# COMPARISON OF SUBJECTIVE AND PHYSIOLOGICAL STRESS LEVELS IN HOME AND OFFICE WORK ENVIRONMENTS

Matthew Harper

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of PhD.

July 2023

# Acknowledgements

I would like to express my deepest gratitude to my supervisors Doctor Fawaz Ghali, Doctor Wasiq Khan, Professor Abir Hussain, and Professor Dhiya Al-Jumeily OBE for their dedicated help, time, and valuable guidance throughout the research project. Credit also goes to my colleagues Doctor Sulaf Assi, Megan Wilson, and Megan Watson for their invaluable suggestions and support in the final stages of the research project.

Many thanks to Tricia Waterson and the administration team of Faculty of Engineering and Technology for their support, especially in organising my attendance at conferences and completing the vital administrative milestones and examinations throughout the project.

I owe special thanks to my family for their endless support and encouragement. I would also like to thank my close friends for being with me at every stage of this PhD journey.

# Contents

## Table of Equations

# Abstract

Work stress is a major problem to individuals and society, with prolonged periods of stress often leading to health issues and reduced productivity. COVID-19 has increased the incidence of individuals working in a mixture of home and office-based environments, with each location presenting its own stressors. Identification of stress levels in each environment will allow individuals to better plan how to mitigate stress and boost productivity. In this project, differences in stress levels are predicted in each work environment from individuals' physiological responses and subjectively reported stress and productivity. Initial work on the project focused upon development of a system for the detection of dementia-related difficulties through the wearable-based tracking of physiological indicators. As such, a review of the available commercial and laboratory devices available for tracking physiological indicators of dementia-related difficulties was conducted. Furthermore, no publicly available physiological dataset for predicting difficulties in dementia currently exists. However, a review of the methods for collecting such a dataset and the impact of COVID-19 found that it is impractical and potentially unethical to conduct an experiment with people with dementia during the pandemic. As such, a pivot in research was necessitated. Comparing the stress levels of individuals working in home and office environments was selected. A data collection experiment was then performed with 13 academics working in combinations of home and office environments. Descriptive statistical features were then extracted from both the physiological and questionnaire data, with the relationships between attributes and features calculated using various advanced data analytics and statistical approaches. The resultant correlation coefficients and statistical summaries of stress were used to evaluate relationships between stress and work environment at different times of day, different days of the week, and while performing different activities. A bagged tree machine learning model was trained over the data, achieving 99.3% accuracy when evaluated using 10-fold cross validation. When tested on the purely unseen instances it achieved 56% accuracy corresponding to inter-class stress classification, however a testing accuracy of 73.7% was achieved using principal component analysis for dimensionality reduction and the dataset is balanced using Synthetic Minority Oversampling Technique.

# Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

# 1. Introduction

## 1.1. Research Background

Work stress is a detrimental physical or emotional response to situations experienced during an individual's working hours that do not match their capabilities, resources or requirements [1]. Work stress has many bad effects on individual employees' health, with long term exposure to elevated stress being associated with an increased risk of experiencing mental health problems [2], hypertension and cardiovascular disease [3, 4], obesity [5], sexual dysfunction [6] and hair loss [7]. Furthermore, the cost of this stress for society is substantial, with estimates of the overall costs of work-related stress being as high as $187 billion, considering the resultant productivity losses and increased health care burdens [8]. One method that has been proposed to help mitigate the impact of work-related stress is early detection and prediction of stress using physiological data collected using wearable computing devices [9]. These wearable devices are highly useful in this regard, due to their unobstructive nature, meaning they can easily and comfortably be worn in daily life while passively collecting high quality physiological data [9-11]. The collected physiological data can be used to build prediction models which help to identify stress occurrences in a timely manner, allowing for stress mitigation measures to be implemented [12].

Initial work on the project focused on detection of dementia-related difficulties using physiological data collected by wearable devices. Reviews were performed to identify the best wearable device for collecting data from people with dementia, as well as the best methods to use in experiments to collect that data. However, it was concluded that performing a data collection experiment with people with dementia during the COVID-19 pandemic was impractical and potentially unethical [13, 14]. As such, stress identification was identified as a new domain for the project to pivot to. This pivot and the justifications for it are explained in section 1.2. The aim of the project is to compare the work stress experienced in home and office work environment.

## 1.2. Pivot and justification

Section 1.1 of this thesis stated that the initial work conducted on this project was focused on researching using wearable devices and machine learning to support people with dementia. However, this work was halted due to constraints imposed by the COVID-19 pandemic and related lockdowns (this is discussed in greater detail in chapter 4 of this thesis) [13, 14]. As such it was decided to pivot the research away from using wearables and machine learning to detect and predict dementia-related difficulties and instead to use wearables and machine learning to detect and predict work stress. It was also decided that a comparison would be made between stress experienced in the home and office-based work environments, as this was very topical at the time the idea for the pivot was being formulated, due to work from home guidance being in force in numerous jurisdictions, and many employees working in a mixture of home and office work environments [15].

There are three main reasons for selecting stress detection as the new domain for research on the project. The first reason is that dementia-related difficulties and stress can both be detected using similar physiological attributes [9, 10]. This is because both are examples of negative affect, which trigger responses from the Autonomic Nervous System (ANS), which is the part of the nervous system which regulates involuntary or unconscious processes [16]. Indicators of ANS activations include heart rate, electrodermal activity, skin temperature, and movement, all of which have been used in literature to detect both dementia-related difficulties and stress [10, 17-19]. Thus, an understanding of the physiological attributes and features which can be used to predict and identify dementia-related difficulties is highly beneficial to understanding the physiological attributes and features which can be used to predict and identify stress, and vice versa.

Another reason for the selection of stress as a new area of focus for the project is its links to dementia, namely: it can be a risk factor of developing the disease; it can be a cause of dementia-related difficulties; and it can be a symptom of dementia and related difficulties [20-23]. Stress management has been identified in previous literature as an important aspect of day-to-day management of dementia [22]. Dementia-related difficulties are often stressful experiences for an individual to contend with, for example, it can be very stressful becoming disorientated as to where you are and forgetting what time or day it is [24]. Furthermore, when individuals with dementia are exposed to stressors, their cognitive impairment can make it difficult for them to address and mitigate the stressor, or to effectively express to others that the stressor is being experienced. This can then lead to the dementia-related difficulty becoming worse, as the individual experiences greater negative affect and more intense negative emotions due to the frustration of being unable to effectively express their distress [25]. As such, detecting and mitigating stress experienced by the individual with dementia in a quick and timely manner can reduce the occurrence and potential intensity of the dementia-related difficulties. Furthermore, there is a link between work stress and dementia, with long-term exposure to occupational stress being linked to an increased risk of developing dementia [26-28]. Thus, the timely detection and mitigation of work stress could potentially reduce the risk of an individual developing dementia in later life, further linking the domains of dementia and stress research.

Moreover, the methods used to collect and analyse the physiological data in both domains are often very similar, with similar devices [10, 11], data collection methods and protocols, data storage methods, and data analyse methods being employed in both domains. This means that lessons learned from performing physiological data collection with generally healthier, working adults will increase researcher knowledge of how to safely conduct experiments with people with dementia, meaning that future research can be conducted more safely- especially regarding COVID-19- and efficiently, with less burden upon the participants and caregivers [10, 11]. The implications of conducting such a collection experiment with people with dementia during a pandemic, such as COVID-19, are explored in depth in chapter 4.

Overall, the pivot of the research project to the domain of work stress is justified by: the similarities in the physiological expression of stress and dementia-related difficulties; the links between stress and dementia, namely stress causing dementia-related difficulties, or vice versa, and long-term exposure to stress being a risk factor for developing dementia; and the similarities in the methods used for research in the dementia-related difficulty and stress detection domain.

## 1.3. Problem Statement

Both home and office work environments present unique stressors and stress-reduction benefits, especially during the COVID-19 pandemic. Though work has been conducted to understand the stressors and individual stress levels in each of these work environments during this time, there remains a number of gaps in the knowledge. The majority of work in this domain are based on subjective questionnaires, which often suffer from bias from the participants and are limited in their predictive insight compared to physiological data. Furthermore, the literature that exists describing the use of physiological data to analyse difference in academic work at home and in an office environment has no published results, meaning that there is still a lack in understanding of how the physiological attributes relate to their stress levels in each environment.

## 1.4. Research Questions

There were 4 research questions related to the topic of work-related stress which are outlined in this section. However, the initial research conducted on detecting dementia-related activities aimed to address 3 different research questions. The work done to answer these initial 3 research questions was vital in informing the work done to answer the research questions post-pivot, and so the inclusion of the initial research questions in this thesis is important for the reader to gain a proper understanding of the research project that was undertaken. As such, the initial research questions have been added as sub-questions to the finalised research question they were deemed most relevant to (further explanation and justification of this is provided later in this subsection).

The research questions are:

**RQ 1:** What differences exist in stress levels between instances of home and office working?

> **RQ 1.1:** What are the best physiological attributes and sensors for predicting the occurrence of dementia-related difficulties?
>
> **RQ 1.2:** What is the best wearable device which can be used to track the indicative physiological attributes of people with dementia, in a comfortable, unobtrusive and unobstructive manner?

**RQ 2:** What, if any, correlation exists between subjective stress, subjective productivity, time of day, day of week, physiological features, and work environment?

**RQ 3:** Can machine learning be used to identify and predict occurrences of stress in both work environments, based on the physiological attributes of the individual, and what is the best model for doing so?

> **RQ 3.1:** What is the best machine learning model for predicting the occurrences of dementia-related difficulties and the context in which they occur?

**RQ 7:** What mitigations can be used to reduce or prevent work stress once it is detected, based on the causal environmental and personal factors identified in the answer to question 2.

Research question 1.1. and 1.2. in the list above were research question 1 and research question 2 in the initial research on detecting dementia-related difficulties. These research questions were answered in the work that constitutes chapter 3 of this thesis. These questions were consequently added as sub-questions of research question 1 in the new research questions following the pivot. The reason for this is that the answers that were found for questions 1.1. and 1.2. were extremely useful in designing and conducting the research which ultimately helped to answer research question 1. For example, the answer to research question 1.1. helped to inform which physiological attributes and features should be used to measure and predict stress so that it can be compared between the home and office work environments, and the answer to research question 1.2. helped to inform which wearable device should be used to collect those physiological attributes and features.

Alternatively, research question 3.1. was added as a sub-question to research question 3 as the answer to research question 3 will likely support the research into answering research question 3.1. As stated in subsection 1.2, the physiological attributes and features, as well as the data analysis methods commonly used in the dementia-related difficulty detection and stress detection are very similar. As such, one could reasonably suspect that a machine learning model which can predict stress with a high accuracy could also do the same in predicting dementia-related difficulty (obviously with some tweaks and alterations to tailor it more to the dementia-related difficulty detection domain). Thus, answering question 3 will provide a good baseline of knowledge of relevant machine learning models and the features which can be used to train them to answer question 3.1.

## 1.5. Research Scope

The purpose of this research project is to understand the differences in subjective and physiological stress levels in individuals and groups when they are working from home as opposed to in an office environment, and vice versa. A wearable device (Empatica E4 smartwatch) will be used to track the physiological data of a population of adult academics and researchers at Liverpool John Moores University while they work from hybrid environments (i.e., work and office environments). Participants will also record their subjective stress and productivity levels on an online questionnaire. The study shall be a rolling study, with each participant recording their work activities, stress, productivity, and physiological data for 10 working days, then another participant being given the device to begin the study. Data will then be analysed using MATLAB, with descriptive statistical features being extracted from the physiological and questionnaire data, which will be used to establish if relationships exist between stress, productivity, and the physiological data features. Furthermore, a variety of machine learning algorithms will be trained to establish if machine learning can be used to identify and predict occurrences of stress in each work environment and to identify and predict occurrences of stress and the work environment of individuals and groups from their physiological data.

## 1.6. Research Aims & Objectives

The main research aim of the project is to compare the work stress experienced in home and office work environment. The objectives outlined in this section also include the objective of the project from the initial stages when the aim was to develop a wearables-based method for identifying and predicting the occurrence of dementia-related difficulties.

1. To develop a wearables-based method for identifying and predicting the occurrence of dementia-related difficulties, following objectives are set:
    a. To identify the best physiological attributes and sensors for predicting the occurrence of dementia-related difficulties.
    b. To identify best wearable device which can be used to track the indicative physiological attributes of people with dementia, in a comfortable, unobtrusive and unobstructive manner
    c. To collect or access a dataset containing the physiological indicators of dementia-related difficulties.
    d. Train a machine learning model for identifying and predicting the occurrence of dementia-related difficulties.
2. To compare the stress experienced by participants in home and office work environments, following objectives are set:
    a. To collect a dataset containing physiological indicators of work stress in home and office environments.
    b. To conduct a data analysis to understand the complex relationships between the collected attributes and extracted features.
3. To develop a method for identifying work stress and suggest mitigations for managing and reducing work stress once identified, following objectives are set:
    a. Train a machine learning model which can predict the occurrences of work stress from physiological data features.
    b. To identify potential mitigation for managing and reducing work stress, based on the implied causes and context of the stress identified as part of objective 2a.

## 1.7. Research Contribution

This research makes 3 main contributions to the field, which are as follows:

1. A novel dataset containing questionnaire attributes- such as subjective stress and productivity- and physiological attributes- such as heart rate, electrodermal activity, skin temperature, and movement, related to work stress. This dataset has a wider, more comprehensive range of physiological attributes than many already available [29], and has higher temporal resolution than other datasets in the literature [9, 11].

2. A novel bagged tree ensemble machine learning model which can predict inter-class work stress with a validation accuracy of 99.3%, higher than the accuracy of comparable models existing in the literature, and a test accuracy of 73.7%, which is relatively high for the limited size of the dataset it was trained on, and compared to the 62% achieved by a random forest model trained on data collected in similar work environments [29].
3. A novel comparison of the subjective and physiological stress levels and indicators in home and office-based work environments, expanding upon the work of Bolliger et al. by collecting and analysing data in the home environment and including more specific activities related to teaching [9, 11].

## 1.8. Published work

In this section, the work published during the conduction of the research project shall be detailed and summarised, as this published work forms the basis of the research project and the results thereof. The publications are split into 2 categories. The first category is the publications related to dementia and dementia-related difficulties, which were published prior to the switch of focus on the project to work stress detection. The second category is those relating to work stress and its detection, published prior to the aforementioned pivot of the project, with the first paper described being initially dementia-related but justifying the pivot to work stress-related research in its later parts.

### 1.8.1. Dementia related publications

In Harper et al. (2019), the ways in which data science can help with prediction, diagnosis and treatment of Alzheimer's disease are reviewed [22]. It is concluded that current data science techniques are useful in aiding the successful management and treatment of Alzheimer's patients in day-to-day life, and though there is much promise to the use of data science techniques to predict and diagnose Alzheimer's disease, no technique yet exists that is able to process all the required data. It is concluded that research should be conducted to develop a data science-based system to help in Alzheimer's prediction and diagnosis. Also, further research should be conducted to improve current data science techniques used to support the treatment and management of Alzheimer's disease, as was the initial aim of this project. Wearable computing devices are identified as an area of particular promise in this regard, with physiological data collected by such devices being useful in predicting and identifying occurrences of several dementia-related difficulties. Continuing on this theme, Harper & Ghali (2020) presents a systematic review of wearable devices used in the prediction or identification of dementia-related agitation [30]. Six relevant devices found in the literature were evaluated to identify their strengths and weaknesses, with all possessing at least one weakness which made it not ideal for future research on the prediction or detection of dementia-related agitation. Two of the wearables utilised the three desired sensing modalities, identified in previous research. One of those devices was identified as prohibitively expensive to be used in a wide array of applications and research studies, and the other device was untested and unavailable to purchase or use. The conclusion was that work in future studies should focus on developing an inexpensive device which utilises the three desired sensing modalities while remaining usable and accessible for use in a wide array of studies and applications. From this paper, the Empatica E4 was chosen as the device that was most suitable for use on this project, as the device collected all the physiological data streams which were found to be useful, as well as having demonstrated usability and comfort for people with dementia.

Harper et al. (2021, a) outlines a methodology for performing a systematic search for a physiological dataset which could be used to train a machine learning model to predict the occurrences of dementia-related difficulties [31]. The results of the search were that no relevant dataset was found to be publicly available to researchers, either freely or upon request. In response to this finding, a methodology for collecting a relevant physiological dataset for training the fore mentioned model and making the dataset publicly available is proposed. Moreover, methods for using the dataset to train classification models that can predict difficulties are also discussed. Three main categories of solutions to overcome the lack of available data are identified. The first is the most important, namely conducting a data collection experiment to collect a novel physiological dataset. Second, using anonymization and pseudonymization techniques to remove any and all identifiable data from the collected dataset, meaning that if the data is shared the identity and privacy of the participants is protected. The third category of solutions is utilising synthetic data generation to create a larger, anonymous training dataset. In conclusion, a combination of all the identified methods should ideally be employed in future solutions. Future work should focus on conducting a data collection experiment to collect a dataset which could be used to train a dementia-related difficulty prediction model. Furthermore, work should aim to ensure that the collected data can be publicly available to other researchers, reducing the cost and time investment required to research this domain.

### 1.8.2. Stress related publications

In Harper et al. (2021, b) past physiological data collection experiments conducted with people with dementia were reviewed and the methods utilised in the various stages of those experiments were evaluated [13]. The impacts and limitations imposed by the COVID-19 pandemic and lockdowns on the efficacy, safety, and burden-imposition on experiment stakeholders of each method was also discussed. It was concluded that the decision regarding which of the methods to utilise in future data experiments depends mostly upon the type and severity of the dementia of the participants. Moreover, the choice of COVID-secure methods, such as remote studies, are preferable during the COVID-19 pandemic, however such increased costs and burdens were considered to make conducting such an experiment impractical in many cases during the pandemic. Furthermore, Harper & Ghali (2021, c) focused in greater depth upon the burden imposed upon dementia caregivers, both formal and informal, by data collection experiments [14]. Literature showed that caregivers were already highly susceptible to burden and negative psychological and physical health outcomes due to their care duties. This burden has increased in many cases during the COVID-19 pandemic due to increased fears of viral transmission to people with dementia, as well as the extra burden of COVID-secure measures now required in many instances during their caring duties. Similar to the conclusion in Harper et al. (2021, b), it was found that these burdens on caregivers made the conductance of a data collection experiment with people with dementia impractical during the pandemic. As such, it was proposed that the project should pivot towards stress detection. This pivot was justified with the following reasons: stress is often the trigger of dementia-related difficulties, meaning the developed stress prediction model could provide transferable knowledge for the dementia-related difficulties detection; lessons gained from performing physiological data collection with healthy adults will increase researcher knowledge of how to safely conduct dementia-related experiments in the COVID-19 era; stress management techniques are useful tools in the treatment and management of dementia [14]. As such, an experiment is proposed to identify differences in stress levels of individuals working from home and from office environments during the latter stages of the COVID-19 pandemic in the United Kingdom.

However, prior to the initiation of the data collection experiment describe in the following paper, Harper et al. (2022), the COVID-19 pandemic related restrictions were lifted in the United Kingdom, meaning the focus of the fore mentioned paper was less on COVID-19 than expected. Instead, the paper presented an intelligent approach to predict the stress occurrences using the physiological data acquired from individuals working in both remote and office locations. Multiple factors were collected related to physiological indicators of stress and subjective performance level. We developed a boosted tree ensemble model which produced binary stress classification accuracy of 99.9%. The statistical outcomes indicate that there is no overall correlation between mental stress and productivity, however there is some indication of mental stress being is influenced by the work environment, the time of day and the day of the week [15].

## 1.9. Thesis Structure

The thesis is structured as follows. Chapter 2 provides an overview of the relevant literature and the state of the art for work on detecting dementia-related difficulties and work stress using wearable devices, while also providing background information on the causes and symptoms of dementia and related difficulties, as well as some non-wearable related research on the causes and implications of work stress. Chapter 3 explains the methodologies and results for a data search attempting to find a relevant dataset for the project- when the focus was still detecting dementia-related difficulties- and then a subsequent literature search to identify the best wearable device to collect such a relevant dataset and physiological attributes. This chapter thus aims to fulfil objectives 1a and 1b of the research project, as outlined in section 1.5. of this thesis. Chapter 4 focuses on developing a protocol for collecting a novel dataset from people with dementia which can be used to develop the fore mentioned difficulty prediction system, first by showing the methodology and results for a systematic review of methods used in similar studies in the literature. Then the results of this review are used to design a novel methodology to conduct the required data collection experiment. The chapter then explains the problems with conducting such an experiment during a pandemic, before justifying why a pivot of the project to work stress prediction is warranted. Thus, this chapter aims to go some way to addressing objectives 1c. and 1d., however as shall be described in greater detail in later areas of this thesis, these objectives were not fully met due to changing circumstances. Chapter 5 outlines a data collection methodology used for collecting and analysing a physiological dataset related to work stress, with the analysis including the training of a number of machine learning models. The work in this chapter thus addresses objective 2a of the research project. The chapter then discusses the results of the analysis and explains how they answer the posited research questions 4-7, also addressing objectives 2b and 3a. in the process. Moreover objective 3 is addressed in this section by a discussion of potential mitigation techniques for managing and reducing work stress, based on the implied causes and context of the stress identified as part of objective 2a. Finally, chapter 6 concludes the thesis by summarising the answers to the research questions and detailing the future work which should be conducted on this project and in the work stress and dementia-related difficulties fields more generally.

# 2. Literature Review

## 2.1. Introduction

The management of dementia and stress are both highly important, and as such, there is a large body of literature detailing research on using wearable-based systems for the management of both potentially negative states. In this chapter, the state of the art in fields relevant to this project is presented, through a review of the existing literature. Section 2.2. explains what dementia is, the types of dementia, the symptoms of each type of dementia, and the difficulties in daily activities and lives which could be experienced due to dementia and its symptoms. Section 2.3. provides an overview of the wearable technology-based systems existing in literature to identify and predict the occurrences of dementia-related difficulties. Section 2.4. then reviews past work on the differences in stress levels experienced working in office and home environments. Section 2.5. gives a view of the state of the art in research using wearable computing devices to detect work stress, and finally, section 2.6. summarises the chapter.

## 2.2. Dementia: Types, symptoms, and related difficulties

Dementia is an umbrella term used to describe a range of neurodegenerative conditions associated with progressive death of neurons [22]. The death of neurons leads to cognitive impairment and a reduction in cognitive and motor skills, causing the individual to experience a number of symptoms and difficulties in their daily lives [32]. The most common condition referred to as a form of dementia in older adults is Alzheimer's disease (AD), with approximately three quarters of dementia cases being Alzheimer's disease [33]. Other conditions in the dementia umbrella include Lewy Body Dementia, Frontotemporal Disorders and Vascular Dementia [34]. The exhibition and variety of symptoms depend upon the type of dementia the individual develops, with it being common that individuals have multiple of the conditions at the same time [22]. Symptoms and progression of the disease can also vary greatly between individuals, even individuals with the same conditions, based on a myriad of other factors. However, generally speaking, the symptoms of each type of dementia progress from mild, to moderate, to severe, with death being the inevitable end of the progression, as the brain becomes unable to function and maintain vital bodily functions [35]. Despite the inter-conditional and interpersonal variety of symptoms, there is generally agreed upon symptoms that are common at each stage of each condition.

