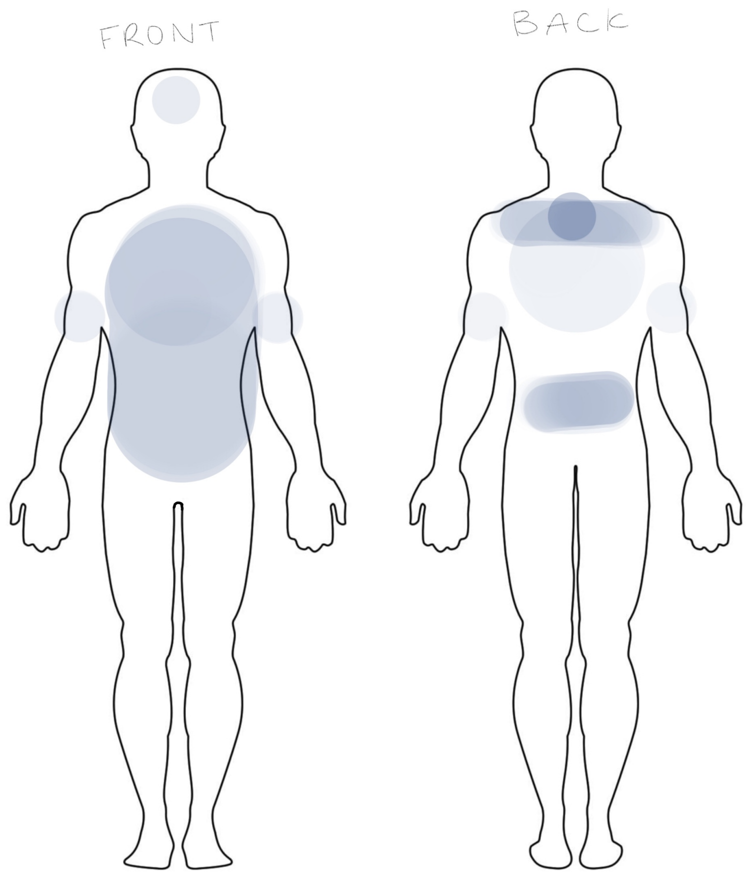

Figure 4.1: Image showing locations where participants may want pressure on the body, as seen on the front and back of a mannequin. Darker regions indicate a higher frequency of interest in that area.

it could be a fun conversation starter". P5 and P11 both brought up wearing the prototype while watching movies to make the experience more *"immersive"* and could act as a *"secondary, complementary stress buster"* while doing an already relaxing activity.

4.1.6 Theme 6: Automation is Encouraged to a Point

Participants were asked how they would feel about automating the garment to inflate and deflate on its own to get a sense of how individuals might respond to a machine learning algorithm determining their emotions. It should be taken into consideration when interpreting these participant answers, that all students are studying computer science, some even majoring in Artificial Intelligence, which may impact their sentiments towards technology as a whole.

Enthusiastic Interest: In general, participants responded positively to the idea of automatic inflation. When asked, P9 was *"fully for automation where possible"* and P11 exclaimed *"that would be awesome!"*. P1 thought that the prototype *"would probably know better when I want to inflate or deflate than I would"*. P3 was interested in the idea of automation as well, saying *"I don't want control over it too much, I just want it to work"*.

A manual override is a must: While there was a general consensus that automation was beneficial, participants thought a manual override was necessary. P6 expressed, *"there's always like a factor of like 'oh ... if I'm not controlling it, then it's kind of scary 'cause like what if it doesn't stop or something, right?'"*. P5 said that there's *"always some comfort to the idea that you have something to stop it"*, referring to a manual stop button. When thinking about the dangers of automation in interactive garments, P4 didn't feel concerned because less could go wrong than other forms of automation: *"it's just clothes, and I feel like it's not as big a deal [as] if we're talking like automated cars"*.

Relieving the mental load: As expressed discussed above, participants found it required significant mental energy to engage in some emotion regulation techniques and some felt that automation would help address this. P12's description summarizes this point, and the sentiment was shared by other participants. They state, *"when I'm stressed I don't realize just how stressed I am . . . and if this garment could control itself . . . I won't have to think about, 'oh, the pressure is a little too much'. It's just like another thing to think about, and if it's something that can be controlled without me having to do anything, then I think that it would be really useful"*. When asked why automating the garment to inflate and deflate automatically would be beneficial, P11 said that they *"don't have to worry about doing it again and again"* referring to applying and removing pressure, which they enjoyed. This indicates that there is a mental toll associated with deciding to press the button that automation would remove.

Negative feedback loop: While pro automation, P8 made the observation that a predictive algorithm could lead to some form of psychological conditioning. P8 thought *"it could be helpful for some people, but other people then might start to associate the inflating with the negative emotion, and then you end up in like this perpetuating cycle"*.

Explanations: Participants expressed that they would likely want to know the cause of the machine learning algorithm's predictions. P6 imagined interacting with it and trying to find a reason why: *"if I was working and it was like automated and then it turned on all of a sudden, I think I'd be like 'oh, am I feeling you know like anxious or something?'"*. P7, who was cautious with automation, would rather do *"stuff manually . . . if something has been changed without me knowing, I'll be like, 'wait, what happened?'"*. P5 thought if it started actuating automatically, they might *"start pondering on like what did I do or what did I feel so that it did that?"*.

Summary of Themes

We have presented six themes extracted from interviews of 12 participants. These six themes are: Stress and Distraction Shape our Daily Lives, Everyone Copes Differently, Emotion Regulation Requires Human Initiative, Physical Interactions Ground Individuals, Emotional Garments Are Not One Size Fits All, and Automation is Encouraged to a Point. These themes provide insight into how users interact with wearables for emotion regulation and validate results found by other researchers in previous works. Further discussion on these themes, and their relationship to the quantitative results discussed below, will occur in Section 5.

4.2 Quantitative Results

Simple and low resource models proved effective in predicting when users had the pumps, and therefore compression, activated under the two train-test split techniques and two binary classification problems. Table 3.3 in Chapter 3 summarizes the different test conditions. The following sections outline the results for each of the 4 conditions. Full code and results are available in a GitHub Repository [1]. To serve as a baseline, we compare accuracy against single class prediction, assuming that the pump is always predicted to be off (the majority case).