Alzheimer's disease is the most common form or cause of dementia in older adults. It was first formally identified by Alois Alzheimer in 1907 and is thought to be caused by abnormal deposits of two specific proteins in the brain [22]. One of the proteins is Amyloid, with abnormally large deposits of this protein causing plaques to form around neurons. The Tau protein is the other, which causes the formation of tangles in the neurons. The resulting damage which occurs to the brain cells is that there is a decrease in levels of neurotransmitters, making the transmission of signals between neurons more difficult [36]. As the disease progresses the damage leads to areas of the brain shrinking, with the first area to be affected being responsible for memories, in most cases [37]. The symptoms of mild Alzheimer's disease include small lapses in short-term memory (such as misplacing items or forgetting why one entered a room), decreased ability to plan and organise, and difficulty performing some cognitive tasks in social or work settings (for example remembering names or performing sums) [38]. The symptoms of moderate Alzheimer's disease tend to be more noticeable to the individual and others than those of the mild condition, with reductions in short term memory and reduced ability to perform complex tasks being more pronounced [39]. For example, the individual may begin forgetting recent events and personal details with greater frequency. Mood swings and personality changes could also begin to be exhibited at this stage, along with social withdrawal. Another common symptom at this stage of the disease is disorientation in regard to location and time, which can mean the individual gets lost and wanders more often or can experience distress and anger due to their confusion or inability to know where they are or to express their needs or feelings adequately. The severe stage of the condition is the final and most impairing stage. The individual will require assistance for most daily tasks, experiencing symptoms such as reduced awareness of surroundings, problems recognising individuals beyond being familiar or unfamiliar, repetitive behaviour, wandering and suspicion. Changes in personality and mood swings will also become far more noticeable and frequent. As the severe condition progresses, the ability to respond to the environment will eventually cease to exist and communication will become extremely limited, if not impossible. Constant care will be required as basic motor functions, such as the ability to swallow, shut down due to the sheer extent of damage to neurons. Death of the individual will occur as the final neurons perish [40].

Vascular dementia is the second most common dementia, accounting for around a fifth of all dementia cases [41]. It is caused by reduced blood flow to cerebral neurons, causing them to become damaged and perish. Subcortical Vascular dementia occurs due to the narrowing of small blood vessels within the brain, which restricts the volume of blood delivered to neurons [42]. Post-stroke or single-infarct dementia occurs after the individual has suffered a stroke. Multi-infarct dementia occurs when an individual suffers multiple mini strokes which cause damage to several areas of the brain [43]. Symptoms of mild vascular dementia may include slowness of thought, trouble understanding concepts and concentrating, and problems with memory. In the later stages of vascular dementia symptoms may include severe personality changes and aggression, significant memory loss, disorientation and confusion, depression and mood swings, incontinence and difficulty walking and remaining balanced [44].

Lewy Body Dementia is either the second or third most common dementia, accounting for approximately a tenth to a fifth of all dementia cases [45]. It is caused by the build-up of Lewy Bodies, which are tiny deposits of proteins [46]. The exact cause of the formation of Lewy Bodies is not yet known nor is how they interact with the brain to cause the symptoms; however, they cause a reduction in levels of neurotransmitters and the eventual death of cerebral neurons. Damage caused by Lewy Bodies can also be the cause of Parkinson's disease and a number of other disorders affecting the brain. As such, Lewy Body dementia shares symptoms with Parkinson's disease, and individuals with either disease suffer from motor function impairments such as trembling, difficulty balancing, and slow, stiff, rigid movement, among other motor function impairments. Lewy Body dementia also causes cognitive symptoms such as fluctuating alertness and periods of confusion or sleepiness, disturbed sleep patterns, fainting spells, and hallucinations. Once again, the symptoms intensify and get worse as the disease progresses, with the patient eventually losing all ability to look after themselves and be independent in any way [47].

Frontotemporal dementia accounts for 15% of dementia cases and, as the name implies, mainly affects the frontal and temporal lobes of the brain [48]. The resulting degenerations in these regions can be caused by a number of diseases, with the main two causes being disorders which involve abnormal levels and activities of the Tau or TDP43 proteins [49]. There are several differing types of the disease, all of which have different symptoms at certain stages of progression. Behaviour variants of the disease are characterised by changes in personality and conduct, as areas in the brain that control conduct, judgement and empathy experience the most prominent neuron loss [50]. Primary progressive aphasia derives from neuronal loss in areas of the brain that govern language, speaking, writing and comprehension, meaning an individual will experience problems understanding written or spoken words or experience great difficulties communicating as speech becomes hesitant, laboured and ungrammatical [51]. Another type of Frontotemporal dementia affects motor skills, but not necessarily behaviour or language. One disorder included in this form of dementia is Amyotrophic Lateral Sclerosis, which has symptoms including muscle weakening and wasting [52]. Another is  Corticobasal Syndrome, which can cause an inability to coordinate one's limbs and stiffness in limbs [53]. Finally, Progressive Supranuclear Palsy is also a form of frontotemporal dementia, with symptoms including muscles stiffness, difficulty walking, changes in posture, and affected eye movement [54].

## 2.3. Wearables for dementia management

Wearable computers are used extensively in literature to track physiological indicators of specific dementia-related difficulties or groups of difficulties [10, 18, 55, 56]. In Harper et al. (2019) and Harper et al. (2020) it is argued that wearable computing devices can be some of the most effective and useful tools for aiding in the management of dementia and related difficulties [22, 30]. Advantages of using systems which incorporate wearable devices include the passive, non-obstructive, comfortable, and convenient manner in which they can collect physiological data. In this section, an overview of the aforementioned literature regarding wearable-based systems for aiding the management of dementia and related difficulties is provided.

The BESI system is one system which supports the detection of agitation by analysing the behaviour and movement of an individual with dementia. The BESI system utilizes the Pebble smartwatch, an accelerometer, which is able to track participants' movements. From the accelerometer sensor data, the system aims to identify when an individual participant exhibits agitated behaviours [57-59]. In the related study, participants wore the Pebble for 30 days. People with dementia and their caregivers were recruited as pairs, with the number of pairs included in each iteration of the study varying between 3 and 10. Machine learning classification models, such as support vector machines (SVM) and a bagged tree ensemble model, where then trained to classify instances of agitation, with an ensemble model achieving the most robust classification of agitated behaviours [59]. Available literature does not provide enlightenment as to the participating individuals' stages of dementia; however, one might assume that from the community-based setting in which participants lived with a not insignificant degree of independence, that mild to moderate dementia is implied. A wrist-worn device was also utilised by Melander et al. (2017) to attain physiological data from individuals with a condition within the dementia umbrella living within institutionalized settings [18]. The researchers requested of several participants that they wear an Empatica E4, which tracked the electrodermal activity (EDA) of the individual donning the device, and concurrently a nurse was tasked with the recording and notation of observations of occurrences of dementia-related agitation and other relevant difficulties on a chart provided to them by the researchers. Researchers found a strong, significant correlation between the physiological data recorded at the time the difficulty occurred and the difficulty. Also, there was a strong, significant correlation between the observed difficulties and the physiological data 1 to 2 hours prior to the time the observation was noted. Thus, EDA data is shown as highly favourable for predicting the occurrence of dementia-related difficulties, even up to 2 hours before the difficulties are observed.

Sefcik et al. (2019) used a chest-worn ECG sensor to monitor the heart rate of people with advanced dementia who had the symptom of persistent vocalizations (PV) [60]. PV are uncontrolled or disruptive vocalizations with no specific purpose regarding communication. Participants wore the device for 2 hours each data collection session and caregivers recorded each time a PV occurred. Heart rates of each participant were compared on days PV were exhibited and days no PV were exhibited. A strong correlation between heart rate and PV was found, with heart rates being generally higher on days PVs were exhibited. Nesbitt et al. also utilised heart rate to identify and predict dementia-related agitation, however they used it in combination with limb movement and vocalisations recorded by a microphone [61]. The data collected by the smartwatch correlated with many of the observations of agitation, whereas the vocalisation data did not correlate with the agitation. The researchers suggest that the reason for this may be due to background noise making the recordings too noisy.

In conclusion, wearable computing devices are demonstrated to be highly useful in the tracking and detection of specific dementia-related difficulties, such as PV and agitation [62-64]. However, there is no system which can detect a truly comprehensive range of dementia-related difficulties, such as wandering, agitation, and PV combined. This is a problem, because as was discussed in section 2.2. of the thesis, people with dementia may exhibit a range of different and changing difficulties depending on the type of dementia they have and the speed of their disease progression. Thus, a system which tracks just one difficulty or group of difficulties may miss some of the difficulties in the user's day to day life or may become obsolete for an individual once the disease progresses and their symptoms change or worsen. All this considered, the development of a system which can detect and support the management of a comprehensive range of dementia-related difficulties and symptoms is vital in future work.

## 2.4. Stress in home and office working environments

Luis-Martinez et al. (2021) performed a descriptive-cross-sectional study to understand the link between physical activity and stress in members of the academic community working remotely from home during COVID-19 pandemic-related lockdowns. The researchers asked second and third year university students to complete surveys via the Moodle Platform, with the Cognitivist Systemic Inventory being applied to study academic stress (SISCO) and the International Physical Activity Questionnaire (IPAQ) being used to assess physical activity. The results show that 93.4% of respondents self-reported experiencing stress of some level during the study, with 15.5% self-reporting mild stress, 65.4% self-reporting moderate stress, and 2.5% self-reporting severe stress. The study found that the main stressors experienced by the respondents were related to the demands of self-directed study and activities complementary to this style of learning, while regular physical activity was shown to be a useful stress reduction technique [65].

Similarly, Marcén-Román et al. (2021) conducted a cross-sectional descriptive study at the University of Zaragoza with a sample of 252 Health Sciences university students, to assess the psychological impact of pandemic-related stress on the students. Participants were asked to complete a self-administered online questionnaire which used a modified scale (PSS-10-C) and the Goldberg scale to evaluate the impact of perceived stress and assess anxiety and depression, respectively. Like Luis-Martinez et al. (2021), the results show an increase in student stress due to the pandemic, however Marcén-Román et al. (2021) found that only 13.1% of students from across the 3 health sciences degrees self-reported stress, considerably less than reported by Luis-Martinez et al. (2021). The results also showed lower self-reported stress than Odriozola-González et al. (2020), who performed a similar study at an earlier stage of the pandemic and found 28.14% of respondents reported stress [66]. The marked reduction in perceived stress between the results Odriozola-González et al. (2020) and Marcén-Román et al. (2021) may indicate that an increased knowledge of the disease and developed methods of increasing remote learning experiences may result in a lowering of perceived stress in the students as the pandemic progresses [67].

Fornili et al. (2021) used online surveys to address pandemic-related lockdowns and the resulting changes in lifestyle of the academic communities in five Italian universities. 220,000 students and 20,000 employees were invited to participate, with 6% of students and 19% of employees invited ultimately participating. The online survey had a number of sections, asking the participants for information relating to their socio-demographic, housing, and habitual characteristics, as well as any symptoms or lifestyle changes due to COVID-19. Furthermore, the researchers assessed psychological distress of the participants using the Hospital Anxiety and Depression Scale (HADS). 31% of respondents reported suffering anxiety due to the pandemic, with 16% reporting borderline anxiety and 15% reporting severe anxiety. Overall, results showed that 20% of participants were either severely anxious or severely depressed. People of younger age and lower socio-economic status were found to have higher levels of socio-economic stress, as well as females and students. Similar to the results of Luis-Martinez et al. (2021), physical activity was found to be a useful stress management technique, with regular physical activity correlating negatively with HADS scores [65, 68].

Salazar et al. (2021) also used questionnaires to assess the levels of stress in the academic community caused by COVID-19 three and a half weeks into a pandemic-related lockdown. They collected sociodemographic data from 677 participants, as well as information on the participants' coping strategies, level of anxiety, stress, depression, perception of COVID-19, and perceived level of social support. The results showed that most participants (80.5%) had the same health status at the time of the study as before the lockdown. However, in a number of demographics there were significant increases in stress, above those experienced by other demographics. As with Fornili et al. (2021) females and younger individuals were shown to have higher stress levels than other demographics, with researchers suggesting this may be due to the more affected demographics having greater sensitivity to the interruption of interpersonal relationships and worries about future career prospects. However, one limitation of this study is that no data on participants' mental health history was collected, thus it is not possible to say the elevated stress levels of some individuals was not in some way influenced by pre-existing mental health conditions [68, 69].

However, outside of academia the correlation between the COVID-19 pandemic/lockdowns and increased stress is not as clearly demonstrated in the literature. Shao et al. (2021) surveyed 127 employees of a software development company, in order to identify the stressors, they experienced working remotely and from the office during the gradual transition from full-time home working during lockdown back to office working. The participants completed 2 daily surveys, recording their work location and the stressors they experienced in their work location, selected from a list of 28 stressor types, pre-determined by the researchers based on a pilot interview study. Using this information, the researchers aimed to predict the stressors most commonly experienced in each location and the participants choice of work environment for the subsequent day following the date the data was collected. The results showed that the greatest predictors of next day work location were interferences from family, difficult work coordination and excessive workload, with the former two being indicative of choosing to work from the office the next day and the former being a predictor of working from home the next day. Moreover, researchers found that though stressors relating to COVID-19 infection are not enough alone to predict the employee's choice of work environment the next day their effects become clearly apparent in tandem with other stressors. Overall, this study found that some stressors, such as family boundary stressors and technology stressors, were more prevalent in home working environments and some stressors, such as excessive workload, were more prevalent in office environments, meaning that the stress levels of the individual may not necessarily be higher in one work environment than another, with all factors considered [70].

Similarly, Galanti et al (2021) found that working from home during the COVID-19 pandemic increased the prevalence of certain stressors experienced by participants, however reduced the prevalence of other stressors. Their cross-sectional study analysed data collected through an online questionnaire completed by 209 employees working from home during the pandemic. The researchers then tested their hypotheses using hierarchical linear regression. As shown by Shoa et al. (2021), stressors relating to social isolation and family-work conflict increase work stress in work from home scenarios. However, working from home also decreased stress resulting in those emotionally affected by COVID-19, including those with concerns of contracting the disease or passing it on to loved ones [71]. As such, the results from both studies could be used to argue that the level of stress experienced working from home during COVID-19 is reliant upon the individual beliefs, preferences, and demographics of the individual participants. This notion is supported by much of the literature covering stress in the academic community during the pandemic [66-69].

Furthermore, Tan et al. (2020) found that Chinese workers returning to working in the workplace after lockdown generally exhibited no significant increase in the level of psychiatric symptoms, such as stress, contrary to the researchers' expectations. Participants completed an online questionnaire which collected data regarding their: demographic and occupational status; physical symptoms and self-reported physical health for the past 14 days; perceived impact of COVID-19 lockdown and related events through the Impact of Event Scale-Revised (IES-R); experiences of depression, anxiety and stress using the Depression, Anxiety and Stress Scale (DASS-21); their sleep quality using The insomnia Severity Index (ISI); other psychiatric symptoms; and the COVID-security and hygiene measures enforced in the workplace. Results showed a prevalence of just 1.5% of stress within the population related to the return to working from an office environment. One possible explanation for this unexpectedly low stress prevalence could be the rigorous COVID-security measures, such as increased social distancing, handwashing and mask wearing that were compulsory as employees returned to work [72].

Moreover, Irawanto et al (2021) found a significant increase in stress during the preliminary stages of lockdown forcing employees to work from home, however this stress lessened as the lockdown progressed. A questionnaire was used to collect 41 features relating to 472 participants' work from home status, work-life balance, work stress and job satisfaction. The data was then analysed using an analysis of the significance of the path coefficient of partial least squares (PLS) to test the hypotheses, all relating to the impact working from home, stress, and work-life balance had upon each other and job satisfaction. During the preliminary stages of the enforced work from home period, employees generally experienced an increase in stress due to the additional demands of adjusting to a new working environment and new working routines and methods. However, increased time spent with family appears to have offset much of this stress as the work from home period continued. Researchers also found that job stress experienced working from home would have a negative impact on job satisfaction [73].

In addition, Moretti et al. (2020) found that there was no notable change in stress experienced by individuals working from home due to lockdowns from when they worked in offices. Participants in their study were administrative officers working remotely due to COVID-19. Questionnaires containing 12 items including demographic information, past experiences of remote working and perceived productivity and work stress were filled in by participants. Researchers further asked if the participants would continue working remotely after COVID-19, given the option. The researchers then used SPSS v. 25.0 software to calculate the descriptive statistical features of the collected data, as well as the correlations between each of the features. Since the beginning of the work from home period and the study, 27.5% of respondents reported experiencing no change in their levels of work-related stress, 39.2% reporting a reduction in work-related stress and 33% reported increased work-related stress. As such, reduced stress is the most common outcome of working from home instead of an office environment in this study, however it is also true a majority of workers did not experience a reduction in stress working at home rather than in the office. Furthermore, participants self-reported lower productivity working at home than in the office, which could mean home working is less conducive to effective and efficient working. Despite this, 62.7% wanted to continue working remotely post-pandemic, implying that for reasons other than stress-though possibly including it in some cases- remote working was preferably to a majority of the surveyed workers [74].

Likewise, Prasad et al. (2020) found that working from home had both benefits and drawbacks for employees, with stressors being different in work and office environments.

Using a questionnaire, the effect of seven occupational stressors on the dependent factor of psychological well-being of employees in the Information Technology Industry were measured. Participants recorded remote working factors and stress on a Likert scale (1-5), and psychological well-being factors measure on a scale of 1-7. Results of the analysis showed remote working is a challenge for employees because of workplace isolation, family disturbance, peer absence, lack of suggestions to the employees, and working too much or not working at all. Alternatively, benefits of remote working were found to include reduced time and stress commuting, and increased job control. These findings of a mixture of positive and negative stress-related factors are consistent with other studies in the field [73, 74]. However, unlike other studies of individuals working or studying from home during COVID-19, no statistically significant differences in stress were identified in different demographic groups [66-69].

Song & Gao (2020) compared the stress experienced by workers working remotely and in the office, using data from the 2010, 2012, and 2013 American Time Use Survey Well-Being Modules, which was collected through phone interviews by the U.S. Census Bureau. Statistical comparisons of correlations were calculated between stress and different work environments, and work and non-work activities. Analysis showed that remote working was associated with higher levels of stress than working in the office, and thus has a negative effect on subjective well-being. However, many of the differences in stress experienced working in different environments may be highly influenced by the participants' demographic status. Parents experience higher stress and lower happiness working remotely, while childless workers experience fewer negative effects when working remotely. The researchers conclude that there is no conclusive evidence of the benefits of remote working to employee's subjective well-being, and that demographics are important considerations when deciding whether individuals need extra support working remotely [75].

Similar to Song & Gao (2020), a study by Messenger et al. (2017) concluded that working remotely generally resulted in increased experience of stress, due to reduced work-life barrier, often longer hours of work and family conflict. Eurofound and the ILO jointly developed a standard expert questionnaire to compile the information on remote telecommunications working across 15 countries. This data was then analysis for trends and statistical features regarding remote workers and the effect of remote working on workers, companies, and society. It was found that positive features of remote work include a general increase in work-life balance, reduced stress and time commuting, and increased autonomy to organise one's own work and work environment. However, 41% of home workers felt stressed compared with 25% of those working in an office, implying remote working can result in an increase in stress for many workers [76].

## 2.5. Wearables to detect work stress

Bolliger et al (2020) proposed using a wearable computing device- the Empatica E4 smartwatch- Ecological Momentary Assessment (EMA) in the form of questionnaires on smartphones, and smartphone usage data to predict occurrences in stress in individuals in the academic community. The researchers aimed to recruit 50 participants, asking each to wear an Empatica E4 device as they complete their daily work activities, to collect their heart rate, electrodermal activity (EDA), skin temperature, and movement data. The participants were also asked to download an app to track their smartphone usage, and to answer a number of EMAs on their smartphone. The EMAs were used to collect ground truth data, such as the activity the participant was completing and their subjective stress level. Researchers aim to combine the data and use it to train a stress prediction model. One limitation of their protocol is that the data collection coincided with the COVID-19 pandemic, however the researchers only collected data on days when the participant was in the office, despite work from home guidance and policies from government and businesses being ubiquitous. This inevitably led to long pauses in their data collection experiment. As such, in Bolliger et al. (2023) the authors updated their protocol to include data collection on days when the participant is working from home, however analysis of the effect of environment on stress levels is somewhat limited, with no analysis of the physiological stress data at all [11]. Furthermore, the study was limited by the high demand the protocol had on participants, with them having to regularly complete EMAs and otherwise adhere to the study protocol [9].

Betti et al. (2018) also explored the usability of wearables for detecting work-related stress. In their study, 15 participants performed a Maastricht Acute Stress Test (MAST) whilst wearing a collection of physiological sensors, which collected electromyography (EMG), Electroencephalogram (EEG), EDA, and heart rate variability (HRV) data. The participants were also asked to provide a saliva sample so researchers could analyse salivary cortisol levels as an objective comparison to their stress predictions from other physiological features. From the various physiological items collected, 15 descriptive statistical features were calculated. These features were then used to train a Support Vector Machine (SVM), which predicted the occurrence of stress or non-stress on based on participant data. Using the 15 features, an accuracy of up to of 86% was achieved for discriminating between stress and relaxation. Furthermore, the results of the data analysis also correlated with levels of salivary cortisol [77]. This is a relatively high accuracy, higher than the 74.5% accuracy in binary classification achieved by Wisjman et al. (2013) [78]. However, Han et al. (2017) achieved 94% accuracy in binary classification, using a combination of SVM and random forest (RF) classifiers on features extracted from ECG and respiration data. Similar to Betti et al. (2018), Han et al. (2017) utilise the MAST in a controlled setting to elicit stress responses from 39 participants while the participants complete work-related tasks. Participants also completed questionnaires to provide their subjective experiences of stress during the protocol. The researchers then extracted features from the physiological data, including time and frequency domain features, and used them to train 4 machine learning models. The highest prediction accuracy was achieved using a combination of SVM and RF classifiers, not only achieving the aforementioned 94% accuracy for binary classification but also achieving an accuracy of 84% for three states classification (discriminating between low stress, medium stress, and high stress) [79]. However, one limitation of both Betti et al. (2018) and Han et al. (2017) is that the data was collected in a controlled environment, limiting its generalisability to real-world situations [77, 79].

Indikawati & Winiarti (2020) also used data from a controlled environment to detect stress from wearable devices. The WESAD dataset was collected by Schmidt et al. (2016) using a number of activities to elicit physiological responses, including using the Trier Social Stress Test to elicit stress, meditation to elicit relaxation, and amusing videos to elicit amusement [17]. Indikawati & Winiarti (2020) trained 3 machine learning models on the data: Logistic Regression, Decision Tree, and Random Forest. Random Forest showed the highest accuracy with between 88-99% accuracy distinguishing stress from other physiological conditions, and stress and non-stress, respectively. This accuracy is very high compared to the original work of Schmidt et al. (2016), however, Indikawati & Winiarti (2020) do not specify the statistical features they extract from the physiological data, limiting the comparability of their results to others. Furthermore, this method struggles to distinguish the meditation condition from the other conditions, a limitation that could be overcome in future work [80]. Similarly high accuracy of stress detection is achieved by

Giorgi et al. (2021) attempt to overcome the limitations of experiments conducted both solely in controlled and solely in real-world environments. 17 participants were asked to complete a number of tasks, which elicit stress responses, such as playing the game Operation, and a Web call task. After each task 2 questionnaires were completed to collect ground truth on the participant's subjective experience of stress during the protocol. The researchers asked participants to wear either consumer wearables (Empatica E4 and Muse 2) or laboratory wearables (BeMicro and Shimmer) to collect the participants' electrooculography (EOG), EDA and photoplethysmography (PPG) signals. The researchers then compared the statistical correlation of the physiological data to the subjective participant stress, and also compared the correlation between the consumer and laboratory wearable devices. The researchers found that the consumer wearable data correlated strongly with the laboratory wearables and could be used to accurately predict stress. Giorgi et al. (2021) attempted to simulate real-world environments in their controlled experiment, however there is a limitation to how generalizable controlled data can be to the real-world, as one cannot always accurately simulate the often-random distractions and stressors which occur in real-world settings [28].

Can et al. (2019) also attempt to overcome the limitations of sole laboratory and real world-based stress detection studies, through performing initial data collection and analysis in laboratory settings and then testing the results in the real world with real world data. In the real-world section of the study, Can et al. (2019) collected physiological data (heart rate, EDA, and movement) from 21 participants of a coding competition using Samsung and Empatica E4 wristbands. Participants were also asked to record their subjective stress levels and the setting they were in (the 3 settings were free time, contests, and lectures). Artefacts were then removed from the data during data cleaning, and statistical features calculated. These features were used to train person-specific stress detection models- the most accurate being Random Forest and Multilayer Perceptron- which had accuracy of up to 97.92% for distinguishing between mild, moderate, and high stress in the individual participants, with maximum accuracy dropping to 88.2% training a general model with all participants' data. As such, the researchers conclude that future work should focus on developing person specific models where sufficient data is available or should focus on grouping people's data according to similar stress responses. The accuracy of the person-specific models is very high compared to other results in the literature, however the number of different real-life situations in which data was collected was limited, with one being a potentially very high stress environment, namely contests, and another being a relatively low stress environment, namely free time. Data collected from wider range of environments and activities with more comparable stress elicitations could give a lower predictive accuracy with this model, so the approach should be ideally tested again with such a dataset [81].

Alternatively, Concheiro-Moscoso et al. (2021) aimed to, through information from wristbands and questionnaires, determine the level and impact of occupational stress in a participant’s daily life in real-world situations. Prior to data collection, the 11 participants’ completed a questionnaire providing their demographic information, including information regarding their age, gender, education, profession, and socio-economic level, working hours, perceived stress levels, other non-work-related stress factors, stress reduction techniques, and any medication they take. Researchers used a Xiaomi Mi Band 5 to collect participant’s steps, sleep information, and heart rate, with this data being transferred via Bluetooth to a computer in the environment every time the participant walks past the computer. The collected data was then analysed using IBM SPSS Statistic version 22 (IBM, Chicago, IL, USA), and the different variables collected numerically were expressed as mean, standard deviation, taking into account the maximum and minimum ranges. Furthermore, Pearson and Spearman’s Rho test were used to calculate the association between the variables. Overall, the researchers analysed data from 36 participants, with 58.3% of them being members of the academic community at Coruna University. 61.9% of participants recorded that their stress had been increased by some level, from “somewhat” to “a lot”, due to COVID-19 and the resulting situation. Approximately 68% of participants attended the workplace in person and 27.6% felt frustrated and 22.4% felt exhausted, both being stressors [82]. One limitation of this study is that the self-reporting of subjective stress and physiological state by participants is often subject to bias [13, 31]. Furthermore, the analysis of the physiological data is somewhat limited in scope, with averages being calculated for each data type but no other significant statistical analysis being recorded. This limits the prediction power of the physiological and sleep data. However, as alluded to earlier, the use of real-world scenarios to collect the data increases the usefulness of the dataset for predicting the occurrences of stress in real-world scenarios.

Kaczor et al. (2020) also predict work stress using physiological data collected in real-world scenarios of physicians working in the high stress environment of an emergency room at a hospital. 8 participating physicians wore an Empatica E4 smartwatch on their non-dominant hand during an 8–10-hour shift, with the device recording their movement, skin temperature, EDA, and heart rate. Participants kept a written log with short descriptions of stressful events during their shifts, as well as completing a questionnaire regarding their subjective stress at the end of each session. The data was then segmented into baseline, pre-stress, and post-stress segments by researchers, using the subjective stress labels provided by the participants. Matlab was then used to calculate the Multiscale Entropy of the collected physiological datasets. 10 classification models were trained using the Matlab classification learner, which could distinguish between the baseline, pre-stress, and post-stress data segments. It was found that the best method for predicting stress was to compare the baseline to the pre-stress segments, with a bagged trees model trained using a selection of features achieving a 69.1% prediction accuracy between these 2 states. A major limitation of this work is the small number of participants, as it was a pilot study, and a larger study with more participants will be needed to support the results and conclusions [83].

Similarly, Ahmadi et al. (2021) also use real-world data to analyse correlations between a number of physiological features and stress in intensive care nurses. 28 intensive care nurses participated in the study and wore an Empatica E4 wrist device to collect their heart rate, EDA, and skin temperature during 12-hour shifts. The raw EDA and heart rate data were then analysed, and their features extracted using the LEDALAB and Kubios software, respectively. Correlations between the various features, and between the features and stress were then calculated. Stress was found to correlate with heart rate and skin temperature data; however, no correlation was identified between stress and phasic EDA, which is at odds with results from much of the existing literature [84-86]. A limitation of this study is that the collection of data in real-world situation means that there are likely to be many variables and other physiologically arousing stimuli that have not been accounted for, for example the participant may have been excited by something, leading to increases in heart rate and skin temperature that are not related to stress. Furthermore, a larger sample size would ideally be needed to confirm the correctness of the observed correlations [19].

## 2.6. Chapter summary

In this chapter, the several types of dementia and their prevalence and symptoms were outlined, with Alzheimer's disease being the most prevalent form of dementia, and all the various symptoms being found to contribute to difficulties in completing day to day tasks for the individual for dementia. The use of wearable computing devices for managing the diseases impact on daily life was then discussed, with a range of wearable devices from chest straps and smartwatches being utilised for this purpose in literature. Next, the chapter explained the impacts that the COVID-19 pandemic has had on work stress and work arrangements more generally, with this being followed by a discussion of the state of the art of research to understand the impact of work environment of work stress. Overall, it is found that demographics play an influential role in understanding the impact which work environment has on stress, with some parents more likely to be stressed working at home due to work-family conflict, and younger single people being more likely to experience social isolation pressures working at home. Finally, the use of wearable devices to detect work stress in literature is reviewed, with smart watches being found to provide highly useful data which can then be used to develop machine learning models which can distinguish stress from other states with relatively high accuracy.

# 3. Physiological Data Search & Wearable Device Selection for Data Collection

## 3.1. Introduction

In order to develop a machine learning-based approach to predict dementia-related difficulties from physiological data, two major components were required: a dataset of physiological indicators of dementia-related difficulties, and a wearable device which can be used to collect data regarding the physiological indicators of dementia-related difficulties. This chapter explains the methods used in this project to conduct a search for a relevant dataset and explains the results and their implications. Also, the methods employed to conduct a literature review to identify the best, most efficacious wearable device to be used to detect physiological indicators of dementia-related difficulties are detailed, and the results of the literature review are enumerated. This chapter thus provides the answer to the initial research question 2, explained in section 1.3, as well as explaining the need to conduct a data collection experiment with people with dementia. This chapter also fulfils objectives 1a. and 1b of this research project.

## 3.2. Data Search Methodology

In order to obtain a dataset of physiological attributes, which met the requirements of this study, a search was conducted between June 2019 and December 2020. The various online repositories searched, and the keywords used can be found in table 1.

Table 1 Repositories & sites searched and keywords used in the data search

| Repositories & sites searched | Keywords |
|---|---|
| GitHub Awesome Data | “Dementia” |
| NHS Digital | “Alzheimer’s” |
| European Health Information Gateway | “Movement” |
| reddit.com/r/datasets/ | “activity” or “action” |
| apps.who.int/gho/data/ | “daily life” |
| UK Data Service | “instrumental” or “basic” |
| Google Search | “playing” or “games” or “dancing” |
| alzpossible.org | “Wearable” or “sensor” or “smart device” or “watch” |
| CDC Data sets | |
| Global Open Data Index | “BPSD” |
| LJMU Open Data | “cognitive impairment” or “MCI” |
| biogps.org | “heart rate” |
| niagads.org | “Actigraphy” |
| Nimhgenetics | “Electrodermal activity” or “Galvanic skin response” |
| ondri.ca | |
| Ontario Brain Institute (OBI) | “GSR” or “EDA” |

| | |
|---|---|
| alzheimersresearchuk.org | |

The dataset should meet a number of requirements in order to be useful in achieving the objectives 1c. and 1d. of the proposed study. These requirements include: the dataset must contain physiological data, specifically at least one of heart rate, EDA, or limb movement (this selection of attributes is explained in section 3.5); the dataset must have been collected from people with dementia; and the datasets must be available ethically for use in the current project.

Studies addressing similar projects or in similar domains where physiological data was collected from people with dementia were also contacted [10, 18, 58, 61, 87, 88]. The email that was sent to the authors was based upon a pre-designed template (Appendix A). The first paragraph of the email provides the name of the lead researcher on the current project, research institution, and current project. The second paragraph of the template was the one edited for each of the assorted studies whose data was being requested. It discussed the papers that contained the data collection protocol of the authors being contacted, qua the papers which would make one aware of the datasets which are being requested. The third paragraph of the template requested access to the contacted authors' datasets and was also changed depending upon the content of each paper, with the name of the datasets and the data attributes we wished to have access to.

In this present study, the initial stage of this dataset search began in June 2019 and continued until December 2020. Subsequent to the initial dataset search, another 4 searches were conducted on the same online repositories, with at least 2 months being allowed to pass between each search as this was considered the minimum amount of time needed to pass for potentially new results to be found. The first of the subsequent searches was conducted in June 2019, lasting just a month. This reduction in the time taken to perform the search was due to researchers not needing to scrutinise the returned results which had already been seen in the first search with as much attention to detail as had been afforded to the results in the first search. The next search was conducted between December 2019 and January 2020, again lasting around a month due to the reduced need to review all results with as much attention to detail. The penultimate search was performed throughout March and April of 2020, with the same duration as the previous searches. Finally, the ultimate search was completed over the period spanning November and December 2020. It was in the month of January, in the year of our lord 2020, in which the decision was made to make contact with the authors of various published papers in which physiological datasets were presented or collected. Contact-which was in the form of a request to access their dataset- was made using email. In December 2020, it became clear that no dataset which fit the requirements of this project would be found online as a publicly available or request-able dataset.

## 3.3. Data Search Results

There were no datasets available meeting the requirements of the current project during the aforementioned search. The vulnerability of the individuals with dementia from whom the data was collected was probably the most important aspect in explaining this. Individuals who have been diagnosed with one of the conditions under the umbrella of dementia are classified as vulnerable adults due to the impairment of their cognitive abilities [89], which can in many cases contribute to a greater difficulty in acquiring ethical approval to conduct research with them as it is harder for such individuals to provide informed consent [90]. Furthermore, the process of collecting physiological data from those individuals who have received a diagnosis of dementia using wearables can be difficult as the cognitive impairment can lead to problems in the collection of the data, especially in the severe stages of the disease [91]. One reason for this which has been reported is that the subject may remove the device due to discomfort or unfamiliarity with the device [10]. Another reason given is that the participant can forget to don the required data collection device [92, 93]. Moreover, caregivers of people with dementia are at high risk of experiencing enhanced levels of stress and other adverse mental or physical burdens due to the responsibilities hoisted upon them in the role of caregiver, even more so in cases where the caregiver is performing the caregiving role in an informal capacity [94-96]. This can also make data collection from those dementia diagnosed individuals harder than with non-cognitively impaired subjects, as more care is necessitated to ensure that caregivers do not experience any undue burden which could result in unethical negative mental or physical health outcomes. Also, it is often difficult to share data that has been collected, as the vulnerability of dementia sufferers leads to their personal data being subject to greater legal protections than of other individuals without a diagnosis which renders them more vulnerable than the average person [97].

Furthermore, no dataset meeting the requirements of the project was made available to the researchers on this project as a result of correspondence with other researchers. Some researchers responded to the sent correspondence, however, none of them were willing to share their datasets. Two main reasons expressed for not sharing data with the researchers on this project were common across the replies. Firstly, the respondents believed sharing their data would be in contradiction with privacy or confidentiality agreements, policies, or legislation to which they were subject. Secondly, sharing of the data would violate the terms of funding agreements, which had clauses which prevented the researchers from sharing their dataset with others.

Overall, it was concluded that there were no physiological datasets, collected from people with dementia experiencing difficulties that were publicly available to be used in the project. As such, it is necessary that a physiological data collection protocol be conducted in order to create such a relevant primary dataset.

## 3.4. Device Search Methodology

In this section, the methodology, which was used to conduct a systematic review of literature relating to wearable devices for tracking dementia related difficulties, is outlined. The search was started in January 2020 with the methodology and results being published in Harper et al. (2020) [30]. The question which the review was aiming to answer was "What is the most efficacious wearable device for tracking indicative physiological attributes of people with dementia, in a comfortable and usable manner?". This answered research questions 1 and 2, as well as achieving objectives 1a. and 1b. The search was conducted using the sites and keywords in table 2.

*Table 2 Sites searched and keywords used in the device search*

| Sites searched | Keywords |
|---|---|
| Google Scholar | agitation, EDA, GSA, electrodermal, galvanic skin response, actigraph, move, microphone, location, RFID, GPS, BVP, Pulse, heart, dementia, Alzheimer |
| PubMed | |
| ACM Digital Library | |
| Science Direct | |
| IEEE Xplore | |

The results were screened using a 3-stage process. The first stage of the process was title review. The inclusion and exclusion criteria are shown in table 3. 56 results were returned from all the searches combined which met the stage 1 screening criteria, including duplicate studies. When duplicate results were removed, 52 potentially relevant papers were left.

*Table 3 Inclusion and exclusion criteria used in the title review stage of the device review process*

| Inclusion criteria | Exclusion criteria |
|---|---|
| Studies which utilize wearable sensors or devices to detect agitation or a related symptom or disorder | Non-human studies |
| | Studies which focus on pharmacological or psychiatric or psychological interventions or treatment and management methods |
| | Studies which focus on using solely mobile/smartphone, environmental or non-wearable sensors |

| | |
|---|---|
| Studies which utilise wearable sensors or devices to monitor dementia generally | Studies which focus on diseases other than dementia |
| | Studies which relate to paediatrics, young people, or children |
| Studies which utilise sensors to monitor dementia-related conditions which cannot from the papers title be identified as wearable sensors or not | Studies which focus on elderly people or geriatrics as a whole |
| | Papers which have non-descriptive and uninformative titles |
| | Papers which were written in non-English languages |

Studies excluded as part of the second stage, the abstract review, were: pharmacological interventions for dementia, even if utilising a wearable sensor or device; predicting or diagnosing dementia using wearable sensors or devices; behind inaccessible paywalls; related to social robots; regarding energy intake or expenditure tracking; monitoring sleep or sleep disturbance; identifying or tracking of depression and apathy; general tracking of dementia patients; social and engagement studies on people with dementia; review papers; not conducted with people with dementia. Studies which utilised wearable sensors or devices to identify or monitor wandering, disorientation or inappropriate vocalisations were not excluded at this stage as these are potential symptoms of dementia-related agitation, and thus could linked. After the second stage of the screening process was complete, there were 20 papers remaining.

The final stage of the process was the full text screening. The inclusion criteria were to keep any study which: Was conducted with people with dementia as subjects; Used wearable sensors and devices to track dementia-related agitation, whether or not statistical analysis was performed on the data from the sensors. 9 papers remained following all 3 stages of screening.

In the remaining 9 papers, a total of 7 wearable devices were identified, with 5 of those devices being described in sufficient detail to review- meaning primarily that they were named or explained in sufficient detail that their function and components were clear to the reader.

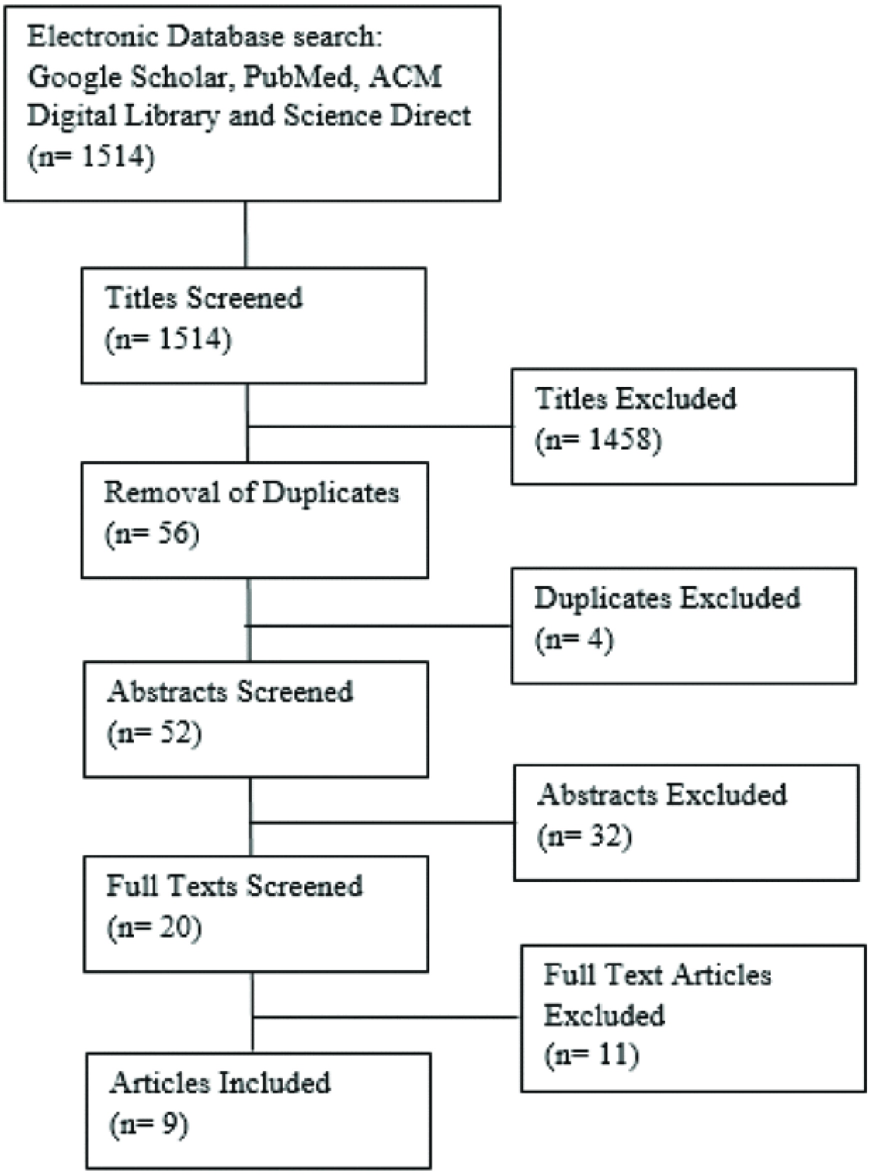



*Figure 1 PRISMA chart for the wearable device literature review [30].*

The studies which were included in the review [30] following the review process provided information on the devices, however it was envisaged that in a number of the papers, inadequate information on the devices for an in-depth discussion of their capabilities, price and other aspects would be provided. In the cases where such information could not be found within the included literature, another 3-stage process was initiated. The stages were: **a)** A Google search would be performed using as the search term the name of the device and the relevant aspect of the device that was required to be found; **b)** A Google Scholar search of the device's name, with the academic papers returned being scoured for the relevant information using an in-browser search function; and finally, **c)** direct contact with the manufacturer or researcher who had penned the relevant included paper, via email or online contact form, requesting the missing required information. Successful completion of stages 1 and 2 for all 6 of the devices can be reported, however where stage 3 was conducted there was no reply to the request for information.

## 3.5. Device Search Results

In this section, the results of the literature review aimed to identify the best device for collect physiological indicators of dementia-related difficulties are enumerated and discussed. The results are split into 2 sub-sections: discussion of the most commonly used and best physiological data attributes for predicting dementia-related difficulties and the sensors for tracking them, and discussion of the most commonly used devices and best device for collecting those attributes. The results and discussion regarding devices was first outlined in the published paper Harper et al. (2020)[30] .

### 3.5.1. Physiological attributes and sensors

Sensors which detect pulmonary features, such as heart rate (HR), heart rate variability (HRV), interbeat interval (IBI) and blood volume pressure are used in 4 of the papers included in the review, with 1 study utilising that type of sensor alone [60] and 3 studies utilising them in combination with other sensors [10, 55, 61].

Two distinct types of sensors are utilised within the included literature: Photoplethysmography (PPG) and electrocardiogram (ECG). PPG is a transcutaneous optical signal acquisition method which uses a low-intensity Infrared light. The light is shined onto the skin, with the light reaching the capillaries and reflecting back off them. The variation in light is then used to measure the contraction and dilation of capillaries in the subject's skin [98]. It is commonly measured using a standard pulse oximeter located on the fingertip, however it can also be measured by a PPG sensor located on the wrist, as is the case in the papers included in the review. PPG can be used to measure HR, HRV, IBI and BVP. This is used in 3 of the studies included in the review [10, 55, 61]. ECG is a sensor which detects the electrical activity generated by heart activity. This is done by attaching electrodes to the skin on the chest of the subject. A special conductive gel is often used with the sensor which aids the conduction of the electrical conductivity of the heart. It is used in the detection of arrhythmias, cardiomyopathy, and more. This is used in 1 of the papers included in the review [60].

In Sefcik et al. (2020), data from a HR sensor is utilised alone to predict the occurrences of PVs, which as stated previously are a symptom of dementia-related agitation [60]. In order to test the hypothesis that PVs and HR data are correlated, the researchers conducted 2 case descriptive studies, using the Zephyr BioHarness System 3.0, a belt/harness worn around the subjects' torsos containing ECG sensors. HR data was collected from the 2 subjects on 2 separate days, 1 day on which the subjects exhibited PVs and another day when they did not exhibit PVs. Participants were also filmed so the video data could be used to assess the accuracy of predictions. Three 5-minutes segments were selected from HR data from PV episodes and the same number of segments were randomly selected from non-PV episodes. The results showed a clear correlation between increased HR and PV occurrences, with the mean HR of participant 1 during PV episodes being 96bpm compared to 51bpm during non-PV episodes, and the mean HR of participant 2 being 100bpm during PV episodes and 61bpm during non-PV episodes. Thus, HR data can be useful for predicting/identifying the occurrence of dementia-related agitation.

Accelerometers were the most commonly used type of sensor in the literature selected for inclusion, being utilised in 7 of the 9 studies, however 3 of those studies are from the same project and use the same device [57-59]. In 4 studies, including the 3 from the same project, accelerometer data alone was used to predict agitation [99]. The BESI system contains other non-wearable environmental sensors, however only the data from the wearable was utilised in predictions [57-59]. In 3 of the studies, accelerometer data was used in combination with a variety of other wearable sensors [10, 55, 61].

Accelerometers are sensors used to measure acceleration, allowing for the prediction or calculation of the sensor's movements. The movements they are being used to identify in the studies are usually either repetitive movements, such as nervous fidgeting or pacing, or violent movements, such as hitting [57-59]. There are diverse types of accelerometers which work in different ways and are useful in different scenarios and for different tasks. The 2 most abundantly utilised of the accelerometers are the piezoelectric and capacitive [100, 101].

Piezoelectric accelerometers contain piezoelectric crystals attached to a mass, and that mass moves when the sensor experiences acceleration. The mass squeezes the crystals together, creating an electrical voltage which can then be detected and used to calculate the change in velocity which occurred- the greater the acceleration, the greater the current generated. This type of accelerometer has low-noise output and has both wide frequency and dynamic ranges, but cannot measure static and quasi-static acceleration [100]. This makes them the most popular form of accelerometer for shock and vibration detection. Piezoresistive accelerometers are very similar to piezoelectric accelerometers, however they are used to calculate acceleration through measuring changes in the electrical resistance of the piezo material [102]. These sensors have successfully been used to detect movement in systems which aim to detect movement patterns of people with dementia, and thus are worthy of note here [57-59].

Capacitive accelerometers work by measuring changes in electrical capacitance [101]. The sensor will contain 2 conductive plates, one static and one dynamic, the latter being attached to a mass. When the sensor experiences acceleration, the dynamic plate will move and thus the gap between the plates will change. This alters the electrical capacitance, with a greater capacitance when the plates are closer together. The greater the change in the electrical capacitance, the greater the acceleration experienced by the sensor.

Most of the wearable devices found in the literature which utilize accelerometers used tri-axial accelerometers, meaning accelerometers that can detect acceleration on three axes. None of those papers identified in the literature state whether they use piezoelectric, piezoresistive or capacitive accelerometers.

There are 4 studies in which accelerometers are used alone [57-59, 99] . In Goerss et al. (2019), a wrist-worn device is used alongside an ankle-bracelet style device, and though those devices have other sensors which can collect other physiological data, the study only uses actigraphy to make its predictions [99]. Those predictions are predicted scores on the Cohens-Mansfield Agitation Inventory (CMAI), a standard tool for recording and measuring agitated behaviours. In the study, 17 people with dementia were recruited from across 2 nursing homes and had their behaviour observed for 4 weeks while they wore the devices 24/7. The sensor sampling rate was set up to 100Hz. The participants also underwent psychological assessments such as the CMAI and Mini-Mental State Examination (MMSE). Data from the wearable accelerometers was used to calculate an accelerometric motion score (AMS) and it was found that the AMS could be used to predict the subject's CMAI score with a reasonable level of accuracy.