Of the 12 participants whose data was collected, 8 participants were used to train the models. Two participants were excluded because of sensor faults, one participant was excluded because they did not understand the activation method, and one participant was excluded because they forgot to activate the pump during the session. The exclusion of these participants does not affect training as all evaluations are done

[1] https://github.com/victoriaaarmstrong/msc_thesis

on a per subject basis, rather than across the entire cohort of participants.

Table 4.2 shows the aggregate results of training different models under the 4 different conditions where the best model was selected from each train-test set. In the case where models are tied, both are listed. For this reason, some reported percentages below may sum to more than 100%, as each instance is counted. The variance in prediction accuracy was also calculated under each condition for all train-test splits and reported below.

Splits	**Condition A**	**Condition B**	**Splits**	**Condition C**	**Condition D**
90-10	XGB/SVM	SVM	**80-20**	Log-Reg/XGB	Log-Reg
80-20	Log-Reg/XGB	Log-Reg	**70-20**	Log-Reg/XGB	Log-Reg/XGB/SVM
70-30	Log-Reg/XGB/SVM	Log-Reg	**60-20**	XGB	XGB
60-40	SVM	Log-Reg/SVM	**50-20**	XGB	XGB
50-50	SVM	XGB	**40-20**	XGB	XGB
40-60	XGB	XGB	**30-20**	Log-Reg	XGB
30-70	SVM	XGB	**20-20**	XGB	XGB
20-80	XGB	XGB	**10-20**	Log-Reg/SVM	XGB
10-90	XGB	XGB			

Table 4.2: The model that achieved the highest accuracy for each train-test split under each condition. Models were only included if they achieved better than baseline accuracy.

4.2.1 Data Exploration

To better understand the data collected, a report was generated for each participant consisting of statistical measures of the data. Additionally, a bivariate analysis was run to observe interactions between different features. On average, pumps were active 25% of the time the participants interacted with the jacket and the average number of pump activations was 18.4. The number of pump activations and the length of active time varied greatly amongst participants, as can be seen in Table 4.3. Taking the standard deviation of kurtosis and skew values indicated that there was a large variation across participants in the biometric data collected.

Hexplots were also used to observe relationships between different features. Some relationships were expected and looking at these plots acted as confirmation of the correctness of our data. For example, beats per minute was correlated with blood volume pressure. As beats per minute is derived from blood volume pressure using libraries built for the heart rate sensor we used, this correlation is expected. Related to the previous observation on data shape, correlations between variables other than bvp and bpm varied across participants.

	Number Activations	Percentage of Time On	Percentage of Time Off
P1	11	1.74	98.26
P2	0	0	100
P3	29	17.07	82.93
P4	13	47.37	52.63
P5	45	34.18	65.82
P6	21	46.42	53.58
P7	10	3.04	96.96
P8	15	40.52	59.48
P9	excluded		
P10	excluded		
P11	excluded		
P12	3	12.81	87.19

Table 4.3: Summary of interactions for participants, including number of times the pump was activated, and the total length of time the pump was active.

4.2.2 Condition A

Condition A uses traditional train test splits in increments of 10% to train each of the three models to predict the label of the current time step. Under this condition, at least one model under at least one training condition was able to achieve better than baseline accuracy for all 8 participants. XGBoost had the highest accuracy across participants more frequently than SVM or Logistic Regression. As seen in Table 4.2, XGBoost was also able to predict above baseline accuracy across participants in 67% of train-test splits, followed by SVM with 56%, and Logistic Regression with 22%. XGBoost also showed the lowest variance in model accuracy across train-test splits under this condition.

Observing F1-scores, 4.4 also gives insights into model performance. The formula for F1-scores can be found in the Equation 2.3. Looking at the aggregate F1-scores across participants, XGBoost had the highest, on average, F1-score for the pump-off label. There was a dramatic decrease in performance for the pump-on label, and Logistic Regression had the highest F1-score on average. There was minimal deviation in the F1-scores across the models, although XGBoost did have the lowest average standard deviation for both pump-off and pump-on F1-scores.

	Average						**Standard Deviation**					
	0			**1**			**0**			**1**		
Split	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**
90-10	0.62	0.83	0.66	0.26	0.14	0.22	0.39	0.15	0.45	0.36	0.35	0.40
80-20	0.65	0.77	0.65	0.36	0.29	0.33	0.41	0.22	0.44	0.35	0.40	0.40
70-30	0.64	0.74	0.61	0.40	0.36	0.39	0.39	0.25	0.43	0.35	0.36	0.41
60-40	0.72	0.71	0.71	0.52	0.30	0.43	0.33	0.31	0.36	0.39	0.37	0.42
50-50	0.71	0.75	0.71	0.38	0.32	0.34	0.33	0.23	0.36	0.36	0.34	0.43
40-60	0.49	0.65	0.46	0.31	0.35	0.29	0.45	0.40	0.46	0.33	0.35	0.37
30-70	0.17	0.76	0.38	0.46	0.33	0.52	0.25	0.17	0.47	0.44	0.15	0.41
20-80	0.48	0.56	0.52	0.35	0.38	0.38	0.33	0.39	0.39	0.38	0.36	0.41
10-90	0.69	0.61	0.65	0.27	0.42	0.27	0.21	0.25	0.21	0.27	0.30	0.38

Table 4.4: The average and standard deviation of F1-Scores aggregated across participants under Condition A.

4.2.3 Condition B

Condition B uses traditional train test splits in increments of 10% to train each of the three models to predict the label of the next time step. Under this condition, at least one model under at least one training condition was able to achieve better than baseline accuracy for 7 out of 8 participants. Logistic Regression and XGBoost had the highest accuracy across participants more frequently than SVM. As seen in Table 4.2, XGBoost was also able to predict above baseline accuracy across participants in 56% of train-test splits, followed by Logistic-Regression with 33%, and SVM with 22%. Logistic-Regression and XGBoost had the lowest variance in model accuracy across train-test splits under this condition.

	Average						**Standard Deviation**					
	0			**1**			**0**			**1**		
Split	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**
90-10	0.73	0.79	0.78	0.25	0.17	0.29	0.43	0.28	0.36	0.46	0.36	0.49
80-20	0.80	0.79	0.56	0.16	0.29	0.18	0.34	0.24	0.46	0.36	0.39	0.35
70-30	0.59	0.65	0.49	0.23	0.33	0.17	0.42	0.37	0.44	0.34	0.34	0.23
60-40	0.55	0.69	0.49	0.29	0.30	0.26	0.41	0.32	0.47	0.36	0.34	0.25
50-50	0.76	0.75	0.52	0.24	0.33	0.20	0.30	0.22	0.44	0.34	0.36	0.26
40-60	0.61	0.65	0.57	0.22	0.42	0.19	0.40	0.42	0.41	0.35	0.32	0.25
30-70	0.53	0.67	0.44	0.33	0.33	0.30	0.38	0.35	0.42	0.34	0.35	0.30
20-80	0.52	0.76	0.47	0.32	0.34	0.44	0.41	0.21	0.40	0.22	0.37	0.26
10-90	0.53	0.71	0.33	0.36	0.19	0.51	0.41	0.19	0.47	0.25	0.27	0.20

Table 4.5: The average and standard deviation of F1-Scores aggregated across participants under Condition B.

Similarly to Condition A, under Condition B, XGBoost had the highest average F1-score 4.5 for pump-off. In this case, XGBoost also had the highest F1-score for pump-on. In terms of standard deviation, XGBoost had the lowest value for the pump-off label and SVM had the lowest value for the pump-on label.

4.2.4 Condition C

Condition C uses incremental backwards increases of 10% with a fixed testing set of 20% to train each of the three models to predict the label of the current time step. Under this condition, at least one model under at least one training condition was able to achieve better than baseline accuracy for 6 out of 8 participants. XGBoost had the highest accuracy across participants more frequently than SVM or logistic-regression. As seen in Table 4.2, XGBoost was also able to predict above baseline accuracy across participants in 75% of train-test splits, followed by Logistic-Regression with 50%, and SVM with 13%. Logistic Regression had the lowest variance in model accuracy across train-test splits under this condition.

	Average						Standard Deviation					
	0			1			0			1		
Split	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**
80-20	0.65	0.77	0.65	0.36	0.29	0.33	0.41	0.22	0.44	0.35	0.40	0.40
70-20	0.66	0.76	0.76	0.39	0.31	0.33	0.40	0.25	0.39	0.34	0.39	0.41
60-20	0.66	0.77	0.64	0.31	0.30	0.20	0.35	0.30	0.37	0.25	0.40	0.29
50-20	0.67	0.79	0.64	0.21	0.30	0.24	0.28	0.26	0.23	0.29	0.39	0.30
40-20	0.84	0.80	0.81	0.21	0.35	0.24	0.15	0.20	0.18	0.31	0.42	0.33
30-20	0.77	0.81	0.71	0.23	0.36	0.27	0.23	0.18	0.28	0.31	0.40	0.31
20-20	0.82	0.82	0.81	0.23	0.47	0.30	0.19	0.34	0.22	0.29	0.41	0.30
10-20	0.88	0.80	0.87	0.49	0.46	0.47	0.13	0.25	0.13	0.34	0.38	0.33

Table 4.6: The average and standard deviation of F1-Scores aggregated across participants under Condition C.

Under Condition C, XGBoost marginally outperforms Logistic Regression and SVM for the pump-off F1-score 4.6, as well as outperforming both the pump-on F1-score. Standard deviations were similar for the pump-off conditions, with XGBoost being the lowest value. For the pump-on condition, Logistic Regression had the lowest standard deviation across participants and train-test splits.

4.2.5 Condition D

Condition D uses incremental backwards increases of 10% with a fixed to testing set of 20% to train each of the three models to predict the label of the next time step. Under this condition, at least one model under at least one training condition was able to achieve better than baseline accuracy for 6 out of 8 participants. XGBoost had the highest accuracy across participants more frequently than SVM or logistic-regression.

As seen in Table 4.2, XGBoost was also able to predict above baseline accuracy across participants in 88% of train-test splits, followed by Logistic-Regression with 25%, and SVM with 13%. Logistic-Regression and XGBoost had the lowest variance in model accuracy across train-test splits under this condition.

In this condition, XGBoost had the highest average F1-score 4.7 for both pump-on and pump-off aggregated across participants. Additionally, for the pump-off label, these scores showed the lowest standard deviation using the XGBoost model. However, for the pump-on condition, Logistic Regression had the lowest standard deviation.

	Average						**Standard Deviation**					
	0			**1**			**0**			**1**		
Split	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**	**LR**	**XGB**	**SVM**
80-20	0.78	0.74	0.55	0.19	0.25	0.18	0.33	0.32	0.47	0.37	0.37	0.35
70-20	0.69	0.77	0.69	0.22	0.28	0.22	0.39	0.28	0.42	0.40	0.37	0.39
60-20	0.53	0.76	0.53	0.14	0.26	0.16	0.48	0.29	0.49	0.24	0.38	0.26
50-20	0.74	0.83	0.75	0.19	0.37	0.19	0.34	0.17	0.38	0.30	0.39	0.28
40-20	0.58	0.83	0.69	0.12	0.34	0.38	0.42	0.14	0.42	0.25	0.39	0.43
30-20	0.40	0.83	0.62	0.19	0.42	0.40	0.45	0.15	0.41	0.27	0.39	0.42
20-20	0.72	0.78	0.71	0.20	0.38	0.20	0.37	0.32	0.38	0.28	0.37	0.29
10-20	0.53	0.90	0.67	0.23	0.43	0.23	0.46	0.13	0.41	0.27	0.43	0.29

Table 4.7: The average and standard deviation of F1-Scores aggregated across participants under Condition D.

4.2.6 Discussion of Quantitative Results

By observing the results under each condition, we can begin to determine which classifier performed the best. It should be noted that while training, some training sets only contained a single label, leading to errors in predictions on the testing set. In these cases, an unseen label error would be thrown during testing. This error arose from the nature of our training and testing splits. Smaller sets did not contain label diversity, especially in the cases where participants chose to have the pump on or off for extended periods of time.