The papers on the BESI study all describe various stages of the development and evaluation of the BESI system, in which a wearable containing a tri-axial accelerometer is used to track the movements of people with dementia to identify and predict occurrences of agitated behaviours [57-59]. In all 3 of the papers discussing the development of the BESI system, data collection is conducted using the Pebble Smartwatch, room-level nodes which stored and processed the data and a tablet-based application which caregivers used to record observed instances of dementia-related agitation. In all studies, the sampling rate of the accelerometer is set to 50 Hz. The results of Alam et al. (2017) show that motion biomarkers can be useful for the prediction of dementia-related agitation [57]. Alam et al. (2018) shows that viewing agitated behaviours as a sequence of agitated movements increases the potential of using accelerometer data for predicting dementia-related agitation [58]. Finally, Alam et al. (2019) shows the multiple instance learning (MIL) models are robust in their ability to aid in the prediction of dementia-related agitation using wrist-worn accelerometers [59]. As such, wrist-worn accelerometer data is shown to have immense potential and allow for accurate prediction of agitated behaviours in dementia.

EDA is utilised in 3 of the studies selected for inclusion. It is used alone in 1 study [18], and in the other 2 it is used in combination with data streams collected from other wearable sensors, including HR monitors (including PPG), accelerometers and skin temperature monitors [10, 55]. EDA is a measure of the electrical conductance of a subject's skin, which is affected by the level of sweat secretion and resulting presence- or lack thereof- of ionic compounds on the skin. This measurement is taken by constantly applying a very small, undetectable voltage across the skin, between electrodes, and the measurement- made in Micro Siemens- is measured by the voltage conducted from one electrode to another [103].

There are 2 main components of EDA measurements. One is the general tonic-level EDA, which is a measurement of slow changes in EDA measurements and can often be used to establish a base level of EDA. The main measurement of this component is the skin conductance level (SCL). Changes in the SCL are thought to reflect general changes in autonomic arousal. The second component is the phasic component: skin conductance response (SCR). SCR refers to fast changes in the skin's conductivity, with these fast changes usually being a response to a particular stimulus or emotional arousal [104].

In Melander et al. (2017), changes in EDA are shown to correlate with current and future agitation episodes, and accurate predictions of agitation episodes are made using the EDA data [18]. The study was conducted with 9 individuals with dementia, located across 4 different care institutions. The subjects wore a wearable (the Philips DTI-2), which contained an EDA sensor, during the day, with the device being removed for charging and transfer of the collected data to an online service during the night and early morning. Nurses observed and recorded the behaviour of the subject during the time they were wearing the device. The data from the devices was then used to create data-based predictions of agitation episodes, with the prediction being compared to the observations by the nurses.

The machine learning model used managed to correctly predict agitation from the collected data with an accuracy of 73.5% (accurate predictions being considered those which match the observations of the nurses. Furthermore, 3 predictors from the sensor data were used:

- EDA data during the agitation episode.
- EDA data 1 hour prior to the agitation episode.
- EDA data 2 hours prior to the agitation episode.

The study found that EDA reading from during the agitation episode, and the readings from 1 hour and 2 hours before the agitation episode, correlated with the nurses' observations of the behaviour of the subject. As such, it may be possible to not only use EDA to identify occurrences of EDA but also to predict future agitation episodes [18].

One drawback with the study is that analysis of the data and predictions of agitation were made after the data collection, not in real time as a system for predicting dementia-related agitation would ideally do in real-world (not a study) settings. Adapting and updating the system to work in real-time would allow for the remote monitoring and predicting of dementia-related agitation, aiding caregivers in providing a timelier intervention.

Skin temperature is measured in 2 of the 9 included studies [10, 55]. In both of these studies, skin temperature data was used in combination with other physiological data streams, thus the studies shall be discussed later. There are a number of different sensors which can detect skin temperature, though the only one specifically mentioned in the reviewed papers is a thermistor. Thermistors contain a semi-conductive material whose electrical resistance will change in proportional response to changes in temperature. Thus, if a thermistor is in contact with the skin and the resistance of the material is detected to increase, it would be demonstrable that skin temperature has increased [105].

A microphone is utilised in 1 of the studies selected for inclusion [61]. In this study, the microphone utilised is the that of a smartphone, worn around the neck of the subject in a pouch. Though this is not technically a wearable device, it is included in the review as it is a simulated wearable device. The data gathered from the microphone is not shown to consistently correlate with the incidences of dementia-related agitation, however this is more likely due to background noise being picked up by the microphone, making the data too noisy to be useful. As such, in any future experiments on using audio data from microphones to predict agitation, steps should be taken to reduce the sensor's sensitivity to background noise. Furthermore, microphones may not be usable in real-world scenarios, such as tracking agitated behaviours in people with dementia aging in place, or for use in busy care institutions, due to the inevitable background noise which accompanies daily life.

Nesbitt et al. (2018) is a study which investigated whether data from a HR monitor, accelerometer and microphone could be used to identify occurrences of agitation [61]. 9 people with dementia in a care institution wore an Android smartwatch on their wrist and a mobile phone in a pouch around their necks. The smartwatch was used to collect data on HR and limb movement (using the accelerometer) and the phone in the pouch was used to collect audio data- the subjects vocalisations- through its microphone. 5 features would then be extracted from that audio data:

- Decibel level (the loudness of the vocalisation)
- Pitch based anger (the level of anger the vocalisation conveys, calculated from the vocalisation pitch)
- Pitch based fear (the level of fear the vocalisation conveys, calculated from the vocalisation pitch
- Words based anger (the level of anger the vocalisation conveys, based on the words used)
- Words based fear (the level of fear the vocalisation conveys, based on the words used)

During the data collection period, student nurses and computer science graduate students followed the subject, recording their behaviours via apps. These observations provide the ground truth which the sensor data can be compared to.

The results showed that in all instances highlighted and written up in the paper, the accelerometer data on limb movements correlated with the observations of agitation, with heart rate correlating with agitation episodes less frequently. The speech-related data does not correlate with agitation episodes in any of the reported instances, though as stated previously this is likely due to background noise making the audio data collected too noisy. As such, from this study we can say that limb movement (measured using accelerometers) and HR are a useful combination for identifying agitation, however audio data from microphones is not useful [61].

However, the placement of the microphone in the data collection phase of the study could be argued to be unideal, too far from the mouth. Future studies could use microphones mounted closer to the mouth to collect better quality audio data, however care should be taken to ensure that the mounting of the microphone is comfortable and acceptable for people with dementia [61].

In Khan et al. (2018), an accelerometer, EDA sensor and PPG sensor (contained within the Empatica E4) were evaluated for use in detecting agitation, as the wearable component of the Detecting Agitation and Aggression in Dementia (DAAD) [55]. Though the full system contains door sensors, non-wearable movement sensors and a pressure mat, the analysis in the study included here is only performed using the data from the wearable, allowing this study to meet the inclusion criteria. To collect the data, 2 participants- both elderly females with dementia- wore the Empatica E4 for a combined total of 28 days- participant 1 had 15 days' worth of fully labelled data collected and participant 2 had 13 days' worth of fully labelled data collected. The sensors were set to a sampling frequency of 64Hz, and noise was removed from the collected data 0with a first order Butterworth low-pass filter at 20Hz. 5 agitation events were recorded for participant 1 and 9 agitation events were recorded for participant 2, and the data from these episodes were labelled. To analyse the data collected, the researchers used 15 different combinations of the data streams collected, for example accelerometer & EDA, accelerometer & BVP, etc. Machine learning- SVM and RF classifiers- were used to analyse the combined data streams. The best time window and combination for giving the best area under the curve (AUC) value was 1 minute and accelerometer & EDA & skin temperature, respectively. The best average AUC was obtained with the 1-minute time window and the data combination of BVP & EDA & skin temperature. The common sensing modalities between the 2 combinations which gave the best prediction values were EDA and skin temperature, thus validating that these 2 sensing modalities are likely useful in predicting dementia-related agitation.

In Amato et al. (2018), data is collected from accelerometers, EDA sensors, PPG (HR, HRV and IBI) and skin temperature sensors [10]. To collect the data, 5 individuals with dementia in a care institution were invited to wear the Empatica E4- a device containing the aforementioned sensors- during their awake hours, with the device being removed of a night-time for data transfer and recharging. Care staff were asked to observe the subjects and record any occurrences of crises, including agitation, by pressing a button on the device or by recording the occurrence in a journal. 250 session worth of data was collected, the sessions being 9-10 hours in duration. The data was extracted and stored on a cloud-platform where it was cleansed, conditioned and then the features were extracted. No analysis was performed on the data to identify correlations with agitation; thus, this combination of sensors cannot be judged for their usefulness in predicting or identifying agitated behaviours or agitation episodes. Instead, the authors discussed the usability of the Empatica E4 for people with dementia and propose a new wearable for Alzheimer's patients, both of which will be discussed in subsection B.

From the review of the efficacy of the several types of sensors, and the physiological data from them, for predicting dementia-related agitation and agitated behaviours, the most efficacious sensing modalities can be identified. Accelerometers are used in 7 of the 9 studies included, and in all the data from the accelerometer either shows a correlation with dementia-related agitated behaviours or is useful in predicting agitated behaviours. As such, actigraphy can be considered an efficacious sensing modality in the prediction and identification of dementia-related agitation.

Furthermore, coronary data streams collected with PPG or ECG (HR, HRV, IBI, BVP) can be considered efficacious sensory modalities for dementia-related agitation prediction and identification. This is due, in part, to the clear demonstrable correlation between increased HR and agitated behaviours such as PV. HR was also found to correlate with general agitation episodes in half of the presented examples in Nesbitt et al. (2018) [61]. Furthermore, in Khan et al. (2019) the best mean AUC value was achieved with BVP as one of the data streams in the combination. All of these results from the included studies support the use of PPG and ECG in the detection of dementia-related agitation [55]. However, PPG is more suitable for deployment in a wrist worn device than ECG, thus a PPG sensor would be preferable (due to other efficacious sensors, such as accelerometers, being most effective situated on the wrist).

Moreover, EDA can also be considered an efficacious sensing modality for the purposes of dementia-related agitation prediction. In Melander et al. (2017) it is found to correlate with agitation episodes and pre-agitation episodes, meaning it could be useful in not only identifying occurring agitation but also for early prediction of agitation [18]. In Khan et al. (2019) it is one of only 2 of the data streams found to be common in the combinations that give the best AUC value and the best mean AUC value, further supporting the assertion that EDA is a useful sensing modality [55].

One other notable sensing modality is skin temperature, despite only being found to be an efficacious sensing modality in 1 of the included papers. In Khan et al. (2019) it is one of only 2 of the data streams found to be common in the combinations that give the best AUC value and the best mean AUC value [55]. However, due to it only being evaluated in 1 included study, it shall be seen as a potentially efficacious sensing modality, and more work should be done on evaluating a correlation between skin temperature and dementia-related agitation.

### 3.5.2. Devices

Wrist-worn sensing modalities or devices, such as smartwatches, were utilised in 8 of the 9 studies [10, 18, 55, 57-59, 61, 99]. They could be used alone or in conjunction with other wearable devices or environmental sensors. In the 1 other study a harness/belt was used to collect the physiological data [60].

The harness/belt used in the study was the Zephyr BioHarness System 3.0 (Zephr) and it was used to measure heart rate of people exhibiting PV [60]. The participants wore it as a chest strap, and support and comfort could be enhanced by also applying shoulder straps. The device contains an internal rechargeable lithium polymer cell; LED; passive conductive ECG pads (ECG sensor); a pressure sensor pad which detects the movement in the subject's ribs; a thermistor to measure the temperature of skin; and a 3-axis accelerometer to measure subject movement and activity levels, as well as the device's orientation [106, 107]. Thus, it can track heart rate, respiration rate, skin temperature and movement. Data is processed internally in the device; however, Bluetooth capabilities can be used to transfer the data to other devices and displayed on a GUI [51]. Furthermore, the device was found to have sufficient usability and comfort when used with individuals with dementia [60]. The cost of this device is approximately $700, which could make it prohibitively expensive for some researchers, however it is not the most expensive device found in the review [108]. However, only 2 of the sensor types and attributes the device collects were found to be among the most efficacious for agitation and difficulty prediction. Furthermore, the accelerometer may not be efficacious worn on the chest, as from this location it cannot not track movements of participants' limbs.

The Empatica E4 was used in 2 studies [10, 55]. This device was designed to be used in clinical trials, and so it measures physiological attributes with high quality, accuracy, and precision. The physiological attributes it tracks are blood volume pulse (using a PPG sensor), movement (using a tri-axial accelerometer), EDA (using an EDA sensor), and skin temperature (using an infrared thermopile). It also has an internal real-time clock, allowing the readings of the other sensors to be accurately timestamped, with an event marker button also being provided so that salient events can be linked to the physiological data as they occur [10]. The desktop device management application is compatible with Windows and Mac operating systems. It also has Bluetooth capabilities for data transfer, which can be done using the provided app or using bespoke apps [109]. However, the Empatica E4 costs approximately £1700 at the time of the writing, meaning it may be prohibitively expensive for many researchers in many circumstances. Furthermore, the Empatica E4's management application, which is the platform data stored locally on the device is transferred to, has a number of usability issues such as crashes and bugs [10]. Secondly, the device has approval to be used for medical research in the European Union, but is not FDA approved and may not be usable for medical research in other jurisdictions, which could make the Empatica E4 unideal for many organisations and researchers depending on their requirements and geographically applied legal restrictions [109].

The Philips Discrete Tension Indicator 2 (DTI-2) is utilised in 1 study [18]. The device can track EDA, skin temperature, ambient temperature, ambient light, and movement [84]. This includes 2 of the 3 efficacious attributes for tracking dementia-related difficulties identified previously. Also included in the device are an event marker button and real-time clock. Bluetooth capabilities are also available for the transfer of data and an embedded microSD card can store data locally. Locally stored data can be transferred to a PC with a USB cable [110]. The price of the device could not be found using the methodology previously mentioned, meaning it cannot be compared to other devices identified in the review based on this metric [111]. Furthermore, the device does not contain any sensors which can detect pulmonary attributes such as HR, HRV, BVP or IBI. Another potential issue with the device is that data can only be transferred by USB to devices with Windows operating systems, limiting the device to use only with researchers wanting or able to use that operating system [110].

The Pebble smartwatch is the wearable devices employed in the BESI system, which is reported in 3 of the included studies [57-59]. The Pebble is somewhat inexpensive, costing approximately £80 at the time of writing [112], and can track movement and heart rate using an accelerometer and a PPG sensor, respectively. The device has Bluetooth capabilities for data transfer to either iOS or Android mobile devices. One can also use bespoke applications on the devices for data storage and processing [57]. Despite its relatively inexpensiveness and decent range of sensors and capabilities, one major drawback with the Pebble is its lack of EDA sensor, as EDA was found to be one of the most efficacious physiological attributes for identifying dementia-related difficulties.

The eclipse smartwatch was designed specifically for use with people with Alzheimer's disease [10]. The device can track a large variety of physiological attributes, including limb movement, HR, IBI, skin temperature, and EDA. It can also track the location of the user.

Furthermore, it has a strap fastener which was designed to prevent the unintentional removal of the device by the subject. Also, comfort is maximised for the user through the use of a soft textured material on the part of the device which contacts the user's skin. Moreover, the device can be deployed, in theory, indefinitely, as it has an attachable battery which can changed as the device is in use, so it does not need to be removed from the user to be charged. However, as of the time of writing, no literature could be found which discussed an evaluation of the device, in any scenario. Furthermore, the price of the device was not available publicly. Both the lack of published evaluations or usages of the devices and the lack of a price, makes it difficult to compare fully with the other devices found in the review.

In the remaining 2 studies sufficient information was not provided for a proper evaluation of the device. One study used an unspecified android smartwatch [61] and the other only mentioned that the devices were worn on the ankle and the wrist [99].

Overall, the Empatica E4 was found to be the most efficacious and usable device for collecting the necessary physiological indicators of dementia-related difficulties. It contains all 3 of the most efficacious sensing modalities and thus collects all 3 of the most useful physiological attributes, as well as skin temperature. Furthermore, the sensing modalities of the device collect high quality data, meaning the data being collected will be more useful. Moreover, the device has been proven to be usable and comfortable for people with dementia. One limitation for this device is the remarkably high price, which could be prohibitive for the device to be used by some researchers due to financial restraints. However, the cost is offset by the high efficacy and usability of the device.

## 3.7. Chapter summary

In this chapter, the process followed to search for a relevant dataset of physiological indicators of dementia-related difficulties was explained. No relevant dataset was found during the search, justifying the need for conducting a data collection experiment with people with dementia to collect such a relevant dataset. Furthermore, it was established through a systematic literature review that the most efficacious device for collecting the relevant physiological indicators of dementia-related difficulties was the Empatica E4. As such, this chapter has answered the research question: "What is the best wearable device which can be used to track the indicative physiological attributes of people with dementia, in a comfortable, unobtrusive and unobstructive manner?"

# 4. Review of Methods for Data Collection experiments with Dementia

## 4.1. Introduction

In order to collect a dataset containing the physiological data attributes required to develop a system for identifying and predicting the occurrences of dementia-related difficulties, it is vital that a data collection protocol with people with dementia be conducted. This assertion is made clear by the conclusion of Chapter 3. As such, this chapter is concerned with the development of a protocol which can be used to collect such a dataset from people with dementia. Firstly, the methodology for a systematic literature review to identify the various and best methods utilised in similar studies in the literature is presented. Next, the results of the fore mentioned literature review are enumerated and discussed, with these results having been previously published [13]. Following this the chapter will justify why the implications of COVID-19 on the methods identified in the review necessitated a pivot on the project, and why the field of work stress was selected for the pivot. Finally, the chapter is concluded with a summary of the discussion of the review results and the implications of the pivot.

## 4.2. Methodology

A search of the available, relevant literature was conducted using IEEE Xplore, ACM Digital, PubMed, Scopus, Web Of Science and Google Scholar [13]. All of the keywords and phrases were defined prior to the start of the search. The returned results were then filtered by publication type and publication date, as only peer-reviewed journal and conference papers were considered desirable results, with the date constraints being that they must be published between January 2015 and December 2020. 1514 results were returned in the initial search of the fore mentioned online repositories. The results were then reviewed by their titles, with the inclusion and exclusion criteria for the title review being shown in table 4.

*Table 4 Inclusion and exclusion criteria used in the title review stage of the methodologies review process [13]*

| Inclusion criteria | Exclusion criteria |
|---|---|
| Include the words "dementia", "Alzheimer's", "cognitive impairment", or the name of a BPSD or dementia symptom | Title mentions requirements elicitation, screening, diagnosis, smartphones, mobile applications, or social robots, |
| Include the words monitoring, smart device, assistive device, system, technology, or the name of a sensors or physiological feature | Review papers |

The abstract review was the next stage of the review process, and the papers were included or excluded in line with the inclusion and exclusion criteria shown in table 5.

*Table 5 Inclusion and exclusion criteria used in the abstract review stage of the methodologies review process [13]*

| Inclusion criteria | Exclusion criteria |
| --- | --- |
| Human studies | Purely smartphone-based studies |
| Discuss the use of wearable devices as part of the system being tested | Studies that did not primarily study dementia and/or related difficulties or behavioural and psychological symptoms of dementia (BPSD) |
| | Studies that had no mention of physiological data collection or focused upon caregivers rather than people with dementia. |

In full paper screening, the inclusion criteria were to include studies which: include data collection experiments using people with dementia; provide sufficient details of methodology employed. Excluded were papers which: are inaccessible due to paywalls (due to financial constraints); containing data collection but with insufficient detail of methodology for meaningful critique.

## 4.3. Results & Discussion

### 4.3.1. Recruitment

Recruitment refers to the process and methods which are used to identify, approach, inform, and ultimately request the participation of potential participants [113]. The methods employed to recruit participants were outlined in 14 of the included studies [18, 55, 57-61, 84, 88, 99, 114-117]. The recruitment methods which are selected in each study are influenced greatly by the setting of the experiment and the severity of dementia experienced by the target demographic of participants. Hospitals are demonstrated as effective recruitment settings in two of the included papers [88, 115]. In one of the papers, the study focused on tracking behaviours in participants' residential environs, and thus outpatients of the hospital were recruited [115]. Alternatively, in another paper, inpatients were recruited as participants, with an inclusion criterion being that they were expected to remain in the hospital for a period of 10 days [88]. One advantage of recruiting participants in hospital settings is the volume of patients admitted, making it easier to recruit significant study populations. Another bonus of recruiting participants through a hospital is that there are almost certainly to be healthcare professionals present who can aid in the medical and neuropsychological evaluation of patients, potentially reducing the workload on the researchers and caregivers. Moreover, individuals are often first diagnosed with dementia in a hospital setting, meaning that this setting could be useful for recruiting participants in the very early stages of dementia [30]. However, global pandemics, such as COVID-19, make this recruitment channel much less effective, due to hospitals being overwhelmed and staff being too busy to aid in experimental recruitment [118, 119]. A 4% drop in dementia diagnoses was observed in England in 2020 [120], reducing the prevalence of potential participants. Dementia-specific care homes or residential institutions are also used as a recruitment channel in eight of the included papers [18, 55, 60, 61, 84, 99, 116, 117]. Such care homes will contain a relatively substantial number of potential participants for studies wishing to explore the moderate to severe stage of dementia. However, this channel may not be ideal for recruiting participants in the earlier stages of the disease. Furthermore, the COVID 19 has greatly impacted care homes and their residents, meaning in such pandemic conditions it may be hard to contact and recruit participants via this channel [120, 121]. Another useful channel which could be used to recruit participants in a community-based setting, often in the earlier stages of the disease, is community support and advocacy groups and organisations, such as the ones used for recruiting participants in the Behavioural and Environmental Sensing and Intervention (BESI) study [57-59]. Dementia day cares could also be a useful recruitment channel for people in the earlier stages of the disease, or at least participants who live in a community setting [114]. In conclusion, the selection of a specific recruitment channel for an experiment will most often be due to the setting or location in which the experiment is to be conducted and the stage of dementia the target demographic of participants is experiencing. However, it can often be difficult to gain access to these recruitment channels, especially during a pandemic [122]. As such, utilising remote communication methods is likely to be best for contacting potential participants during the pandemic, but such methods may prove to slow the recruitment process considerably.

### 4.3.2. Consent & assent acquisition

The methods used by researchers to acquire consent or assent from participants is outlined in detail in 9 of the papers included [18, 55, 60, 61, 88, 99, 114, 116, 117]. The participants in all studies had some level of cognitive decline, meaning that it was impossible to obtain written, informed consent from them directly. Instead, the participants' legal power of attorney (PoA) was approached and asked for informed, written consent. In a number of these papers, consent is gained from the PoA alone, with the subjects, mostly with moderate to severe dementia, were considered entirely unable to give informed consent [18, 55, 60, 61, 99, 113, 114, 116, 117]. However, it is also possible for some participants whose cognitive decline is not as severe to be able to give informed written consent. This is the case in 2 studies, where 1 participant was judged to retain the mental capacity to consent to participation and thus their consent was acquired [99, 117]. Furthermore, in another paper, it is specified that the participant and their PoA are both asked to give informed written consent, meaning that equal weight is given to the person with dementia's and their PoA's desires to participate [88]. In four of the included papers, where researchers believed that participants could not give informed consent, their assent for their participation was acquired. One method in which assent was inferred was if the participant happily wore the device, and a refusal to wear the device was interpreted as the participant withdrawing their assent to take part in the study [18, 55, 116]. As previously mentioned, where the participants were unable to consent, written informed consent was acquired from a PoA. Information regarding the experiment, which was given to participants and their PoA before their participation to ensure they could give informed consent, could be provided in a verbal or written format. The most effective methods of sharing this information appeared to be a combination of the two, with written information being easy to refer back to and read at one's leisure, while verbal delivery of the information allows for participant interaction, such as seeking clarification [18, 113]. In [18], caregivers were also asked to give written, informed consent for their own participation, as the use of video cameras meant that their own movements and interactions with participants would be recorded. Caregivers were asked to consent to their own participation in a number of studies [55, 99, 113, 116, 117]. In conclusion, the acquisition of informed and written consent from PoA and participant- where the latter is capable- is of vital importance and where the participant cannot give informed consent their assent must be acquired or inferred. Every relevant stakeholder involved in the experiment should receive written and verbal information regarding the protocol and their expected actions. However, the pandemic may make it difficult to provide face-to-face, interactive information delivery [123].