As mentioned in each condition's results above, a small number of participants under the different conditions were always outperformed by the majority class baseline. These participants were P4 under Condition B, and P1 and P7 under both Conditions C and D. The majority class prediction for P4 was 53%, while the majority class predictions for P1 and P7 were 97% and 98% respectively. P1 and P7 have a proportionally larger dataset imbalance compared to other participants. This imbalance was magnified by the application of the sliding window for next step predictions used in Conditions C and D. This could explain why models for these participants saw a drop in accuracy compared to the baseline only under these two conditions. It should also be noted that under conditions A and B for P1 and P7, the performance above baseline in terms of accuracy was marginal. In the case of P4 there appear to be no strong indicators in the statistical analysis as to why performance on their data was sub-baseline. There is the possibility of an undetected sensor fault, and further exploration could be conducted to determine this.

Observing the aggregate data, trends in model choice become evident. Looking first at the variance in model accuracy, XGBoost and Logistic Regression show more stability across train-test splits over SVM. When choosing a model that has the best accuracy for each condition and train-test split, it doesn't appear to matter much whether Logistic Regression or XGBoost are selected. However, when cases that don't perform better than baseline are removed, a pattern emerges. XGBoost does better at predicting better than baseline accuracy than both Logistic-Regression and SVM. Additionally, under Conditions B and D, where the next time step was predicted, XGBoost performs better as the training dataset becomes smaller. This trend is not mirrored in Conditions A and C, however, XGBoost still performs best at better than baseline accuracy for the majority of train-test splits. Across participants, there was not a distinct difference between predicting the current time step in Conditions A and C, or the next time step in Conditions B and D.

Observing F1-scores across participants reinforces the use of XGBoost, as it generally had higher F1-scores across participants and train-test sets under all 4 conditions. Additionally, F1-scores show lower variance on average under XGBoost, than Logistic Regression and SVM. Interestingly, F1-scores drop significantly when predicting pump-off versus pump-on. This could be due to an imbalance in the dataset. It was also observed that using the sliding window reduced the size of the dataset, thereby magnifying any misclassifications.

Overall, XGBoost outperforms other models across the testing conditions. Under traditional train test splits with larger amounts of training data, XGBoost can be outperformed by other models. While Logistic Regression is able to outperform in some of these instances, XGBoost provides consistency and performs better when smaller amounts of training data are available. Further discussion of quantitative results in conjunction with findings from the user study will be discussed in Chapter 5.

Chapter 5

Discussion

In order to explore the relationship between emotions and physiological data, we have constructed two interactive jackets that collect heart rate, temperature, and movement. These jackets also apply pressure to the sides with the push of a button. We had 12 participants interact with the garment for 50 minutes, followed by an interview portion. From this study, both qualitative and quantitative results were collected. Qualitative results were analyzed using thematic analysis [7]. Quantitative data was used to train three classifiers, Logistic Regression, SVM, and XGBoost, to predict when users decided to turn on and off the pump. Here, we discuss the results found from both the qualitative and quantitative data, situate our work within the existing space, as well as present the limitations of our current study and possible future directions.

5.1 Reflections

Here, we reflect on the three research questions posed in Section 1.

(RQ1) How can interactive garments support wearers in managing daily stress?

Our work has validated the findings of other authors. Participants discussed different emotions that they felt while completing school-related tasks. These feelings can be extrapolated to other areas of life, such as the work environment. This relates to works done by [24] that shows that workplace stressors can impact both physiology and behaviour.

In Theme 4, participants echoed the desire for a felt interface that has been captured by other researchers [41, 37, 19]. Emotion regulation habits across participants using movement reminders situates itself with other tangible reminders and notifications crafted by other authors [32, 26, 58]. However, our work differs from these as we focus on stimulating another sense other than vision, as vision is highly active when performing laptop or pen+paper work. As an inadvertent by-product of our use of air pumps for compression, we touched on the use of vibrotactile feedback, which has been explored for notification use and heart rate awareness [62].

Unlike [23, 38, 65] the biological data collected through this garment is not merely being observed but extracted to be used as training data. While compression has been explored in other settings [45, 33] our work is unique as it focuses on a general negative emotion, rather than notifications or awareness of breathing. Additionally, pump activation was done manually by participants at times they felt was best for them, rather than following a fixed pattern.

As is evident in Themes 2 and 5, participants' likes and dislikes vary greatly. What may be beneficial to some, like the sound of pumps being meditative, can be distracting and undesirable to others. The same need for customization arises within the quantitative results. When observing the physiological data from participants, there are differences in heart rates, movement patterns, and pump activations. This leads to a need to train models on a participant-by-participant basis. In a practical setting, unique customization to a wearer must be done efficiently and effectively. This leads into our quantitative analysis and how we have laid the foundation for on-device predictions customized to each wearer.

(RQ2) Can we train machine learning models to predict when wearers choose to activate garment compression? By looking at different training and testing splits for each participant's data, we were able to observe the amount of data necessary to train a model to predict when the wearer would like to activate compression on their sides. On the tail ends, when there is either 10% training or 10% testing, all models tend to behave poorly. In some cases, unseen label errors occur because the size of the testing set does not include both on and off labels. As the amount of training data increases, we see an increase in performance using XGBoost as our classifier. Even with an increase to 30% training with a 20% or 70% testing, the model was able to predict when the wearer would like to activate the pump with a better-than-baseline accuracy, especially under next-step prediction. These results give us an understanding of the type and amount of data necessary to train a model to make accurate predictions for a wearer. First, both labels must be present in the dataset. Secondly, 30% is a reasonable amount of data to start making better-than-baseline predictions. In a 50 minutes session, 15 minutes corresponds to 30% of the data. Additionally, the use of XGBoost allows us to train a model in under a minute for this amount of data.

(RQ3) If so, is this training and inference able to work in a low-resource setting? In Theme 6, participants feel that automating the garment to apply compression automatically would be beneficial, reducing the mental burden that comes along with identifying and treating one's emotions. This presents a low resource setting to explore, where on-device predictions can be made using a compressed machine learning model. We have shown that training of such a model can be done efficiently, given that training done using XGBoost is fast. We have also demonstrated that a small amount of training data can be used to train a model effectively. Indeed, as discussed above, a 15-minute data sample is enough to achieve better-than-baseline accuracy using XGBoost.