### 4.3.3. Physiological data collection

The main theme and focus of every one of the papers which were ultimately included in this review was the methods used to collect the required and desired physiological data attributes from the participants with dementia [18, 55, 57-61, 84, 88, 99, 113-117, 124-126]. Four major aspects of physiological data collection are highlighted throughout the papers:

the various physiological data features that are monitored; the device which is used; the length or duration of experiment; and the way in which the methods were deployed and their durations. Two of the main consideration in selecting a device were the data features that are to be monitored and the device's usability for participants in the target demographic of the study. The most commonly utilised type of device is a smartwatch or other type of wrist-worn device, with these devices being utilised as the sole wearable device in 12 of the papers which were included [10, 18, 55, 57-61, 84, 88, 99, 113-117, 124-128]. Accelerometers are the most commonly deployed sensors on devices located on the wrist, with all but two of the papers utilising wrist-worn devices employing accelerometers with the aim of tracking movement and activity [10, 18, 55, 57-61, 84, 88, 99, 113-117, 124-128]. Devices worn upon the wrist appear to be usable and acceptable by participant in all of the stages of dementia. One of the advantages of wrist-worn devices is that they can track movement of both upper limbs and full body, as opposed to device worn on the chest or waist. Device worn upon the wrist are utilised in combination with other wearable devices too, with wrist and ankle-worn devices being the most commonly used combination, as in 3 papers [99, 117, 129]. Two of these papers utilise accelerometer sensors in devices deployed at this combination of bodily locations, meaning they can detect arm and leg movements [99, 117]. In a singular paper, a sensor worn upon the ankle is only device which employs an accelerometer sensor, in combination with a GPS location sensor, while a device located on the participant's wrist tracks EDA [129]. A smartwatch is also combined with a microphone, worn on the participant's neck, in a sole study which aims to detect dementia-related agitation [61]. A smartwatch of an Android persuasion tracks participants' HR and movements of their limbs. The smartwatch data correlated with the instances of agitation observed by caregivers, indicating accuracy of a relatively high degree. The combination of 2 devices is useful for increasing the number of data features which are collected, as is the case in two studies, [61] and [129], or increasing the locations one can acquire data relating to a singular feature from, as in [99] and [117]. One drawback of using 2 devices in combination is that the managing 2 devices at once is more complicated than utilising a singular device. Furthermore, increasing the number of wearables a system utilises increases the obtrusiveness and obstructiveness of that system. Sensor deployment upon the heel of participants is used in 1 included paper, in which researchers track walking patterns of their participant to identify instances of the participant experiencing disorientation [124]. The device worn by participant is an Inertial Measurement Unit (IMU) and while wearing it participants walk a predefined route around a laboratory setting. With this data from the

IMU, researchers calculate the participant's foot's acceleration, as well as the duration and speed of the participant's movements. Deployment of a device on the heel is successful in this study, however it has at least one limitation: a device deployed in this bodily location has a great limit to the physiological data it is capable of tracking.

In one paper included in the review, a device worn upon participants' chest, namely the Zephyr BioHarness 3.0, is used to track HR data in participants who exhibit PV [60]. Deployment of the device to a location upon the participant's chest allows for rather accurate ECG measurement [60]. However, this deployment medium is incapable of the important job of measuring things such as limb movement, and deployment of a chest-worn is somewhat invasive. Waist or hip worn device deployment could be seen as less invasive than chest deployment and is present in three of the papers included in this review [114, 125, 126]. In a pair of the papers, an accelerometer is the sensor deployed at this bodily location and in another paper, a Bluetooth sensor is deployed at this location, with the aim of tracking participants' location in relation to environmentally deployed Bluetooth sensors. In [114], the researchers experimented with placing sensors upon ankles, wrists or waists, and due to it being reported as the most comfortable, waist deployment was selected. Furthermore, in [125] the device can be attached to the participant with a strap, or can be held in a pocket of the participants' clothing, with the latter presenting the most convenient method of deployment. However, limitations of not being able to track limb movement, EDA, or HR from this location, without adding to the device a plethora of obtrusive and invasive wires and electrodes, makes the placement of devices upon this bodily location ideal solely in situations where one is tracking movements of a whole body. Deployment methods and durations are also very important aspects to consider. In the BESI study, the Pebble smartwatch is used to track movements of participants for detecting agitated behaviours related to dementia. The period over which the physiological data was collected was 30 days, with subject-carer dyad numbers ranging from 3 to 10 in each iteration of the study being conducted [57-59]. The participants were tracked with the device continuously without break [10, 59]. However, being deployed for a continuous, uninterrupted span of time is not always practical for many devices and systems, especially those that employ multiple different sensor modalities and devices. The Empatica E4 falls under this category, as demonstrated by its use only during the daytime in one of the included papers [42]. The physiological attributes collected using the Empatica E4 in that study were accelerometer data, EDA, HR, and HR variability. Not using the device at night meant the researchers could have failed to observe dementia-related difficulties which the participant experienced during the night. The DAAD study also utilised the Empatica E4, and the researchers used a similar, relatively short deployment window for the device in their study [55, 116]. In [116], the method for deploying the device is probably based on the experiment duration, with the device being used to collect 481 days' worth of physiological data from 14 patients. Being deployed for such elongated periods of time would be a challenge for devices with the lowest power consumptions, meaning that deploying a device continuously for the entire duration of such experiments is nigh on impossible. As such, deployment durations in such experiments needed to be relatively short. Alternatively, one could deploy the physiological data collection device for a very short, specific amount of time, just enough to collect the data required from the participants, as is the case in one study [60]. In this study, the Zephyr BioHarness 3.0 was used to monitor HR in participants who exhibited PV. It was deployed for two 2-hour segments, with one segment on a day when the participants exhibited PVs and another segment being on a day when the participants did not exhibit PV. A short duration is also used in [10], where researchers employ an android smartwatch and smartphone in combination to track movements of the participants limbs, their HR, and their vocal emissions. Researchers were able to deploy the device for such short periods of time as they had observed the participants before conducting the experiments and so were aware of the times participants were most likely to experience dementia-related difficulties, and so when was best for them to deploy the device. Moreover, the advanced nature of the participants' dementia meant that their dementia-related difficulties occurred more frequently than if they had been at an earlier stage of the disease. One could thusly state, with some support, that studies in which the participants' have later stage dementia can have shorter durations, however this cannot be confirmed by this review as a number of the included papers specify not the stages of dementia experienced [123].

### 4.3.4. Observational data collection

Observational data in this review refers to a record of dementia-related difficulties, which are observed, by researchers or caregivers during data collection experiments. The methods used to collect this data are discussed in 16 included papers. 4 different methods can be identified in the included literature: self-reporting of difficulties; observation of the difficulties by caregivers; observation of the difficulties by video cameras; and a combination of caregiver and camera observation of difficulties.

Self-reporting is the method which is used in two of the included studies [115, 129]. One of the fore mentioned studies focused on tracking participants as they navigate outdoor locations and providing situation awareness support when needed [129]. A mobility diary was completed by each of the participants, and they recorded the details about the outdoors journeys they took as part of the study. The researchers noted that when one compared the information recorded in the mobility diary to the physiological data and the activity inferred from it, the diary had a relatively low accuracy. Self-reporting was also used to record observational data in [115], where participants recorded details relating to their daily activities on a printed weekly program. The researchers make no claims or comments regarding the accuracy of this form of self-reporting method, and no such comments or claims can be made on reading their paper. Self-reporting has a number of advantages, including that it is relatively low cost [113]. Another advantage is that there are relatively few ethical concerns with this method of observational data collection compared to the other methods as no one need impose upon the privacy of the participants [56, 113]. Moreover, it is also the superior method in terms of COVID-security, as the participant need have no interaction with anyone else. However, problems with self-reporting do exist. For example, a participant experiencing disorientation or agitation participant may be incapable of recording observations [130]. Furthermore, those experiencing mild dementia are sometimes hesitant to admit to experiencing problems greater than what one would consider to be normal for an elderly individual [131, 132], meaning self-reporting skew observational data to only include difficulties which were undeniably experienced, meaning the record will not be complete. Finally, the cognitive impairment of participants may also make it more likely that they lose the medium for self-reporting [92, 93].

Observation of participants solely by caregivers is the observational method used in eight of the papers included [18, 57, 59, 61, 114, 128, 133, 134]. Two main categories of caregiver observation recording can be identified: methods which are paper-based and methods which are app-based. Caregivers using paper based methods will record their observations in a journal [10, 127] or observation chart [18]. In one study, the paper-based method used was a journal, however this was only used to record extra contextual information regarding the difficulty occurring, as the primary method of recording difficulties occurring was pressing an event marker button on the physiological collection device [10]. Paper-based recording is also used in another paper, where the caregiver of the participant recorded the sleep patterns of the participant in a paper-based sleep diary. The caregiver's recording of observational data was found to be highly accurate when a strong correlation between the diaries content and the collected physiological data was found. An observation chart printed upon paper is the medium utilised for recording observations by the caregiver in another paper [18], and this overcomes the difficulty of properly and concisely quantifying and recording dementia-related behaviours which occur inherently with the utilisation of free-form methods. The caregiver, using a 24-hour observation chart, marked down specific colours, with each colour corresponding to a different dementia-related difficulty being observed. This leads to a standardisation of observational data across participants, making it easy to compare one participant's results with another participant's results. However, one drawback is that the observer will record less contextual information regarding the difficulties than in a journal or diary.

Caregivers using an app to record their observations is the selected method in four papers [57-59, 61]. In the study referred to by the name BESI, caregivers record the time, place, and characteristics of any agitation episodes which they observe during the study. They do this by utilising an android app-based daily survey. The authors of the paper provide no information on the exact nature of the survey, rendering any evaluation of the app-based survey impossible [57-59]. Also utilising an Android app is another paper [61], in which the person observing the participants record dementia-related difficulties that occur by selecting the difficulty from a predetermined list. The use of a list of predetermined options works not only to standardise the observations across participants and observes, but also simplifies the process of recording for the observer. Another important thing to consider is that if the observer is a caregiver, is that caregiver formal or informal. Generally speaking, formal caregivers (FC) are caregivers who provide care to participants in a formal, professional capacity, while informal caregivers (IC) are ones who provide care in an informal manner, usually being family or friends of the person with dementia. FCs, who are the caregivers providing the observational data in three studies [10, 18, 61], are trained professionals and so will have a higher likelihood of communicating their observations using accepted medical terms, meaning their observations should be standardised and be understandable to a high degree [135]. Moreover, FCs are commonly available to observe participants for extended periods of time when the participant is based in institutional settings [52, 53]. Alternatively, ICs, who are the caregivers who provide the observational data recording services in multiple studies [10, 57-59], are more likely to be caregivers to participants with more mild dementia who are based in a home or community setting. This is the case in one study [127] where the caregiver doing the observations is the sister of the participant. This means that ICs are going to be commonly available to observe participants in experiments of longer durations, as long as the caregiver lives with the participant. However, ICs often experience increased burden and poor mental and physical health, resulting from their caregiving responsibilities, which may be increased and made worse by the additional responsibilities of being an observer for a data collection experiment [136, 137]. Moreover, the COVID-19 pandemic, and other future pandemics, may cause a restriction in the amount of time ICs and PCs can spend with participants [123].

Cameras are the observation method selected in two of the papers included [60, 126]. In a singular paper [126], researchers deploy a number of cameras in a mock-up of a waiting room, in which participants perform a variety of tasks. The video recordings made by the cameras are then later used to identify what activities the participants performed and for how long. Similarly, in another paper, cameras where used to record participants who exhibited PV [60]. In both cases, there was a high degree of correlation between the physiological data and the video recording, implying this method of observation is relatively accurate. Another advantage of utilising cameras for observational data collection is that the videos collected can be watch as many times as convenient, with observations being refined for greater accuracy with each viewing [55] Furthermore, passive nature of this camera-based method means that there is no increase in burden on participants or caregivers when it is used. However, the use of cameras and the recording of individuals is not without its ethical concerns- many people are uncomfortable being surveilled with cameras- thus meaning that cameras should be deployed only to shared spaces and not utilised in any private areas in which participants may perform sensitive, personal activities [116]. Another disadvantage of using camera for observation in studies is that they have a relatively high cost, compared with self-reporting and caregiver observations [113]. Caregivers and camera-based observation methods are used in combination in four papers [55, 99, 116, 117]. In the DAAD study [55, 116] caregivers used observational charts to record instances of dementia-related difficulties, highlighting contextual information which cannot necessarily be captured by a camera. In conjunction with this, cameras located in the communal spaces of the care home were used to record behaviours of participants, with the recorded video later being used to refine the initial caregiver observations. Another paper used a similar combination of caregivers and cameras [99]. Researchers used cameras to record the participants' behaviours while a FC recorded any relevant difficulties or activities they observed on a paper-based observation chart. The cameras were used in a communal room in the institution, making it vital for all users of said communal space to consent to being recorded in their daily lives as they use the space. One member of staff did not consent to being recorded with cameras, citing concerns relating to their privacy, and so it was decided not to deploy cameras [99]. Similar concerns relating to privacy are also expressed in another paper [117], in which cameras and caregivers are used in combination also. However, the authors state that in order to mitigate the privacy concerns expressed, they strictly limited access to the recordings that were made to 2 qualified individuals who were the only people required to view the footage for the needs of the study to be meet. However, if such reassurances are not enough to assuage the privacy concerns of participants, observers, or other stakeholders from whom consent is required, caregivers can still be used to obtain valuable observational data.

### 4.3.5. Data transfer & storage

Two distinct methods of data storage were identified in the literature. Seven papers utilised the device's internal memory to store the data as it is collected, with the data being later transferred off of the device. In three of these papers, the data is transferred via a wired connection to a computer, and in the others the same process happens, however the data is then further transferred onward, onto a cloud service. COVID-19 made the transfer of data more difficult, as extra steps were required to safely get the device from the participant, without an increased risk of COVID transmission [138, 139]. Another potential issue with transferring data, which exists beyond the constraints of a global pandemic, is the need of licenced software to transfer the data to a desktop device [109]. Automatic transfer of data via a wireless medium, such as Bluetooth or Wi-Fi, is a valid alternative, with the data being sent to edge computing devices or nodes, as with the BESI study [57-59, 140, 141]. In one study the data is transferred to Bluetooth "anchors" and then is transferred onwards from these via Wi-Fi to servers were they can be stored more permanently [125]. Similarly in another paper, the data is transferred to a local base station, within the study's immediate environs, from where it is then sent onwards to storage in the cloud [128]. Local storage of the data on the device requires little to no environmental or additional infrastructure, and can help to extend battery life, as no extra battery power is being used to operate wireless capabilities. However, researchers or caregivers will be left with the additional duties of transferring the data, increasing workload and burden. Alternatively, the workload for researchers and caregivers is reduced with wireless data transfer and is useful for continuous collection of data over an extended period. Despite these advantages, this approach does have its drawbacks, including more infrastructure requirements and increased complexity and cost [57-59, 140, 141]. Furthermore, cloud-based services can be helpful for when data to be stored is of large proportions [26]. However, use of such cloud-based resources is dependent upon the implementation of good security and access controls.

## 4.4. Proposed methodology

In this section, a data collection protocol is proposed, with the utilised methods being based on the results of the review of methods and methodologies in section 3 of this chapter.

### 4.4.1. Proposed novel data collection methodology

Recruitment for the proposed experiment shall be performed via the Liverpool NHS Memory Clinic, with an NHS doctor who is also an advisor on our project acting as a gatekeeper. The recruitment criteria will be that the potential subject shall be dwelling in a community setting i.e., their own home, and shall have had a very recent diagnosis of dementia- within the month proceeding their recruitment into the trial, with their dementia being in the mild stage. One could argue that a better approach is to recruit people via the community support and advocacy groups for community dwelling people with dementia, as attendees of such groups are almost certain to be community dwelling. However, the use of the Memory Clinic is best for the proposed experiment as it increases the likelihood we shall recruit people with recent diagnose, and furthermore, we already have contacts established with the memory clinic through the NHS doctor advising on our project.

Consent for the subject to participate will need to be acquired from both the subject and their next of kin/PoA. The reason for this is that the subject will just have been diagnosed with a disease that causes cognitive impairment, and thus it would likely require expert opinion and psychological testing to ensure that the person is capable of consenting. However, to simply have the next of kin/PoA consent could be equally distressing to the subject as they are likely to have only mild cognitive impairment at the early stage of dementia in which we wish to recruit them. As such, best practice will be to give both the next of kin/PoA the information about the procedures, duration, and other information about the experiment- including that they can withdraw their consent and participation at any time- both in written form and in verbal form. The former could be a leaflet which gives them the basic information about the experiment, and the latter should be a semi-formal verbal briefing in which both subject and caregiver can ask any questions they wish and seek whatever clarification they need. Then, the next of kin/PoA will be asked for informed, written consent, and the subject will be asked to give verbal consent, if they are deemed capable by an attending medical professional. Thus, both parties will have given informed consent where possible.

The caregiver will have been given an overview of the experiment in the consent acquisition process, however once consent is given and the subject recruited, their caregiver will be given more information on their specific duties and activities they will need to carry out. They will also be shown how to do them in a semi-formal training session in which they can ask questions and seek any necessary clarification. The caregiver will be required to remove and charge the device every evening as the subject goes to bed and reapply the device of a morning as the subject wakes up. They will also be required to collect observational data in an observation diary. The caregiver is likely to be an informal caregiver- a family member or friend the subject lives with- thus they will be unlikely to already know how to take down meaningful observational data. Thus, they shall receive training on how to do so. One could say that it would be better to utilise a formal caregiver who already possesses such skills, but such caregivers would increase the cost of the experiment and would be less likely to be available for prolonged periods of observation. Thus, training informal caregivers to make the observations is the preferable approach.

The subjects shall be monitored for a period of 30 days. Data collection shall take place during the waking hours of the subject, but while the subject is sleeping the device shall be removed to be charged, and so that the data collected that day can be transferred to a computer. As previously discussed, this means that no data could be collected on any difficulties that occur at night, however it allows for a more long-term recording of the subject's daytime activities. During the data collection period, the caregiver shall be asked to observe the subject and record any difficulties which occur in the subject in a paper journal. The record shall contain a time, a description of the difficulty and some contextual information such as what the subject was doing prior to the difficulty. It could be argued that the paper-based journal could be less efficient than an app-based recording system and less reliable and reviewable than a camera-based system. However, paper-based journals do not require charging as tablets or phones do, and paper-based journals have less privacy concerns inherent in their use than video cameras and tend to be cheaper and easier to implement.

The collected physiological data shall be initially recorded to the device during the day and then transferred to a computer or PC of the evening/night. The Empatica E4 does allow for the streaming of data right to the computer via Bluetooth, however this was considered to not be ideal, as it would mean having an external device running in the home of the subject, which would have to be powered by their mains electricity, which could incur them not inconsequential costs over the course of the experiment. As such, it is better that the data be stored locally on the device and then transferred to a computer that only needs to be used at the residence for the duration of the data transfer. Once transferred the data shall be stored locally and securely on the computer. The observational data will be recorded on paper, and these paper observations will be digitised by a researcher- converted to a word document or spreadsheet with the original kept in PDF form- and stored securely, in digital format on the computer. An alternative approach could be to keep the paper documents as paper documents, with the digitisation process presenting potential chances for information loss- a document could scan incorrectly or the researcher typing the observations into a word document could make a mistake. However, digitisation of the documents allows them to be stored more easily and search and used for labelling of the physiological data more efficiently.

Thus, a step-by-step methodology for the proposed experiment is as follows:

1. Recruit 5 subjects via the memory clinic through the gatekeeper.
2. Inform the potential subject and their next of kin/PoA of the experimental process and implications for them in written and verbal formats.
3. Acquire written and informed consent from the PoA and informed verbal consent from the subject if considered appropriate by attending medical professional.
4. Train the caregiver in managing the device and taking observational records of dementia-related difficulties using a paper-based journal.
5. The caregiver will attach the Empatica E4 (device) to the left wrist of the subject in the morning when the subject gets up, and ensure the device is on.
6. The caregiver will observe the subject throughout the subject's daily routine, recording in the journal any dementia-related difficulties the subject experiences.
7. The caregiver will remove the device from the subject immediately before the subject goes to bed.
8. Transfer the data from the device to the computer for storage.
9. Take the observations recorded for that day and give the caregiver a new journal for the next day.
10. Put the device on charge to ensure it is charged for the next day.
11. Repeat steps 5 to 10 for another 29 days.
12. Convert all the recorded observations into a digital format, to be stored on the computer storing the physiological data.

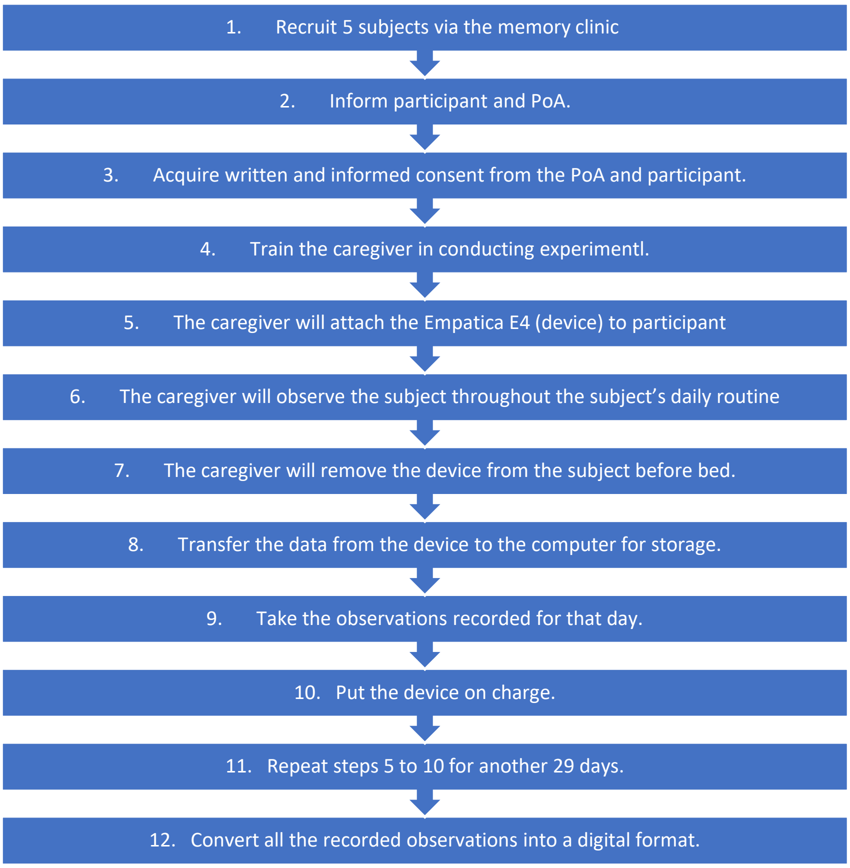



*Figure 2 Flow diagram of the proposed experimental methodology*

## 4.5. Pivot of research to work stress

In section 4.3. of this thesis, a variety of complications and limitations were found to exist in utilising many of the necessary methods for conducting the required data collection experiment with people with dementia. As such, the conducting of an experiment of that nature during the COVID-19 pandemic was found to be problematic and potentially unethical. In this section a justification for the pivot of the research project to focus on work stress detection from physiological data is provided. Furthermore, section 4.5.2. briefly outlines the aims and the proposed methodology of the new project direction, with the latter being discussed in much greater detail in Chapter 5 of the thesis. The pivot and justifications for it were first outlined in Harper et al. (2021,c) [14].

### 4.5.1. Justification

The COVID-19 pandemic resulted in a number of lockdowns and social distancing-related laws and guidance in many different regions [39]. Due to this, it became difficult to conduct experiments with people with dementia, as that demographic is one of the most at risk of death from the virus, and thus need to be given even greater and more strictly enforced protections against it [7] [39]. Furthermore, lockdowns and social distancing lead to an increase in feelings of stress and isolation across the locked-down populations. This meant that certain key stakeholders such as caregivers were likely to be negatively impacted by being part of an experiment due to increased responsibilities [27]. Therefore, it was considered potentially unsafe and unethical to conduct a data collection experiment with people with dementia during the pandemic. The need to collect such a dataset to progress the research project was discussed in depth in chapter 3 of this thesis, and as collecting the dataset became impractical within the time and budgetary constraints of the project, it was decided that a pivot to a different but related research topic was necessary to continue the work on the project.

The research domain selected was work stress detection from physiological data, and more specifically the differences in work stress experienced by individuals as they worked in home-based and office-based environments. The decision to pivot to work stress was based upon a few justifying notions. Firstly, both stress and dementia-related difficulties can be tracked with wearables which detect activations of the Autonomic Nervous System (ANS) to identify emotional arousals indicating negative affect [16, 35], so processes employed in both domains are similar. Indeed, the wearable devices employed, physiological attributes collected, and the machine learning models trained in the work stress domain and dementia-related difficulties domain have high degrees of similarity. Therefore, lessons learnt from developing stress detection models will be useful for detecting dementia-related difficulties. Another justification is that stress management can be highlighted as both a mitigation to and preventative measure of dementia-related difficulties [22]. For example, a person may become agitated due to stress that is unresolved, and long-term stress has been shown to speed up the progress of dementia, and thus lead to increased difficulties [50] [51]. Finally, the selection of the specific topic of comparing work stress experienced by individuals between working in home and office-based environments was supported by two arguments. One argument is that as the pandemic lead to an increase in individuals working from home, or from mixed home office environments, research on this topic would be timely and provide needed insight to work stress in the pandemic and post-pandemic working situations. Secondly, working age individuals tend to have lower risks of mortality from COVID-19 than people with dementia, and asking them to record their own activities and physiological attributes while they perform work tasks they had to perform regardless leads to no greater risk to those individuals of contracting COVID-19 [40]. A final justification which could also be provided is that conducting a physiological data collection experiment with working age people would allow the researchers to learn how to conduct such an experiment and work out any potential issues, for example, in regard to the operation of the devices, before performing similar experiments with similar equipment with more vulnerable individuals with whom the issues could result in more disruption, distress, or difficulties [41].

## 4.5. Chapter summary

This chapter aimed to present a novel data collection methodology which could be used to collect a dataset from people with dementia, with the aim of using it to develop a system which could predict the occurrences of dementia-related difficulties. To that end, a systematic literature review was conducted to identify the various methods that were used in previous similar studies existing in literature. 5 distinct stages or elements of the data protocols were identified: recruitment; consent and assent acquisition; physiological data collection; observational data collection; data transfer and storage. The selection of methods at each stage of the methodology depended mainly upon the stage of dementia or difficulty being studied. For example, studies including participants with severe dementia were more likely to use caregiver observation than studies of participants with mild dementia. In the conductance of the review, it was identified that many of the methods could not realistically be utilised in a COVID-secure manner, making conducting a data collection experiment with people with dementia potentially unsafe and unethical during the COVID-19 pandemic. As such, a pivot to detecting work stress is proposed, with the justifications that: the participants will be working age adults at lower risk of COVID-19 morbidity than people with dementia; the physiological indicators of dementia-related difficulties such as agitation and stress are similar; and stress can be a result of dementia-related difficulties as well as a cause of them.

# 5. Comparison of Work Stress in Home and Office Work Environments

## 5.1. Introduction

This chapter is structured thusly. Firstly, the methodology for attaining and analysing a physiological dataset from academics working in a mixture of home and office work environments is described. Included in this analysis is the development of machine learning models that can classify the occurrences of high and low, stress observations. Next, the results of the data analysis and machine learning are shown and then discussed, along with a discussion on the potentials mitigations to work stress, based on the causes implied be the results of the data analysis. Finally, the chapter is concluded with a summary of the results of the analysis and the potential mitigations.

## 5.2. Methodology

The methodology proposed in this chapter entails participants wearing a wrist-worn device which collects data related to the physiological indicators of stress, as they go about their average working day in home and office-based working environments. The aspects of the methodology outline in this section include the ethical approval of the study and the strategy used to recruit participants. Furthermore, the protocol used to collect the data from the participants is also described, as well as the methods used to pre-process and analyse the data. Finally, the methods used to train a variety of machine learning models that can be used to predict instances of work stress from physiological indicators of work stress are also discussed.

### 5.2.1. Ethical approval & recruitment

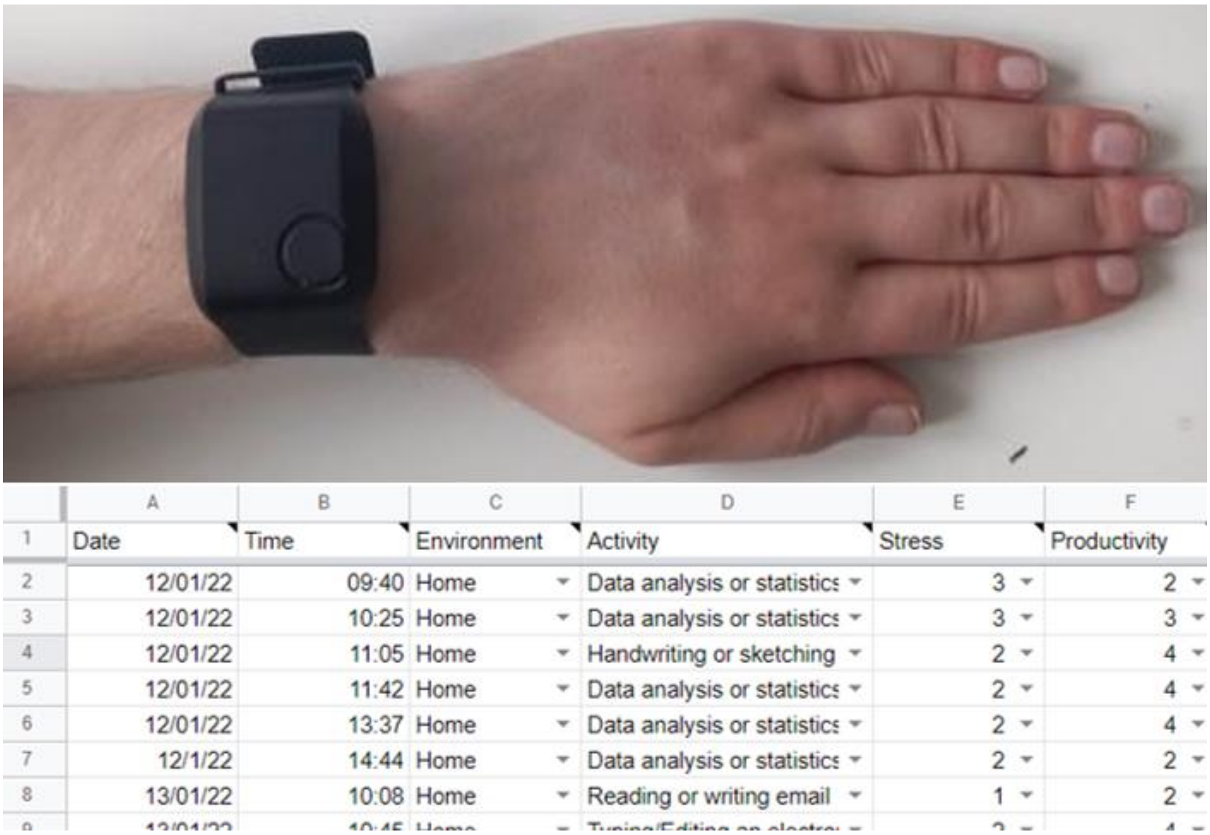
To collect the primary dataset of physiological stress indicators and subjective stress and productivity measures, an ethical approval is obtained from the Research Ethics Committee, Liverpool John Moores University (Reference number: 21/CMP/001). Participants are recruited with emails sent via internal university email groups. The invitation emails explain the study and what potential participants would be required to do if they took part and have a participant information sheet as an attachment. Emails are sent to academic staff and postgraduate research students at LJMU, with inclusion criteria being that participants be academic staff or postgraduate researchers and able to work from home and office settings for at least 3 days each over the course of the protocol.

### 5.2.2. Data collection protocol

The Empatica E4 smartwatch was chosen as the physiological data collection device. It was selected for comfort and accuracy of the physiological features collected [14]. The device collects skin temperature, electrodermal activity, heart rate, heart rate variability, blood volume pulse data, inter-beat interval (IBI) and movement/accelerometer data from the participants wrist.

The data collection protocol is largely based on a protocol outlined in earlier published work on the project [29]. The protocol is planned to last approximately 10 working days for each participant. Participants are asked to wear an Empatica E4 on their non-dominant wrist, with the aim of reducing movement artefacts in the physiological data streams as opposed to the dominant wrist, while they complete their daily work-related tasks. The participants are given a list of 10 activities (that included meeting, meeting preparation, typing/editing an electronic document, reading or writing email, handwriting or sketching by hand, researching/learning, data analysis or statistics, teaching (lecture), teaching (tutorial/lab session), lesson/lecture/tutorial preparation) and asked to turn the Empatica E4 on when they begin one of the listed activities and turn it off at the end of that activity. Many of the activities selected are similar to those selected in the Stress at Work (STRAW) study by Bolliger et al. (2020) [26], with emails, data analysis, and the planning and teaching task selections in the current study being inspired by Bolliger et al. (2020). Unlike Bolliger et al. (2020) the current study does not capture the participants’ travel from location to location, or free time and breaks, reducing potentially high participant burden reported in the STRAW study. Furthermore, the current study has more specific options for teaching or presenting, with teaching during the lectures being considered different from teaching in tutorials or lab session to warrant the two having separate selection options.

At the end of each work session, the participants complete a questionnaire. The questionnaire is a Google Sheets document which is shared with the participants via email. The document asks the participant to record six attributes regarding the activity they are recording: the date of the recording; the time of the recording; whether the participant is recording the session at home or in an office environment, from a drop-down selection menu; the activity being performed, from a drop-down selection menu; the subjective stress of the participant, on a Likert scale of 1 to 5, similar to Galanti et al (2021) [22]; and subjective productivity, also on a Likert scale of 1 to 5. In the first column the attribute name is provided, the second column shows all possible values that could be returned for each attribute, the third column shows the number of instances- qua the number of work sessions in which each value of each attribute was recorded- column 4 shows the percentage of all sessions in which each value of each attribute was recorded, and column 5 provides a description of each attribute and possible value of the attributes.



| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | Date | Time | Environment | Activity | Stress | Productivity |
| 2 | 12/01/22 | 09:40 | Home | Data analysis or statistics | 3 | 2 |
| 3 | 12/01/22 | 10:25 | Home | Data analysis or statistics | 3 | 3 |
| 4 | 12/01/22 | 11:05 | Home | Handwriting or sketching | 2 | 4 |
| 5 | 12/01/22 | 11:42 | Home | Data analysis or statistics | 2 | 4 |
| 6 | 12/01/22 | 13:37 | Home | Data analysis or statistics | 2 | 4 |
| 7 | 12/1/22 | 14:44 | Home | Data analysis or statistics | 2 | 2 |
| 8 | 13/01/22 | 10:08 | Home | Reading or writing email | 1 | 2 |

*Figure 3 (Top) An Empatica E4 worn upon a wrist. (Bottom) Screenshot of an example of the Google Sheet-based questionnaire for gathering the contextual/observational data*

### 5.2.3. Data preparation, pre-processing, and initial data analysis

Prior to the training of the machine learning algorithms, an initial data analysis was conducted in order to uncover any correlation between any of the physiological signals and questionnaire data points, proceeded by preparatory steps to clean and scale the data. Firstly, a visual inspection of the data was conducted to remove any missing data or artefacts. MATLAB 2020A was then used for the first stages of the initial data analysis. The naming convention of the folders containing the collected data files meant that all of the questionnaire data relating to each session was contained in the sessions folder name. As such, that is where the questionnaire data was loaded from. Each physiological data signal was then loaded from the relevant file and filtered were necessary. All outlier values were removed from the data attributes, with outliers being defined as any value that lay more than three scaled MAD from the median, and these outliers were replaced with the nearest non-outlier value in the attribute. Furthermore, all data attributes were smoothed using the

MATLAB “Smoothdata” function, reducing the influence of noise on the data analysis and machine learning models. The EDA signal was initially filtered to reduce noise in the signal using a 4th order Butterworth filter with a cut-off frequency of 2Hz. From the EDA signal the mean, standard deviation, dynamic range, mean slope, and the maximum and minimum values are extracted, as in Simons et al. (2020). From the tonic component the mean and standard deviation are extracted. The mean and standard deviation are also calculated from the phasic component [30]. Furthermore, an algorithm inspired by Healey (2000) is used on the phasic component to extract the number of skin conductance responses (SCRs), the sum of the magnitudes of the SCRs, and the sum of the durations of SCRs.

Descriptive statistical features were then calculated from each of the physiological data signals, including the separated tonic and phasic EDA components. The features extracted are shown a table provided in appendix B. Due to the varying length of the data collection sessions, the number of SCRs, SCR magnitude and SCR duration were all normalised using equation 1. This method of normalisation was to divide the number of peaks, magnitude, or duration by 60 and multiply the result by the sampling frequency, thus giving us the number of peaks/magnitude/duration of the SCRs per minute of each work session. This was to reduce bias, as longer sessions may have more SCRs, and thus greater SCR magnitude and duration, than a shorter session, yet at a lower frequency. As such, a shorter session with a high number of SCRs, likely a high stress session, could be erroneously considered the same as a much longer session with the same number of SCRs, which is more likely to be a low stress session. In equation 1, Xi is the normalised signal, X is the non-normalised signal, and fs is the sampling frequency at which the EDA data was collected.

$$Xi = \frac{X}{60}fs$$

*Equation 1 Normalisation equation for SCR magnitude and SCR duration for the initial data analysis*

Similarly, NN50 and NN25 were also normalised for the same reason, using the normalisation equation:

$$Xi = \frac{N}{length(X)/60}$$

*Equation 2 Normalisation equation for NN50 and NN25*

Again, in this equation, Xi is the normalised signal, X is the non-normalised signal, and N is the number of inter-beat intervals that meet the required criteria (differing more than 25 for NN25 and more than 50 for NN50). The reason for this method being employed is that the number of intervals varying by more or less than the required value will be more likely to be higher in longer sessions than shorter ones, regardless of whether the longer session is a high or low stress one. This will make it more likely that duration of a session will create a bias. As such, equation 2 addresses this bias by dividing the number of relevant inter-beat intervals by the number of values in the raw data stream, divided by 60, meaning we have the number of relevant interbeat intervals per minute. The correlations and dependences of the variables were then calculated using a variety of methods. The resulting dataset of combined questionnaire-based and physiological attributes and features were then imported to SPSS, in order that the relationships and correlations between the attributes could be calculated. The relationship between binary stress level and environment was calculated using the Phi Coefficient, as both variables are non-dichotomous nominal variables. The process for doing so was to first cross-tabulate the data creating a contingency table and then calculating the Phi value using the equation:

$$Phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

*Equation 3: Equation for calculating the Phi squared coefficient*

Where a is the count of low stress occurrences in home environments, b is the count of low stress occurrences in office environments, c is the count of high stress occurrences in home environments and d is the count of high stress occurrences in office environments. The correlation between the ordinal scale stress and all of the physiological variables was calculated using the Kendall Tau B coefficient, as this is the recommended coefficient for continuous to ordinal variable correlation calculations. The point-biserial coefficient was used to calculate the relationship between physiological data and binary stress, and the relationship between physiological data and environment. Goodman & Kruskal's Lambda coefficient was used to calculate the relationship between day of week and binary stress, time of day and binary stress, and activity and binary stress. T-tests were also performed to assess the relationship between stress reported in the home and office environments, both for all participants combined, and for each participant individually.

### 5.2.4. Feature extraction & machine learning

The physiological data is pre-processed to eliminate any anomalies and artefacts. All data streams are normalised using the normalization equation:

$$Xi = \frac{X - Xminimum}{Xmaximum - Xminimum}$$

*Equation 4 Normalisation equation for all physiological data attributes*

Where X is the non-normalized data and Xi Is the normalized data. The exception is the phasic EDA component which is normalized using the equation:

$$Xi = \frac{X}{Xmaximum}$$

*Equation 5 Normalisation equation for the phasic EDA component*

As was the normalization method for phasic EDA apply by Simons et al. (2020), and as described by Dawson et al. (2000) [17, 142]. Descriptive statistical features are extracted from each of the physiological data streams, using a sliding window of 60s length, with a 0.25 second overlap. A summary of all features extracted from the physiological and questionnaire data can be found in a table in appendix B. The features extracted from the skin temperature data were also based on the features extracted by authors in previous literature [17, 143]. These features were extracted using the built-in functions of MATLAB and were the mean skin temperature value (ST_MEAN), the maximum skin temperature value (ST_MAX), the minimum skin temperature value (ST_MIN), the median skin temperature value (ST_MED), variance of the skin temperature values (ST_STD), and the range of skin temperature values (HR_RANGE).

The features extracted from the heart rate data were also based on the features extracted by authors in previous literature [17, 29]. These features were extracted using the built-in functions of MATLAB and were the mean heart rate value (HR_MEAN), the maximum heart rate value (HR_MAX), the minimum heart rate value (HR_MIN), variance of the heart rate values (HR_STD), and the range of heart rate values (HR_RANGE). Similarly the features extracted from the blood volume pulse data were based on the features extracted by authors in previous literature [17, 143]. These features were extracted using the built-in functions of MATLAB and were the mean blood volume pulse value (BVP_MEAN), the maximum blood volume pulse value (BVP_MAX), the minimum blood volume pulse value (BVP_MIN), the median blood volume pulse value (BVP_MED), variance of the blood volume pulse values (BVP_STD).

Prior to normalization, the EDA signal is initially filtered to reduce noise in the signal using a 4th order Butterworth filter with a cut-off frequency of 2Hz [30]. The features extracted from the EDA signal were mean EDA (EDA_MEAN), the maximum EDA value (EDA_MAX), the minimum EDA value (EDA_MIN), the median EDA value (EDA_MED), the variance of the EDA data (EDA_STD), and the range in EDA values (EDA_RANGE). All these features were chosen as they were used to great effect in stress detection in Schmidt et al. (2016), and they were calculated using the relevant built-in MATLAB functions [17]. The raw EDA signal was then split into its phasic and tonic components, using an algorithm inspired by Dawson et al. (2000) [144]. From the tonic component the features extracted were the mean value of the tonic component (MEAN_TONIC), the variance of the tonic component (STD_TONIC), as these features were utilised in previous literature [17, 145]. From the phasic component, the number of skin conductance responses-normalised to be per minute of the session- (NUM_SCRs_PERMIN), the combined magnitude of skin conductance responses- normalised to be per minute of the session- (SCR_MAG_NORMALIZED), and the combined duration of skin conductance responses- normalised to be per minute of the session- (SCR_DUR_NORMALISED), using an algorithm inspired by Dawson et al. [144].

The accelerometer data consisted of three axes (X,Y,Z), and features are extracted from the each axis and from the sum of all the axis, as was done in previous literature [17, 142]. From each of the axes the mean value (ACC_X_MEAN, ACC_Y_ MEAN, ACC_Z_ MEAN, ACC_SUM_ MEAN), variance of the accelerometer values (ACC_X_STD, ACC_Y_ STD, ACC_Z_ STD, ACC_SUM_ STD), median value ((ACC_X_MED, ACC_Y_ MED, ACC_Z_ MED, ACC_SUM_ MED), the maximum values (ACC_X_MAX, ACC_Y_MAX, ACC_Z_MAX. ACC_SUM_MAX) and minimum values (ACC_X_MIN, ACC_Y_MIN, ACC_Z_MIN. ACC_SUM_MIN), and dynamic range (ACC_X_RANGE, ACC_Y_RANGE, ACC_Z_RANGE, ACC_SUM_RANGE).

From the IBI data, many of the features extracted were similar to those extracted from the forementioned attributes, namely the mean IBI value (IBI_MEAN), the maximum IBI value (IBI_MAX), the minimum IBI value (IBI_MIN) and the variance of the IBI data (IBI_STD)- all these features were extracted using built-in functions of MATLAB [142]. Also extracted was the root mean square of the IBI data (IBI_RMS), which was also extracted using built-in features of MATLAB. However, the number of interbeat intervals varying by more than 50ms (NN50) and the number of interbeat intervals varying by more than 25ms (NN25), both features extracted in previous literature, were calculated using a bespoke algorithm coded in MATLAB by the researcher [17]. The percentage of IBIs varying by more than 50ms (pNN50) and the percentage of IBIs varying by more than 25ms (pNN25) were calculated by taking the number of IBIs which met each criteria and dividing it by the number of IBIs in the relevant window [142].

MATLAB classification learner was used to train a variety of machine learning models. 3 models were chosen to be of particular interest as all 3 had achieved binary stress classification accuracies of over 99% in previous work on the project outlined in Harper et al. (2022) [15]. To select these models, a preliminary study was initially conducted using a subset of what would be the complete dataset (the first 5 participants). In this preliminary study, 28 different classification models were trained using the MATLAB Classification Learner toolbox, with the models being trained on all 71 features and a 10 cross-fold validation method being used to validate the models. The full list of models trained is: fine tree, medium tree, course tree, linear discriminant, quadratic discriminant, binary GLM logistic regression, efficient logistic regression, Efficient linear SVM, Gaussian Naive Bayes, Kernel Naïve Bayes, Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse Gaussian SVM, Fine KNN, Medium KNN, Coarse KNN, Cosine KNN, Cubic KNN, Weighted KNN, Boosted trees, bagged trees, subspace discriminant, subspace KNN, RUSBoosted trees. All models which achieved a validation accuracy of over 99% were considered to be worthy of note and were therefore chosen to be used in the later study with a more extensive dataset. These models were fine tree, bagged tree ensemble, boosted tree ensemble and RUSBoosted tree ensemble. The reason for only investigating the relatively small number of models which had been validated in the previous published study was to reduce the amount of time needed to train models (it is much quicker to train 4 or 5 models than 28), with time reduction being important due to time and financial constraints on the project. A bagged tree ensemble model was also trained to classify stress level and work environment combined. All trained models were evaluated using the 10-fold cross validation methods.

The hyperparameters for the fine tree, bagged tree ensemble, boosted tree ensemble, and RUSBoosted tree ensemble models were set as follows. For the fine tree, the maximum number of splits was set at 100 and the split criterion was Gini's diversity index. For the boosted tree model, the ensemble method was AdaBoost, the learner type was decision tree, the maximum number of splits was set to 20, the number of learners was 30, and the learning rate was 0.1. For the bagged tree model, the ensemble method was bag, the learner type was decision tree, the maximum number of splits was 11956 and the number of learners was 30. Finally, for the RUSBoosted tree model, the ensemble method was RUSBoost, the learner type was decision tree, the maximum number of splits was 20, the number of learners was 30, and the learning rate was 0.1. For all the models, the training/test split of the data was 80/20 (80% of the data was used for training and 20% of the data was used for testing).

A K-left out method was then utilised to test the machine learning models and reduce their subjectivity. This method worked by randomly selecting 3 participants- excluding participants 12 and 13 due to limitations with their data- whose data would be reserved as a testing dataset. This random selection was performed using a random number generation function in MATLAB, which randomly chose 3 numbers between 1 and 11. The number of 3 was chosen for participants to include in the training of the models as the data of 3 participants was approximately 20% of the data. The models are then retrained on the remaining 10 participants' data, and then the withheld test datasets are then used to evaluate how the models perform when confronted with unseen data.