5.2 Limitations

There are some limitations to the scope of the work conducted. Given time constraints, two jackets were constructed that may not be suitable for everyday wear and tear and did not fit the styles of all participants. This could impact the likelihood of wearers using the jacket in everyday life. To address this, additional time and materials should be used to ensure the durability of the prototypes. Further studies could also be completed to demonstrate the quantitative effectiveness of the interactive garment sensors used. Additionally, a minimal analysis into feature interactions was completed. Future work could provide a more thorough analysis between features. By understanding interactions between features, unimportant features could be removed and their sensing would not be required, allowing for a simpler jacket and algorithm design.

A small number of participants were collected from the School of Computing and asked to complete school tasks. A larger range of participants of different ages, and less familiarity with technology could present different results, especially in terms of sentiments towards garment automation. To address this, further studies can be completed with a broader range of participant demographics.

Lastly, the study lasted for a maximum period of 2 hours. Currently, the interactive garment needs to pass data through a cable to a laptop to accurately collect it for training. This means it is not mobile. Adding functionality to the prototype to allow participants to move around their daily lives with the garment, while still collecting data, could allow participants to interact with the garment in their own space and over a longer period of time. An interaction period greater than 50 minutes could allow participants to gain more familiarity with the garment. This could add greater insights into how participants use the garment for emotion regulation. Additionally, this could help address any confounding factors that were introduced by having participants in an unfamiliar setting while being observed.

5.3 Future Work

We envision two major areas for further research. First, we see the exploration of customized predictive models *on device*. The board used in this study, the Bluefruit Circuit Playground, can be loaded with a trained machine learning model using a compressed C++ model [1]. By collecting a 15-minute sample from a wearer, a model can be trained in under a minute and easily loaded onto the device. In addition to this customized model, techniques from reinforcement learning could help update our policy over time to further adapt to the user. From here, exploration of how effective the automation is can be explored, as well as participant sentiments towards this beyond their perceived response.

Secondly, further work into the customization of actuation for emotion regulation can be explored. This could give insights into how jacket style and appearance could affect the effectiveness of emotion regulation, as well as willingness to engage in this type of emotion regulation. In addition, an inadvertent side effect of air pumps was the desire for vibrotactile feedback on the body. Others desired the sensation of a massage, or the use of sound to help regulate feelings of stress or distraction. The use of customizable actuations in conjunction with automated actuation could be further explored to help emotion regulation. This exploration into customizability could also be broadened to include other factors associated with school work, including effects on productivity. Work could explore connecting to other devices typically used while completing school work, like a laptop, and employ unique actuations to increase productivity and decrease negative emotions like stress.

5.4 Conclusion

In this work, we explored the use of interactive jackets that provide compression to the sides of a user as an emotion regulation technique during school-related tasks. We have reported interview findings with 12 participants and analyzed their physiological data to accurately predict when they would like to experience compression. This work presents an implementation of compression in a wearable garment and serves as a stepping block for further iterations of garments with onboard predictions that are customized to the wearer.

Reflections on emotions and participant interactions are categorized into 6 themes: (1) Stress and Distraction Shape our Daily Lives, (2) Everyone Copes Differently, (3) Emotion Regulation Requires Human Initiative, (4) Physical Interactions Ground Individuals, (5) Emotional Garments Are Not One Size Fits All, and (6) Automation is Encouraged to a Point. Within these themes, we uncover that participants liked to apply compression for different reasons. Participants also have strong preferences towards design qualities and actuation. Compression was useful, being compared to a hug, however, others did not feel the compression and enjoyed the vibrotactile sensations produced by the air pumps. Lastly, participants expressed interest in interactive garment automation, as long as there are ways to override the system if it actuates at the wrong time.

Upon analyzing the data collected from participants, we observed that XGBoost is able to accurately predict when users have the pump on for the current time step and the next time step. XGBoost also requires minimal training data to achieve higher accuracies than majority-class predictions. Not only does XGBoost provide consistency, but it is a low-resource and fast model. This facilitates future work of implementing machine learning predictions on microcontrollers by compressing the machine learning model. After exploring both the qualitative and quantitative aspects of interactive garments for emotion regulation, we found that participants are interested in the idea of automated compression. Additionally, easy-to-implement machine learning models are capable of determining when participants would like to engage with a jacket using personal training data for each participant. Ultimately, our work demonstrates the future of interactive garments that support customized machine learning predictions for emotion regulation on the go.

Bibliography

[1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

[2] R. A. Abbasi, N. Javaid, M. N. J. Ghuman, Z. A. Khan, S. Ur Rehman, et al. Short term load forecasting using xgboost. In *Workshops of the International Conference on Advanced Information Networking and Applications*, pages 1120–1131. Springer, 2019.

[3] Adafruit. Finger pulse oximeter. In *Available online: https://www.adafruit.com/product/4582 (accessed on 20 April 2021)*, 2022.

[4] Apple. Watch series 6. In *Available online: https://www.apple.com/ca/apple-watch-series-6/ (accessed on 20 April 2021)*, 2021.

[5] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, pages 59–66, 2018. doi: 10.1109/FG.2018.00019.

[6] J. Berzowska and M. Coelho. Kukkia and vilkas: Kinetic electronic garments. In *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, pages 82–85. IEEE, 2005.

[7] V. Braun and V. Clarke. Thematic analysis. *APA handbook of research methods in psychology*, 2012.

[8] P. Bruno, V. Melnyk, and F. Völckner. Temperature and emotions: Effects of physical temperature on responses to emotional advertising. *International Journal of Research in Marketing*, 34(1):302–320, 2017.

[9] L. Buechley and H. Perner-Wilson. Crafting technology: Reimagining the processes, materials, and cultures of electronics. *ACM Trans. Comput.-Hum. Interact.*, 19(3), oct 2012. ISSN 1073-0516. doi: 10.1145/2362364.2362369. URL `https://doi-org.proxy.queensu.ca/10.1145/2362364.2362369`.

[10] R. Y. Cai, A. L. Richdale, C. Dissanayake, and M. Uljarević. Resting heart rate variability, emotion regulation, psychological wellbeing and autism symptomatology in adults with and without autism. *International Journal of Psychophysiology*, 137:54–62, 2019.