Following the relatively low performance metrics for the trained model, two methods were chosen to attempt to investigate the reasons for the low metrics: principal component analysis (PCA) and Synthetic Minority Oversampling Technique (SMOTE). PCA was used to analyse the components of the dataset which most explained the variance in the dataset, and the features which had the greatest weighting or influence on the decision to classify each session as high or low stress. The features with the highest weighting were then used to re-train machine learning models. This reduction in dimensionality reduced the number of potentially unhelpful, undescriptive features. SMOTE was used to address the imbalance of the data, synthetically generating more instances of the minority class for binary stress level. A preliminary investigation was then performed with the balanced data to identify which machine learning models would be useful to train on the balanced data, with all 28 models in the MATLAB classification learner toolbox being trained on the new dataset. The bagged tree model was found to have the highest accuracy of the 4 models trained previously on the unbalanced dataset, and the logistic regression model was found to achieve similar accuracy to the bagged tree model. As such, these two models were selected for a more in-depth exploration and to be re-trained and their performance fully evaluated.

## 5.3. Results

### 5.3.1 Dataset & basic relationships

The collected proof of concept dataset contained over 11,000 minutes of data from 13 participants, with 295 work sessions recorded. 184 of the recorded work session were low stress (71.32%) and 74 sessions were high stress (28.68%). The median and modal Likert scale stress value were 2, the mean value was 2.095 and the maximum and minimum values were 5 and 1, respectively. The median and modal Likert scale productivity value were 3, the mean value was 3.0915 and the maximum and minimum values were 5 and 1, respectively.

*Table 6 Median, modal, mean, maximum and minimum subjective stress and productivity for the overall dataset*

| Attribute | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Stress | 2.00 | 2.00 | 2.10 | 5.00 | 1.00 |
| Productivity | 3.00 | 3.00 | 3.09 | 5.00 | 1.00 |

141 (47.8%) work sessions were recorded in the home environment, with a total duration of 5554.95minutes of data being recorded. The mean duration of home sessions was 39.4 minutes, the maximum and minimum durations were 197.55 and 8.33 minutes, respectively. 154 (52.2%) work sessions were recorded in the office environment with a total duration of 5233.63. The mean duration of office sessions was 41.35 minutes, the maximum and minimum durations were 154.93 and 6.13 minutes, respectively. The median and modal Likert stress value in both environments were 2, and the maximum and minimum values in home and office environments were 4 and 1, and 5 and 1, respectively. There was a difference in the mean Likert stress values between each environment, with the mean value for the home environment being 2.128 and the mean value in the office environment being 2.065. This could imply a slightly higher level of stress in the home environment, however there are certain limitations in the interpretation of mean values for ordinal variables. The lack in any major difference between the median, modal, maximum, minimum and mean stress values recorded in each environment implies no overall correlation or relationship between stress and environment.

*Table 7 Median, modal, mean, maximum, and minimum stress in the home and office environments*

| Environment | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Home | 2.00 | 2.00 | 2.13 | 4.00 | 1.00 |
| Office | 2.00 | 1.00 | 2.07 | 5.00 | 1.00 |

A two-tailed two-sample t-test with an alpha value of 0.05 was also performed to assess the differences in stress between the home and office environments (using built-in data analysis features of Microsoft Excel). The t Stat was 0.627 and the two-tail p-value was 0.531. These results support the already stated notion that there is very little relationship between stress experienced and work environment, as there are no statistically significant differences in stress experienced in the two environments.

The median and modal Likert productivity value in both environments were 3, and the maximum and minimum values in both environments were 5 and 1, respectively. The difference in mean value for Likert scale productivity has a relatively insignificant difference of approximately 0.0975, with the mean value for the home environment being 3.142 and the mean value in the office environment being 3.045. As was the case with stress, the lack in any major difference between the median, modal, maximum, minimum and mean productivity values recorded in each environment implies no overall correlation or relationship between stress and environment.

*Table 8 Median, modal, mean, maximum and minimum productivity in home and office environments*

| Environment | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Home | 3.00 | 3.00 | 3.14 | 5.00 | 1.00 |
| Office | 3.00 | 3.00 | 3.05 | 5.00 | 1.00 |

125 (42.37%) of the sessions were recorded in the morning, with a total duration of 4440.25 minutes of data recorded and the mean recorded duration being 35.52 minutes. 155 (52.54%) of the work sessions were recorded in the afternoon, with a total duration and mean duration of 6733.5 minutes and 43.4minutes, respectively. 15 (5.43%) of the work sessions were recorded in the evening, with total and mean durations of 748.6 minutes and 49.9 minutes, respectively. The median and modal Likert stress values for morning and afternoon sessions were 2, with the median and modal Likert stress value for evening session being higher at 3. Similarly, while the mean recorded stress for morning and afternoon sessions is similar at 2.04 in the morning and 2.10 in the afternoon, the mean recorded stress value in the evening was 2.53. However, the maximum value for the morning and afternoon sessions is higher than the evening, with the former 2 times of day having maximums of 4 or 5, while the evening has a maximum value of 3. This, coupled with the fact that the minimum stress value for morning and evening being 1 and the minimum value for evening being 2, implies that though working before 6pm had the potential to demonstrate higher levels of stress than working in the evening, evening sessions had consistently elevated stress levels than the other times of day.

*Table 9 Median, modal, mean, maximum, and minimum stress in morning, evening, and afternoon*

| Time Of Day | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Morning | 2.00 | 2.00 | 2.04 | 4.00 | 1.00 |
| Afternoon | 2.00 | 2.00 | 2.10 | 5.00 | 1.00 |
| Evening | 3.00 | 3.00 | 2.53 | 3.00 | 2.00 |

Recorded productivity appeared more consistent across the times of day, with the median value of productivity recorded at each time of day being 3. The modal productivity values for afternoon and evening session were also the same at 3, however modal productivity was higher in the morning at 4. The maximum and minimum values of productivity were 5 and 1, respectively, for morning and afternoon sessions, while the range of responses was smaller in the evening, with the maximum and minimum being 4 and 2, respectively. The biggest differences in productivity across the various times of day was in the mean values, with the highest being 3.36 in evening sessions, followed by 3.18 in morning sessions and 2.99 in afternoon sessions.

*Table 10 Median, modal, mean, maximum, and minimum productivity in morning, evening, and afternoon*

| Time Of Day | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Morning | 3.00 | 4.00 | 3.18 | 5.00 | 1.00 |
| Afternoon | 3.00 | 3.00 | 2.99 | 5.00 | 1.00 |
| Evening | 3.00 | 3.00 | 3.33 | 4.00 | 2.00 |

75 of the morning sessions were recorded in the home environment, with 20 being recorded as high stress and 55 being recorded as low stress. The mean, modal and mean stress value recorded in these sessions was 2, 2 and 2.08 respectively, with the maximum and minimum values being 4 and 1. 50 of the morning sessions were recorded in the office environment, with 15 being recorded as high stress and 35 being recorded as low stress. The mean, modal and mean stress value recorded in these sessions was 2, 2 and 1.98 respectively, with the maximum and minimum values being 3 and 1. This implies stress levels are slightly lower in sessions recorded in the office of a morning as opposed to at home in the morning. A similar trend is seen in the afternoon sessions, with an implication of slightly higher stress levels being recorded in the home environment at this time of day. 54 of the afternoon sessions were recorded in the home environment and 101 in the office environment. The median stress values in both environments are the same at 2, as is the maximum and minimum values at 4 and 1, respectively, however the modal and mean values are slightly lower in the office environment. The modal value is 2 in the home environment and 2 in the office environment, while the mean is 2.13 in the home environment and 2.08 in the office environment. This gives some support to the idea of marginally higher stress in home environments during this period of the day. However, the trend is reversed in the evening, with significantly higher stress being recorded in the office environment of an evening. 12 of the evening sessions were recorded in the home environment and 3 were recorded in the office environment. The median, modal and mean stress values recorded were all higher in the evening, as shown in table 11. The minimum value is also higher in the office environment at 3, as opposed to 2 in the home environment. This implies significantly higher levels of stress in office environments compared to home environments of an evening, however one limitation of this comparison is that there are only 3 sessions recorded in the

office environment in the evening, which is too small a sample size to make any truly conclusive conclusions.

*Table 11 Median, modal, mean, maximum, and minimum stress based on time of day in each working environment*

| Environment Time of Day | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Home Morning | 2.00 | 2.00 | 2.08 | 4.00 | 1.00 |
| Home Afternoon | 2.00 | 2.00 | 2.13 | 4.00 | 1.00 |
| Home Evening | 2.00 | 2.00 | 2.42 | 3.00 | 2.00 |
| Office Morning | 2.00 | 2.00 | 1.98 | 3.00 | 1.00 |
| Office Afternoon | 2.00 | 2.00 | 2.08 | 5.00 | 1.00 |
| Office Evening | 3.00 | 3.00 | 3.00 | 3.00 | 3.00 |

58 (19.66%) sessions were recorded on Monday (Total duration = 2087.8 minutes, mean duration = 36.00 minutes), 72 (24.41%) sessions were recorded on Tuesday (Total duration = 3056.23 minutes, mean duration = 42.45 minutes), 61 (20.68%) sessions were recorded on Wednesday (Total duration = 2579.18 minutes, mean duration = 42.28 minutes), 60 (20.34%) sessions were recorded on Thursday (Total duration = 2216.98 minutes, mean duration = 36.95 minutes), 36 (12.20%) sessions were recorded on Friday (Total duration = 1707.98 minutes, mean duration = 31.64 minutes), 5 (1.94%) sessions were recorded on Saturday (Total duration = 158.18 minutes, mean duration = 31.64 minutes), and 3 (1.16%) sessions were recorded on Sunday (Total duration = 116.10 minutes, mean duration = 38.69 minutes). The median stress for every day of the week was the same, equalling 2. The minimum and maximum values for most weekdays (Tuesday to Friday) were the same, being 4 and 1, respectively, and the minimum stress value was 1 on every day Monday to Friday. However, the minimum value increased to 2 for Saturday and Sunday, and the maximum fell to 3 during Saturday sessions and 2 for Sunday sessions. The modal values for all days were 2. Finally, the mean values showed the greatest difference across each of the days, with the lowest being recorded on Thursday (1.93) and the highest being recorded on Saturday (2.4). The median, modal, mean, maximum, and minimum values of stress for each day of the week are shown in table 12.

*Table 12 Median, modal, mean, maximum, and minimum stress on each day of the week*

| Day Of Week | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Monday | 2.00 | 2.00 | 2.12 | 5.00 | 1.00 |

| | | | | | |
|---|---|---|---|---|---|
| Tuesday | 2.00 | 2.00 | 2.08 | 4.00 | 1.00 |
| Wednesday | 2.00 | 2.00 | 2.10 | 4.00 | 1.00 |
| Thursday | 2.00 | 2.00 | 1.93 | 4.00 | 1.00 |
| Friday | 2.00 | 2.00 | 2.31 | 4.00 | 1.00 |
| Saturday | 2.00 | 2.00 | 2.40 | 3.00 | 2.00 |
| Sunday | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 |

8 of the Monday sessions were recorded in the home environment and 50 were recorded in the office environment. The median value of stress in both environments is 2, and the minimum value is also the same at 1. However, modal stress (Home = 1, Office = 2), mean stress (Home = 1.88, Office = 2.169), and maximum stress (Home = 3, Office = 5) are all higher in the office environment. 36 of the Tuesday sessions were recorded in the home environment, while 36 were recorded in the office environment. Again, the median stress in both environments is the same (Home = 2, Office = 2), as is the minimum stress (Home =1, Office = 1). However, modal stress (Home = 1, Office = 2), and mean stress (Home = 1.86, Office = 2.31) are greater in the office environment, implying higher levels of stress in the office environment on Tuesday. 39 of the Wednesday sessions were recorded in the home environment and 22 in the office. Median and modal stress both equal 2 in each environment. Mean stress (Home = 1.87, Office = 2.5), maximum stress (Home = 3, Office = 4), and minimum stress (Home = 1, Office = 2) are all higher in the office environment. 27 of the Thursday sessions were recorded in the home environment while 33 were recorded in the office environment. All of median (Home = 2, Office = 1), modal (Home = 2, Office = 1), mean (Home = 2.52, Office = 1.45), maximum (Home = 4, Office = 3) were higher in the home environment than the office environment, while minimum stress was the same in both environments, equalling 1. Similarly, the Friday sessions showed the same trend, with the median (Home = 2, Office = 1), modal (Home = 2, Office = 1), mean (Home = 2.57, Office = 1.85) were higher in the home environment than the office environment, while maximum and minimum stress were the same in both environments, being 4 and 1, respectively. Saturday and Sunday sessions were only recorded in the home environment as the University campus on which the offices were located was closed on inaccessible on these days. As such, though the statistics regarding stress on these days can be found in table 13, they will not be discussed in great detail here. Overall, it appears that stress is higher working from the office as opposed to home earlier in the week (Monday to Wednesday), and then the trend is reversed and the office becomes the environment with the lower stress on Thursday and Friday.

*Table 13 Median, modal, mean, maximum, and minimum stress on each day of the week in each working environment*

| Environment Day of Week | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Home Monday | 2.00 | 1.00 | 1.88 | 3.00 | 1.00 |
| Office Monday | 2.00 | 2.00 | 2.16 | 5.00 | 1.00 |
| Home Tuesday | 2.00 | 1.00 | 1.86 | 4.00 | 1.00 |
| Office Tuesday | 2.00 | 2.00 | 2.31 | 4.00 | 1.00 |
| Home Wednesday | 2.00 | 2.00 | 1.87 | 3.00 | 1.00 |
| Office Wednesday | 2.00 | 2.00 | 2.50 | 4.00 | 2.00 |
| Home Thursday | 2.00 | 2.00 | 2.52 | 4.00 | 1.00 |
| Office Thursday | 1.00 | 1.00 | 1.45 | 3.00 | 1.00 |
| Home Friday | 2.00 | 2.00 | 2.57 | 4.00 | 1.00 |
| Office Friday | 1.00 | 1.00 | 1.85 | 4.00 | 1.00 |
| Home Saturday | 2.00 | 2.00 | 2.40 | 3.00 | 2.00 |
| Home Sunday | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 |

The most commonly recorded activity in the dataset was “Typing/Editing an electronic document” (Sessions = 72, Total duration = 3222.4 minutes), with “Data analysis and statistics” being a close second (Session = 62, Total duration = 2634.9 minutes). The least commonly recorded activity was “Handwriting or sketching by hand” which was only recorded thrice (Total duration = 66.08 minutes), and all three times were in the home environment. The activity which had the highest recorded stress appears to be “Data analysis or statistics” having the highest median, modal, mean stress value, and joint highest maximum, and minimum stress values. The activity with the lowest recorded stress is “Handwriting or sketching by hand”, with the lowest median, modal, mean, and maximum stress, and joint lowest minimum stress. The median, modal, mean, maximum and minimum stress for all activity can be found below in table 14.

*Table 14 Median, modal, mean, maximum, and minimum stress for each activity in the dataset*

| Activity | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Meeting | 2.00 | 2.00 | 2.00 | 4.00 | 1.00 |
| Meeting Preparation | 1.00 | 1.00 | 1.78 | 4.00 | 1.00 |

| | | | | | |
|---|---|---|---|---|---|
| Typing/Editing an electronic document | 2.00 | 2.00 | 2.10 | 5.00 | 1.00 |
| Reading or writing email | 2.00 | 2.00 | 2.17 | 4.00 | 1.00 |
| Handwriting or sketching by hand | 1.00 | 1.00 | 1.33 | 2.00 | 1.00 |
| Researching/learning | 2.00 | 2.00 | 2.08 | 4.00 | 1.00 |
| Data analysis or statistics | 3.00 | 3.00 | 2.50 | 4.00 | 1.00 |
| Teaching (lecture) | 2.00 | 3.00 | 2.00 | 3.00 | 1.00 |
| Teaching (Tutorial/lab session) | 1.00 | 1.00 | 1.50 | 4.00 | 1.00 |
| Lesson/Lecture/Tutorial Preparation | 2.00 | 2.00 | 1.83 | 4.00 | 1.00 |

Table 15 shows the median, modal, mean, maximum and minimum stress values for each activity in both the home and office environments. N/A is used in the table where insufficient numbers of sessions of a certain activity performed in a certain environment were present in the dataset to calculate the relevant measure of central tendency.

*Table 15 Median, modal, mean, maximum, and minimum stress for each activity in each working environment*

| | Median | Modal | Mean | Maximum | Minimum |
|---|---|---|---|---|---|
| Meeting Home | 2.00 | 2.00 | 2.46 | 4.00 | 1.00 |
| Meeting Office | 2.00 | 2.00 | 1.74 | 3.00 | 1.00 |
| Meeting Preparation Home | 1.50 | N/A | 1.50 | 2.00 | 1.00 |
| Meeting Preparation Office | 1.00 | 1.00 | 1.86 | 4.00 | 1.00 |
| Editing Document Home | 2.00 | 2.00 | 1.90 | 4.00 | 1.00 |
| Editing Document Office | 2.00 | 3.00 | 2.33 | 5.00 | 1.00 |

| | | | | | |
|---|---|---|---|---|---|
| Emails Home | 2.00 | 2.00 | 2.11 | 4.00 | 1.00 |
| Emails Office | 2.00 | 2.00 | 2.22 | 4.00 | 1.00 |
| Handwriting Home | 1.00 | 1.00 | 1.33 | 2.00 | 1.00 |
| Handwriting Office | N/A | N/A | N/A | 0.00 | 0.00 |
| Researching/Learning Home | N/A | N/A | N/A | 0.00 | 0.00 |
| Researching/Learning Office | N/A | N/A | N/A | 0.00 | 0.00 |
| Data Analysis Home | 2.00 | 2.00 | 2.43 | 4.00 | 1.00 |
| Data Analysis Office | 3.00 | 3.00 | 2.59 | 4.00 | 1.00 |
| Teaching Lecture Home | N/A | N/A | N/A | 0.00 | 0.00 |
| Teaching Lecture Office | 2.00 | 3.00 | 2.00 | 3.00 | 1.00 |
| Teaching Lab Home | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 |
| Teaching Lab Office | 1.00 | 1.00 | 1.44 | 4.00 | 1.00 |
| Lesson Preparation Home | 2.00 | 2.00 | 1.91 | 4.00 | 1.00 |
| Lesson Preparation Office | 2.00 | 2.00 | 1.71 | 3.00 | 1.00 |

The most commonly recorded activity in the home environment was “Data analysis and statistics”, with 35 sessions in the home environment and 27 recorded in the office environment. Stress appears to be greater in the office environment for this activity, as the median, modal and mean stress are higher in the office environment than the home environment, while maximum and minimal stress are the same in both environments. However, the opposite is true for the activity “Meeting”, for which mean and maximum stress are higher in the home environment than the office environment. A less clear but similar trend is shown in the “Researching/Learning” activity, with mean and maximum stress being higher in the home environment than the office environment. Overall, there is no activity for which median, modal, mean, maximum, and minimum stress are simultaneously equal in the office and home environments, and thus one could argue there is reason to think that there could be a relationship between stress and environment when accounting for the work activity being completed.

When looking at the relationship between stress and environment and productivity and environment for individual participant one sees that there are significant differences in how the work environment relates to each individual participants experience of stress and productivity. Participant 1 shows a slight difference in stress levels between the home and office environment, with mean and maximum stress being higher in the office environment.

*Table 16 Median, modal, mean, maximum, and minimum stress experienced by participant 1 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 2.00 | 2.00 | 2.00 | 3.00 | 1.00 |
| Office | 2.00 | 2.00 | 2.35 | 4.00 | 1.00 |

A t-test was also performed to assess the difference in stress experienced in each work environment by participant 1 alone. The t Stat resulting from the analysis was -1.761 with a two-tailed p-value of 0.086, which is above the value set for alpha, meaning the difference does not reach the set definition for statistical significance. However, the one-tail p-value is below 0.05, implying some potentially statistically significant variations in stress between the 2 work environments for participant 1.

In terms of productivity there is also a slight difference between the home and office environments for participant 1, with modal productivity in the office environment being 2 greater than in the home environment, and mean productivity also being higher in the office environment.

*Table 17 Median, modal, mean, maximum, and minimum productivity experienced by participant 1 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 2.00 | 2.77 | 4.00 | 1.00 |
| Office | 3.00 | 4.00 | 3.17 | 4.00 | 1.00 |

For participant 2, there is a different trend, with median, mean, and minimum stress all being higher in the home environment than in the office environment.

*Table 18 Median, modal, mean, maximum, and minimum stress experienced by participant 2 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 2.00 | 2.80 | 4.00 | 2.00 |
| Office | 2.00 | 2.00 | 2.53 | 4.00 | 1.00 |

A t-test was also performed to assess the difference in stress experienced in each work environment by participant 2 alone. The resultant t Stat was 1.067, with the two-tail p-value being 0.29, implying no statistically significant difference in recorded stress for participant 2 in each environment.

In terms of productivity, the relationship between productivity and the environment is less clear, as while median and modal productivity is higher in the office environment, mean and maximum productivity is higher in the home environment. It is also of note that the mean productivity is slightly higher in the office than in home environment, but the difference is relatively small.

*Table 19 Median, modal, mean, maximum, and minimum productivity experienced by participant 2 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 3.00 | 3.40 | 5.00 | 1.00 |
| Office | 4.00 | 4.00 | 3.47 | 4.00 | 1.00 |

Participant 3 shows a much clearer relationship between stress and environment, with median and modal stress being 2 greater in the home environment than the office environment and mean and minimum stress being at least 1 greater in the home than office environment.

*Table 20 Median, modal, mean, maximum, and minimum stress experienced by participant 3 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 3.00 | 2.60 | 3.00 | 2.00 |
| Office | 1.00 | 1.00 | 1.58 | 3.00 | 1.00 |

A t-test was also performed to assess the difference in stress experienced in each work environment by participant 3 alone. The resultant t Stat supports the idea of participant 3 experiencing considerably greater stress in one environment, with the t Stat equalling 3.032 and the two-tailed p-value being 0.011, implying a statistically significant difference between the stress experienced by participant 3 in each of the work environments.

Productivity is also greater in the home environment than the office environment for participant 3, with median, modal, mean, and minimum productivity all being higher in the home than office environment.

*Table 21Median, modal, mean, maximum, and minimum productivity experienced by participant 3 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 4.00 | 4.00 | 3.60 | 4.00 | 3.00 |
| Office | 3.00 | 3.00 | 2.50 | 4.00 | 1.00 |

Participant 4 shows no significant difference in stress experienced in each environment, with a slightly higher mean and minimum stress in the office environment compared to the home environment. This is supported both by the measures of central tendency shown in table 22, but also by the results of the t-test, with the t Stat of -0.499 and a two-tailed p-value of 0.632 implying no statistically significant differences in stress between the two environments for participant 4.

*Table 22 Median, modal, mean, maximum, and minimum stress experienced by participant 4 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 2.00 | 2.00 | 2.18 | 3.00 | 1.00 |
| Office | 2.00 | 2.00 | 2.33 | 3.00 | 2.00 |

For productivity, there is also no clear relationship between participant 4's environment and how productive they were. While median and modal stress are both higher in the office environment, mean and minimum stress are higher in the home environment.

*Table 23Median, modal, mean, maximum, and minimum productivity experienced by participant 4 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |

| | | | | | |
|---|---|---|---|---|---|
| Home | 3.00 | 3.00 | 3.45 | 4.00 | 3.00 |
| Office | 3.50 | 4.00 | 3.17 | 4.00 | 2.00 |

Participant 5 also demonstrated no significant differences in stress experienced in each environment, with mean stress being less than 0.2 higher in the home environment and no other metric differing in each environment. This was supported by the results of the t-test, with the t Stat of 0.869 and a two-tailed p-value of 0.391 implying no statistically significant differences in stress experienced by participant 5 between work environments.

*Table 24 Median, modal, mean, maximum, and minimum stress experienced by participant 5 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 1.00 | 1.00 | 1.39 | 3.00 | 1.00 |
| Office | 1.00 | 1.00 | 1.21 | 3.00 | 1.00 |

Participant 5 had comparable results for productivity, with the only difference in metrics between the home and office environment being a slightly higher mean productivity in the home environment.