[11] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408:189–215, 2020.

[12] L. Chao, J. Tao, M. Yang, Y. Li, and Z. Wen. Long short term memory recurrent neural network based multimodal dimensional emotion recognition. In *Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge*, AVEC '15, page 65–72, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450337434. doi: 10.1145/2808196.2811634. URL `https://doi-org.proxy.queensu.ca/10.1145/2808196.2811634`.

[13] T. Chen and C. Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4232-2. doi: 10.1145/2939672.2939785. URL `http://doi.acm.org/10.1145/2939672.2939785`.

[14] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342322. doi: 10.1145/2939672.2939785. URL `https://doi-org.proxy.queensu.ca/10.1145/2939672.2939785`.

[15] Y. Cho, S. J. Julier, and N. Bianchi-Berthouze. Instant stress: detection of perceived mental stress through smartphone photoplethysmography and thermal imaging. *JMIR mental health*, 6(4):e10140, 2019.

[16] H. Christensen and U. Lundberg. Musculoskeletal problems as a result of work organization, work tasks and stress during computer work. *Work & Stress*, 2002.

[17] A. Chunawale and M. Bedekar. Human emotion recognition using physiological signals: A survey. In *2nd International Conference on Communication & Information Processing (ICCIP)*, 2020.

[18] N. Dabby, A. Aleksov, E. Lewallen, S. Oster, R. Fygenson, B. Lathrop, M. Bynum, M. Samady, S. Klein, and S. Girouard. A scalable process for manufacturing integrated, washable smart garments applied to heart rate monitoring. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 38–41, 2017.

[19] L. Devendorf, J. Lo, N. Howell, J. L. Lee, N.-W. Gong, M. E. Karagozler, S. Fukuhara, I. Poupyrev, E. Paulos, and K. Ryokai. " i don't want to wear a screen" probing perceptions of and possibilities for dynamic displays on clothing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 6028–6039, 2016.

[20] V. Dissanayake, D. S. Elvitigala, H. Zhang, C. Weerasinghe, and S. Nanayakkara. Comprate: Power efficient heart rate and heart rate variability monitoring on smart wearables. In *25th ACM Symposium on Virtual Reality Software and Technology*, pages 1–8, 2019.

[21] Empatica. E4 wristband. In *Available online: https://store.empatica.com/products/e4-wristband?variant=17719039950953/ (accessed on 20 April 2021)*, 2021.

[22] S. Endow, H. Moradi, A. Srivastava, E. G. Noya, and C. Torres. Compressables: A haptic prototyping toolkit for wearable compression-based interfaces. In *Designing Interactive Systems Conference 2021*, DIS '21, page 1101–1114, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450384766. doi: 10.1145/3461778.3462057. URL `https://doi.org/10.1145/3461778.3462057`.

[23] O. Ettehadi, L. Jones, and K. Hartman. Heart waves: A heart rate feedback system using water sounds. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 527–532, 2020.

[24] M. Fellmann, F. Lambusch, and L. Pieper. It-supported stress management: Mediating between sensor data and individual appraisal. In *Proceedings of the 5th International Workshop on Sensor-Based Activity Recognition and Interaction*, iWOAR '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450364874. doi: 10.1145/3266157.3266223. URL `https://doi-org.proxy.queensu.ca/10.1145/3266157.3266223`.

[25] Fitbit. Charge 4. In *Available online: https://www.fitbit.com/global/en-ca/products/trackers/charge4?sku=417BKBK (accessed on 20 April 2021)*, 2021.

[26] E. Fransén Waldhör, P. Vierne, P. Seidler, B. Greinke, and K. Bredies. E-textile production of wearable ambient notification devices. In *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*, DIS '17 Companion, page 309–312, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450349918. doi: 10.1145/3064857.3079181. URL `https://doi-org.proxy.queensu.ca/10.1145/3064857.3079181`.

[27] V. Ganti, S. Rachakonda Kirthana, G. S. Kumar, and R. Sumalatha. Human emotion recognition using bio signals. *International Journal*, 8(5), 2020.

[28] Garmin. Vivo active 4. In *Available online: https://buy.garmin.com/en-CA/CA/p/643382 (accessed on 20 April 2021)*, 2021.

[29] M. Harraghy, D. Calderon, R. Lietz, J. Brady, F. Makedon, and E. Becker. A review of wearable heart rate sensors in research. In *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, pages 315–316, 2019.

[30] iMotions. Shimmer gsr+. In *Available online: https://imotions.com/hardware/shimmer3-gsr/ (accessed on 20 April 2021)*, 2021.

[31] A. Industries. Bluefruit. In *Available online: https://learn.adafruit.com/adafruit-circuit-playground-bluefruit?view=all (accessed on 3 April 2022)*, 2022.

[32] L. Jones, J. McClelland, P. Thongsouksanoumane, and A. Girouard. Ambient notifications with shape changing circuits in peripheral locations. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, page 405–408, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346917. doi: 10.1145/3132272.3132291. URL `https://doi.org/10.1145/3132272.3132291`.

[33] A. Jung, M. Alfaras, P. Karpashevich, W. Primett, and K. Höök. Exploring awareness of breathing through deep touch pressure. In *Proceedings of the 2021*

CHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380966. URL `https://doi.org/10.1145/3411764.3445533`.

[34] E. Kamioka et al. Emotions detection scheme using facial skin temperature and heart rate variability. In *MATEC Web of Conferences*, volume 277, page 02037. EDP Sciences, 2019.

[35] M. E. Kemeny. The psychobiology of stress. *Current directions in psychological science*, 12(4):124–129, 2003.

[36] C. Kirschbaum, S. Wüst, and D. Hellhammer. Consistent sex differences in cortisol responses to psychological stress. *Psychosomatic medicine*, 54(6):648–657, 1992.

[37] Y. S. Lee. Spiky starfish: Exploring 'felt technology' through a shape changing wearable bag. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '15, page 419–420, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450333054. doi: 10.1145/2677199.2690878. URL `https://doi-org.proxy.queensu.ca/10.1145/2677199.2690878`.

[38] M. Mauriello, M. Gubbels, and J. E. Froehlich. Social fabric fitness: the design and evaluation of wearable e-textile displays to support group running. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2833–2842, 2014.

[39] J. L. Moraes, M. X. Rocha, G. G. Vasconcelos, J. E. Vasconcelos Filho, V. H. C.

De Albuquerque, and A. R. Alexandria. Advances in photopletysmography signal analysis for biomedical applications. *Sensors*, 18(6):1894, 2018.

[40] S. Muthukumarana, M. A. Messerschmidt, D. J. Matthies, J. Steimle, P. M. Scholl, and S. Nanayakkara. Clothtiles: A prototyping platform to fabricate customized actuators on clothing using 3d printing and shape-memory alloys. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380966. doi: 10.1145/3411764.3445613. URL `https://doi-org.proxy.queensu.ca/10.1145/3411764.3445613`.

[41] S. Nabil, A. Everitt, M. Sturdee, J. Alexander, S. Bowen, P. Wright, and D. Kirk. Actueating: Designing, studying and exploring actuating decorative artefacts. In *Proceedings of the 2018 Designing Interactive Systems Conference*, DIS '18, page 327–339, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450351980. doi: 10.1145/3196709.3196761. URL `https://doi-org.proxy.queensu.ca/10.1145/3196709.3196761`.

[42] S. Nabil, J. Kučera, N. Karastathi, D. S. Kirk, and P. C. Wright. Seamless seams: Crafting techniques for embedding fabrics with interactive actuation. In *Conference on Designing Interactive Systems*, pages 987–999. ACM, 2019.

[43] T. pandas development team. pandas-dev/pandas: Pandas, Feb. 2020. URL `https://doi.org/10.5281/zenodo.3509134`.

[44] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos,

D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[45] H. Pohl, D. Becke, E. Wagner, M. Schrapel, and M. Rohs. Wrist compression feedback by pneumatic actuation. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, page 315–318, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450331463. doi: 10.1145/2702613.2725427. URL `https://doi-org.proxy.queensu.ca/10.1145/2702613.2725427`.

[46] H. F. Posada-Quintero and K. H. Chon. Innovations in electrodermal activity data collection and signal processing: A systematic review. *Sensors*, 20(2):479, 2020.

[47] I. Posch, L. Stark, and G. Fitzpatrick. Etextiles: Reviewing a practice through its tool/kits. In *Proceedings of the 23rd International Symposium on Wearable Computers*, ISWC '19, page 195–205, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368704. doi: 10.1145/3341163.3347738. URL `https://doi.org/10.1145/3341163.3347738`.

[48] I. Poupyrev, N.-W. Gong, S. Fukuhara, M. E. Karagozler, C. Schwesig, and K. E. Robinson. Project jacquard: interactive digital textiles at scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4216–4227, 2016.

[49] R. Rakshit, V. R. Reddy, and P. Deshpande. Emotion detection and recognition using hrv features derived from photoplethysmogram signals. In *Proceedings of*

the 2nd workshop on Emotion Representations and Modelling for Companion Systems, pages 1–6, 2016.

[50] R. Rakshit, V. R. Reddy, and P. Deshpande. Emotion detection and recognition using hrv features derived from photoplethysmogram signals. In *Proceedings of the 2nd Workshop on Emotion Representations and Modelling for Companion Systems*, ERM4CT '16, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450345583. doi: 10.1145/3009960.3009962. URL `https://doi-org.proxy.queensu.ca/10.1145/3009960.3009962`.

[51] S. Reynolds, S. J. Lane, and B. Mullen. Effects of deep pressure stimulation on physiological arousal. *The American Journal of Occupational Therapy*, 69(3): 6903350010p1–6903350010p5, 2015.

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

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

[54] T. S. Saini, M. Bedekar, and S. Zahoor. Circle of emotions in life: Emotion mapping in 2dimensions. In *Proceedings of the 9th International Conference on Computer and Automation Engineering*, ICCAE '17, page 83–88, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450348096. doi: 10.1145/3057039.3057046. URL `https://doi-org.proxy.queensu.ca/10.1145/3057039.3057046`.

[55] P. Sarkar and A. Etemad. Self-supervised learning for ecg-based emotion recognition. In *2020 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2020, Barcelona, Spain, May 4-8, 2020*, pages 3217–3221. IEEE, 2020. doi: 10.1109/ICASSP40776.2020.9053985. URL `https://doi.org/10.1109/ICASSP40776.2020.9053985`.

[56] P. Sarkar and A. Etemad. Self-supervised learning for ecg-based emotion recognition. In *2020 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2020, Barcelona, Spain, May 4-8, 2020*, pages 3217–3221. IEEE, 2020. doi: 10.1109/ICASSP40776.2020.9053985. URL `https://doi.org/10.1109/ICASSP40776.2020.9053985`.

[57] P. Schmidt, A. Reiss, R. Dürichen, C. Marberger, and K. V. Laerhoven. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In S. K. D'Mello, P. G. Georgiou, S. Scherer, E. M. Provost, M. Soleymani, and M. Worsley, editors, *Proceedings of the 2018 on International Conference on Multimodal Interaction, ICMI 2018, Boulder, CO, USA, October 16-20, 2018*, pages 400–408. ACM, 2018. doi: 10.1145/3242969.3242985. URL `https://doi.org/10.1145/3242969.3242985`.

[58] A. Singh, S. Nabil, A. Roudaut, and A. Girouard. Co-designing tangible break reminders with people with repetitive strain injury. In *IFIP Conference on Human-Computer Interaction*, pages 289–311. Springer, 2021.

[59] S. Speer, E. Hamner, M. Tasota, L. Zito, and S. K. Byrne-Houser. Mindfulnest: Strengthening emotion regulation with tangible user interfaces. In *Proceedings of the 2021 International Conference on Multimodal Interaction*, ICMI '21, page

103–111, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450384810. doi: 10.1145/3462244.3479917. URL `https://doi-org.proxy.queensu.ca/10.1145/3462244.3479917`.

[60] G. Udovičić, J. Derek, M. Russo, and M. Sikora. Wearable emotion recognition system based on gsr and ppg signals. In *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care*, pages 53–59, 2017.

[61] M. Umair, C. Sas, and M. Alfaras. Thermopixels: Toolkit for personalizing arousal-based interfaces through hybrid crafting. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, pages 1017–1032, 2020.

[62] M. Umair, C. Sas, N. Chalabianloo, and C. Ersoy. Exploring personalized vibrotactile and thermal patterns for affect regulation. In *Designing Interactive Systems Conference 2021*, DIS '21, page 891–906, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450384766. doi: 10.1145/3461778.3462042. URL `https://doi.org/10.1145/3461778.3462042`.

[63] Z. Wang, Z. Yu, B. Zhao, B. Guo, C. Chen, and Z. Yu. Emotionsense: An adaptive emotion recognition system based on wearable smart devices. *ACM Transactions on Computing for Healthcare*, 1(4):1–17, 2020.

[64] J. C. Y. Wong, J. Wang, E. Y. Fu, H. V. Leong, and G. Ngai. Activity recognition and stress detection via wristband. In *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia*, pages 102–106, 2019.

[65] B. Yu, N. Bongers, A. van Asseldonk, J. Hu, M. Funk, and L. M. G. Feijs. Livingsurface: Biofeedback through shape-changing display. In S. Bakker, C. Hummels, B. Ullmer, L. Geurts, B. Hengeveld, D. Saakes, and M. Broekhuijsen, editors, *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction, Eindhoven, The Netherlands, February 14-17, 2016*, pages 168–175. ACM, 2016. doi: 10.1145/2839462.2839469. URL `https://doi.org/10.1145/2839462.2839469`.

[66] S. Zaman, X. Tao, C. Cochrane, and V. Koncar. Development of e-textile electrodes: washability and mechanical stresses. In *IOP Conference Series: Materials Science and Engineering*, volume 827, page 012025. IOP Publishing, 2020.

Appendix A

Appendix

A.1 Interview Questions

The following list of questions was asked to participants during the interview portion. Questions 1-7 focus on general experiences and questions 8-19 focus specifically on garment interaction.

1. Do you experience stress, anxiety, distraction or other similar negative emotions while completing school related tasks (ex. assignments, lectures, answering emails)? How frequently do you experience these negative emotions?

2. Are there certain times or settings where you experience these negative emotions more frequently or intensely?

3. How do you manage these emotions? Is there one technique that helps you more than others? Which one?

4. How do these emotions affect your ability to go about your day? Which tasks become more difficult?

5. Which, if any, modifications have you made to your school life to avoid feelings of stress, anxiety, distraction or negative emotions?

6. What are your thoughts on: controlled breathing, breathing timers, movement reminders, meditation, weight/pressure (like a weighted blanket)? (go through each individually)

7. Can you elaborate on any of the above further, or reflect on any regulation techniques you engage in that aren't listed above?

8. Without going into specifics, what task were you working on during this experiment? (ex. Answering emails, starting an assignment)

9. At which points did you decide to inflate the airbags and apply pressure? What triggered this desire to apply pressure?

10. When did you decide to deflate the airbags and stop applying pressure? What triggered this desire?

11. What did applying and removing pressure cause you to experience?

12. How do you feel that the pressure-actuation from the garment impacted your emotions? Would you like it in another location on the body?

13. What do you think about automating the garment to inflate/deflate on its own without you pressing the button (ex. randomly, according to some threshold)?

14. How do you see yourself integrating this device into your school or everyday life?

15. How do you see the design qualities of this interactive garment (e.g., in terms of fit/size, design/style, colour/pattern, material/texture)?
16. What changes or additions do you suggest?
17. What other shape changing notifications would you like to see?
18. How do you see this type of wearable interaction as different than using mobile app notifications (or stress relief)?
19. What other recommendations would you like to add?

A.2 Ethical Clearance

Figure A.1 shows the e-mail from the Queen's General Research Ethics Board approving this study, file number 6033235.

Queen's UNIVERSITY

September 15, 2021

Miss Victoria Armstrong
Queen's University

Title: "GCOMP-110-21 Designing physiologically-responsive e-textiles with machine learning techniques;" TRAQ # 6033235

Dear Miss Armstrong:

The General Research Ethics Board (GREB), by means of a delegated board review, has cleared your proposal entitled **"GCOMP-110-21 Designing physiologically-responsive e-textiles with machine learning techniques"** for ethical compliance with the Tri-Council Guidelines (TCPS 2) and Queen's ethics policies. In accordance with the Tri-Council Guidelines (Article 6.14) and Standard Operating Procedures (405), your project has been cleared for one year.

You are reminded of your obligation to submit an annual renewal form prior to the annual renewal due date (access this form at http://www.queensu.ca/traq/signon.html/; click on "Events;" under "Create New Event" click on "General Research Ethics Board Annual Renewal/Closure Form for Cleared Studies"). Please note that when your research project is completed, you need to submit an Annual Renewal/Closure Form in Romeo/traq indicating that the project is 'completed' so that the file can be closed. This should be submitted at the time of completion; there is no need to wait until the annual renewal due date.

You are reminded of your obligation to advise the GREB of any adverse event(s) that occur during this one-year period (access this form at http://www.queensu.ca/traq/signon.html/; click on "Events;" under "Create New Event" click on "General Research Ethics Board Adverse Event Form"). An adverse event includes, but is not limited to, a complaint, a change or unexpected event that alters the level of risk for the researcher or participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GREB within 48 hours.

You are also reminded that all changes that might affect human participants must be cleared by the GREB. For example, you must report changes to the level of risk, applicant characteristics, and implementation of new procedures. To submit an amendment form, access the application by at http://www.queensu.ca/traq/signon.html; click on "Events;" under "Create New Event" click on "General Research Ethics Board Request for the Amendment of Approved Studies." Once submitted, these changes will automatically be sent to the Ethics Coordinator, GREB, at University Research Services for further review and clearance by GREB or the Chair, GREB.

On behalf of the General Research Ethics Board, I wish you continued success in your research.

Sincerely,

Professor Dean A. Tripp, PhD
Chair, General Research Ethics Board (GREB)
Departments of Psychology, Anesthesiology & Urology
Queen's University

Figure A.1: The letter granting ethical approval for this study.