*Table 25Median, modal, mean, maximum, and minimum productivity experienced by participant 5 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 3.00 | 2.83 | 4.00 | 2.00 |
| Office | 3.00 | 3.00 | 3.11 | 4.00 | 2.00 |

Participant 6 however, does show some difference in stress between the home and office environments, with median, mean, and minimum stress all being higher in the home environment. However, the t-test does not show any statistically significant differences in stress experienced by participant 6 in each environment. This trend is similarly seen in regard to productivity, with median, modal, and mean productivity all being higher in the home environment as opposed to the office environment.

*Table 26 Median, modal, mean, maximum, and minimum productivity experienced by participant 6 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 |

| | | | | | |
|---|---|---|---|---|---|
| Office | 3.00 | 3.00 | 3.00 | 4.00 | 2.00 |

Participant 7 likewise shows an increased level of stress in the home environment as opposed to the office environment, with median, modal, and mean stress being higher in the home environment. However, maximum stress is higher in the office environment and there is no statistically significant difference in stress demonstrated by the t-test, with the t Stat being 0.59 and the two-tailed p-value being 0.56.

*Table 27 Median, modal, mean, maximum, and minimum stress experienced by participant 7 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 2.00 | 2.00 | 1.59 | 2.00 | 1.00 |
| Office | 1.00 | 1.00 | 1.47 | 3.00 | 1.00 |

In terms of productivity, participant 7 shows a slightly higher level of productivity in the office than in the home environment, with median and mean stress both being higher in the office environment.

*Table 28 Median, modal, mean, maximum, and minimum productivity experienced by participant 7 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 4.00 | 3.41 | 5.00 | 2.00 |
| Office | 4.00 | 4.00 | 3.60 | 5.00 | 2.00 |

Participant 8 shows higher levels of stress in the office environment than in the home environment, with median, modal, and mean stress all being higher in the office. However, the t test shows no significant differences in the stress recorded in each environment for participant 7, with the t Stat equalling -1.125 and the two-tailed p-value being 0.272.

*Table 29 Median, modal, mean, maximum, and minimum stress experienced by participant 8 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 2.00 | 2.00 | 2.46 | 4.00 | 2.00 |
| Office | 3.00 | 3.00 | 2.75 | 4.00 | 2.00 |

In contrast, participant 8's productivity is higher in the home environment, with the mean and minimum productivity being higher there than in the office.

*Table 30 Median, modal, mean, maximum, and minimum productivity experienced by participant 8 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 3.00 | 3.15 | 4.00 | 2.00 |
| Office | 3.00 | 3.00 | 2.92 | 4.00 | 1.00 |

Participant 9 experienced higher median, modal, mean, and maximum stress in the home environment than the office environment. However, the t-test shows no significant differences in recorded stress across the environments for participant 9, with the t Stat being 1.427 and the two-tailed p-value being 0.203.

*Table 31 Median, modal, mean, maximum, and minimum stress experienced by participant 9 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 2.00 | 3.00 | 2.00 | 3.00 | 1.00 |
| Office | 1.00 | 1.00 | 1.33 | 2.00 | 1.00 |

However, participant 9 also experienced higher productivity in the office environment, with median, modal, mean, and minimum productivity all being higher in the office environment.

*Table 32 Median, modal, mean, maximum, and minimum productivity experienced by participant 9 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 2.50 | 2.00 | 2.88 | 4.00 | 2.00 |
| Office | 3.00 | 3.00 | 3.00 | 3.00 | 3.00 |

Participant 10 shows no significant difference in stress between each environment, with a slightly higher mean stress in the home environment being the only indication of a difference in stress being experienced in each environment. This lack of relationship between environment and stress is supported once again by the t-test, with the t Stat of 0.352 and the two-tailed p-value of 0.732 implying no significant differences in stress experienced between the two working environments.

*Table 33 Median, modal, mean, maximum, and minimum stress experienced by participant 10 in the home and office work environments*

| Median, Modal, Maximum and Minimum Stress in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 2.00 | 2.00 | 2.63 | 4.00 | 2.00 |
| Office | 2.00 | 2.00 | 2.50 | 4.00 | 2.00 |

Alternatively, there does appear to be some difference in the levels of productivity experienced in each environment by participant 10. Mean and minimum productivity are both higher in the home environment as opposed to the office environment.

*Table 34 Median, modal, mean, maximum, and minimum productivity experienced by participant 10 in the home and office work environments*

| Median, Modal, Maximum and Minimum Productivity in each environment | | | | | |
|---|---|---|---|---|---|
| Environment | Median | Modal | Mean | Maximum | Minimum |
| Home | 3.00 | 3.00 | 3.38 | 5.00 | 2.00 |
| Office | 3 | 3 | 2.56 | 5.00 | 1.00 |

Participants 11, 12, and 13, were considered not to have enough sessions recorded on their own to be able to properly compare the relationship between environment and stress and productivity, and so analysis of those 3 participants on their own is not provided here. The main point to take away from the analysis of the stress and productivity experienced by each individual participant in each environment is that each individual participant experiences stress and productivity differently depending upon their work environment. For some people, such as participant 3, stress and productivity are both clearly higher in the home environment as opposed to the office. For some people, such as participant 7, there is a slightly higher level of stress at home, while there is a slightly higher level of productivity in the office. Furthermore, for people such as participant 5, there is no clear relationship between their work environment and their experiences of stress and productivity. This indicates that stress and the influence of work environment upon it is highly dependent upon the individual worker and their individual circumstances.

### 5.3.2. Correlations & relationships

In the proof-of-concept dataset of 13 participants, no strong correlation was found between stress and productivity using Kendall's Tau B (Tb), however there was correlation between these attributes for some individual participants. The Kendall Tau b result for the entire dataset was Tb = -.034, P value = .496, implying no significant correlation. However, the result of participant 2 (Tb = -.462, P value = <.001) implies a medium negative correlation between the attributes. Alternatively, the result for participant 5 (Tb = .306, P value = .046) shows a medium positive correlation between stress and productivity. This appears to imply that the relationship between stress and productivity is dependent upon the individual in some circumstances, which may be caused by interpersonal preferences and biology, as shall be explored further in the Discussion section of this chapter.

Furthermore, no significant correlation was found between binary stress level and environment in the overall dataset (Chi = .153, P Value = .696). A significant correlation was however found between binary stress and environment for participant 1's data (shown in figure 4). Similarly, using the Goodman and Kruskal's Lambda (L) no significant correlation was found between binary stress level and time of day (L = .004 p-value = .796), day of week (L = 0, P value = N/A), or activity (L = .042, p value = .330).

| | Value | df | Asymptotic Significance (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | 260.652[a] | 4 | <.001 |
| Likelihood Ratio | 261.615 | 4 | <.001 |
| N of Valid Cases | 244 | | |

a. 4 cells (44.4%) have expected count less than 5. The minimum expected count is 1.51.

*Figure 4 Chi-square results for relationship between stress and environment for Participant 1's data*

Moreover, no significant correlation between stress and session duration was found using Kendall's Tau B (Tb), with Tb equalling -.042 and P-value being equal to .342. This lack of significant correlation between stress and session duration is further supported by the results of participant 1 (Tb = -.004, P value = .974), participant 2 (Tb = -.013, P value = .916), participant 3 (Tb = -.026 ,P value = .894), participant 4, (Tb = .168, P value = .405), participant 5 (Tb = .125, P value = .346), participant 6 (Tb = .007, P value = .972), participant 7 (Tb = .005,P value = .970), participant 8 (Tb = -.112,P value = .482), and participant 9 (Tb = -.324,P value = .207). As such, it is probable that there is no correlation to be found between stress and session duration, thus the amount of time participants spent doing an activity is unlikely to affect the levels of stress experienced while doing the activity.

There are a number of significant correlations between the ordinal stress attribute and the extracted physiological data features. The physiological features related to electrodermal activity that had a significant correlation with stress were: mean EDA (Tb = .141, p value = .001), minimum EDA (Tb = .194, p value = <.001), median EDA (Tb = .15, p value = <.001), and mean tonic (Tb = .141, p value = .001). The physiological features relating to heart rate which significantly correlate with stress are: Mean HR (Tb = -.123, P-value = .006), HR maximum (Tb = =.193, p value = <.001), HR variance (Tb = -.165, p value = <.001), and HR range (Tb = -.178, p value = <.001). The physiological features relating to blood volume pulse are: BVP maximum (Tb = -.122, p value = .006 and BVP variance (Tb = -.124, p value = .005). Furthermore, the accelerometer-related features which are correlated to stress are: accelerometer y axis mean (Tb = -.127, p value = .004), accelerometer y axis max (Tb = -.127, p-value = .004), and Accelerometer y axis median (Tb = -.117, p value = .009). Finally, the features extracted from IBI which correlate significantly with stress are: NN50, NN25, pNN50 and pNN25 (Tb = .131, p value = .004).

In the home environment, the features extracted from EDA which correlate significantly with stress are: EDA mean (Tb = .205, p value = .002), EDA Maximum (Tb = .199, P-value = .002), EDA minimum (Tb = .246, p-value = <.001) EDA median (Tb = .207, p value = .001), mean tonic (Tb = .205, p value = .002), Only one other feature correlated with stress in the home environment, namely maximum HR (Tb = -.181, p-value = .005) In the office environment, the list of extracted features which correlate with stress is: minimum EDA (Tb =.164, p value =.001), maximum HR (Tb = -.2065, p value = <.001), HR variance (Tb = -.214, p value = <.001) HR range (Tb = -.229, p value = <.001), maximum BVP (Tb = -.186, p value = .003), minimum BVP (Tb =.167, p value = .007), BVP variance (Tb = -.210, p value = <.001), Accelerometer Y-axis mean (Tb = -.164, p value =.008), Accelerator Y axis median (Tb = -.178, p value =.004), NN50 and pNN50 (Tb = -.205, p value =.001), and pNN50 andpNN25 (Tb =.204, p value = .001).

### 5.3.3. Machine learning

In Harper et al. (2022), the highest machine learning prediction accuracies (>99%) were achieved using the fine tree, bagged tree, and boosted tree models, and thus these models were trained on the complete dataset. An RUSBoosted tree model was also trained, as the dataset was class imbalanced (considerably more low stress than high stress instances, and RUSBoosted tree models have been shown to be effective at classifying class-imbalanced data [146]. All models were evaluated using 10-fold cross validation to calculate the accuracy scores, with 20% of the data being withheld to be used as a test set.

The model in the current study which achieved the highest prediction accuracy was the Bagged Tree Ensemble model, with an accuracy of 99.3%, which is 0. 6 lower than the accuracy achieved by the same model in Harper et al. (2022). 113 misclassification instances occurred, with 80 of those false predictions being high stressed sessions being identified as low stress, and 33 being low stress sessions being misidentified as high stress.

The other models from this study achieved accuracies lower than those achieved for those models in Harper et al. (2022). Trained on the complete dataset, the fine tree model achieved an accuracy of 96.4%, a precision of 0.953, a recall of 0.922, and an F-Score of 0.937. The model had lower misclassification with regards to low stress observation than high class observations. The boosted tree ensemble method achieved only 95% accuracy, an approximately 4% reduction from the fore mentioned previously published work. The low accuracy appears to be the result of relatively high instances of misclassification of high stress observations as low stress.

The accuracy, precision, recall, F-score, area under the curve (AUC) for the ROC curve, PPV for high stress observations, and PPV for low stress observations are all shown the table 35.

*Table 35 The accuracy, precision, recall, F-score, area under the curve (AUC) for the ROC curve, PPV for high stress observations, and PPV for low stress observations for each of the trained machine learning models*

| Model | Accuracy | Precision | Recall | F-score | AUC | PPV High Stress | PPV Low Stress |
|---|---|---|---|---|---|---|---|
| **Bagged tree** | **0.993** | **0.992** | **0.982** | **0.987** | **0.999** | **0.992** | **0.993** |
| Fine tree | 0.964 | 0.953 | 0.922 | 0.937 | 0.979 | 0.953 | 0.969 |
| Boosted tree | 0.950 | 0.981 | 0.842 | 0.907 | 0.992 | 0.982 | 0.939 |
| RUSBoosted trees | 0.874 | 0.746 | 0.858 | 0.798 | 0.948 | 0.746 | 0.938 |

As a further proof of concept, another bagged tree ensemble model was trained using the data of participants 1 to 9 which could predict the work environment and binary stress level, and the accuracy achieved was 99.6%. The PPV for home high stress observations was 99.9%, the PPV for home high stress observations was 99.6%, PPV for office high stress observations was 99.9%, PPV for office low stress observations was 99.5%. The area under the curve was 1.00 and the confusion matrix for this model is shown in figure 5.

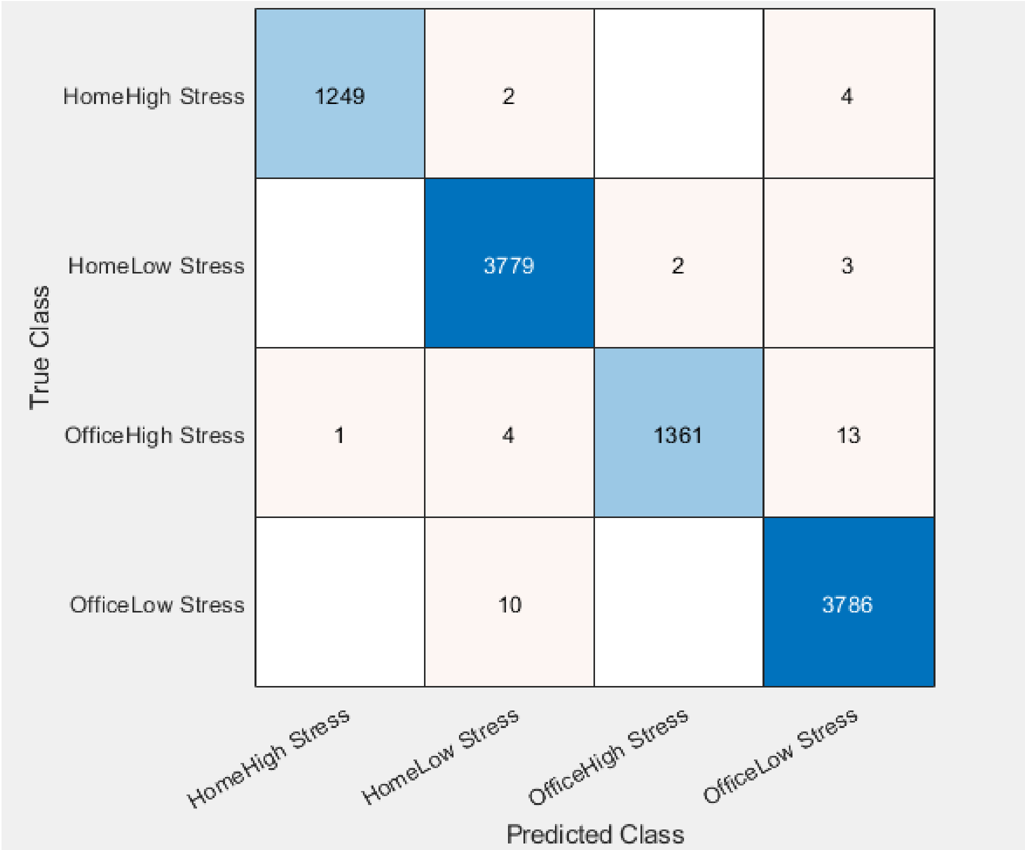



*Figure 5 Confusion matrix for the bagged tree ensemble model for classifying combined environment and stress levels*

A K-left out method is then utilised to evaluate the models and reduce their subjectivity. This method worked by randomly selecting three participants- excluding participants 12 and 13 due to limitations with their data- whose data would be reserved as a testing dataset. This random selection was performed using a random number generation function in MATLAB, which randomly chose 3 numbers between 1 and 11. In this instance the three randomly chosen participants whose data would be withheld for testing were participants 6, 7, and 11. Two models were chosen to be trained and tested in the manner discussed. These models were the Bagged Tree model, which scored the highest accuracy and other evaluation metrics in the initial machine learning model training, and the RUSBoosted Trees model, in an attempt to address the class imbalance between high and low stress. Table 36 shows the validation accuracy and the test accuracy of the two models.

*Table 36 Validation and testing accuracy of the machine learning models*

| Model | Validation Accuracy (%) | Test accuracy (%) |
|---|---|---|
| Bagged Trees | 99.8 | 56.0 |
| Fine Tree | 99.4 | 62.2 |

| | | |
|---|---|---|
| RUSBoosted trees | 94.8 | 50.6 |

The validation accuracy of all 3 of the models trained on 10 participants data are very high, with 10-fold cross validation giving validation accuracies of 99.8% and 99.4% for Bagged trees and the Fine tree model, respectively. However, when the test dataset is used to evaluate the accuracy of the model in detecting stress in previously unseen data, the highest classification accuracy is achieved by the fine tree model, achieving just 62.2%. The bagged tree model, which both previously in this study and in the paper by Harper et al. (2022) achieved the highest validation accuracy, achieves a test accuracy of only 56.0% [15]. This is most likely due to the relatively small sample size used to train and test the models, with the sample size being limited to 13 participants. This limited sample size is confounded by the fact that each participant likely has different physiological stress profiles, as in they can exhibit physiological stress in different ways, as demonstrated in the data analysis previously presented in this thesis.

PCA was then performed to identify the components of the dataset which explained the majority of the variance, and which features had the most weighting in explaining the variance of the dataset. The first PCA was performed in the full dataset of 295 sessions of data and 71 descriptive statistical features. The results showed that 84.6% of the variance was explained by principal component 1 (PC1), 10.5% was explained by principal component 2 (PC2), and 4.06% by principal component 3 (PC3).

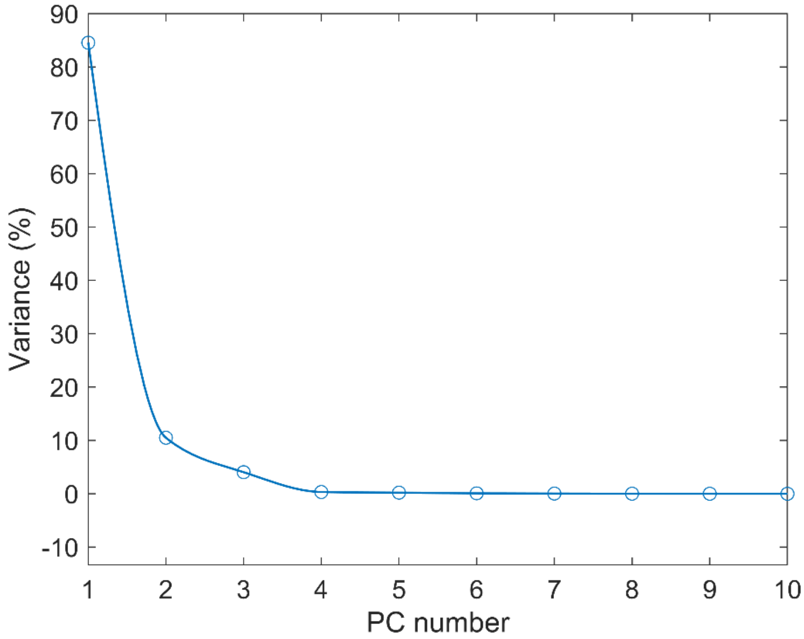



*Figure 5 Percentage variance against number of principal components*

For PC1, the feature with the greatest weight was IBI_MAX, followed by IBI_RMS, followed by IBI_MEAN, followed by IBI_STD, then IBI_MIN, followed by BVP_MAX, followed by BVP_MIN, and finally duration.

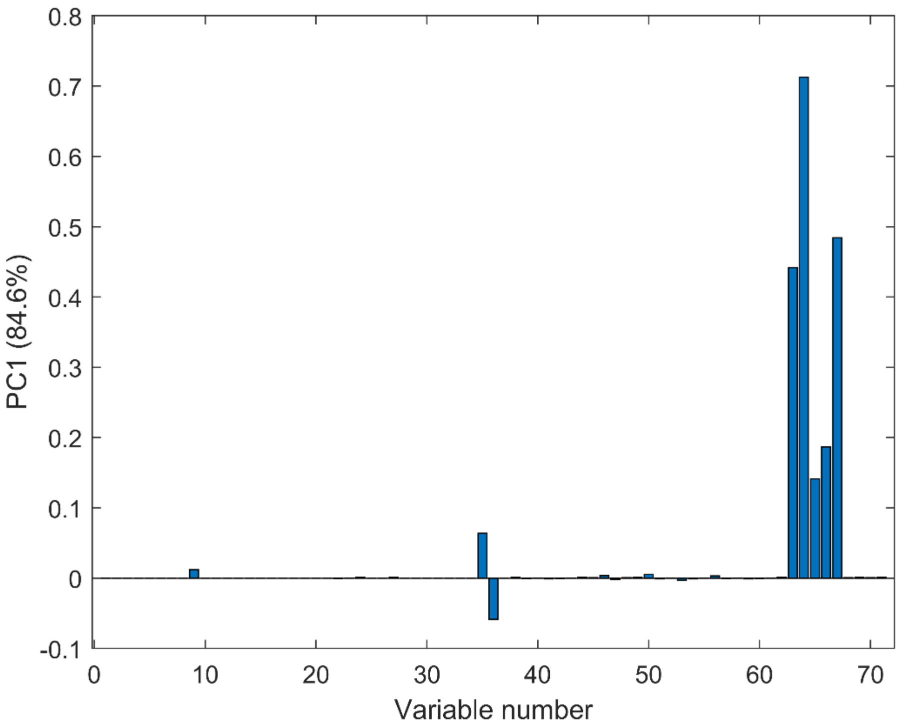



*Figure 6 Loading of the features for PC1*

For PC2, the feature with the greatest weight was BVP_MAX, followed by BVP_MIN, followed by IBI_MIN, followed by IBI_MAX, followed by IBI_STD.

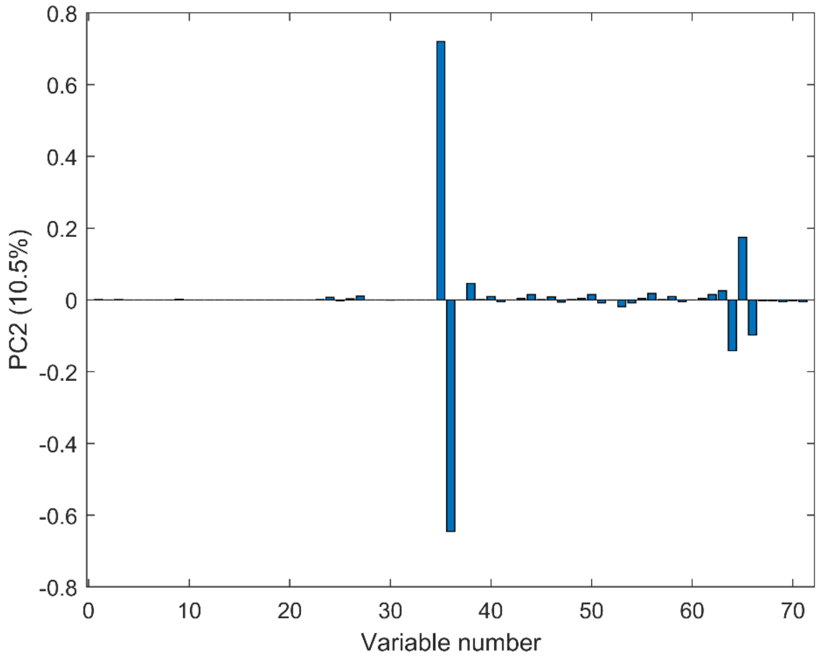



*Figure 7 Loading of features for PC2*

When the PCA scores for each instance for PC1 and PC2 are plotted together, we see a considerable overlap. Figure 8 shows the scores plotted on a scatter graph, with low stress instances being depicted by blue circles and high stress instances being depicted with red squares. There were a number of outlying points, which fall outside the blue circle on the graph, and outliers shall be discussed later in this chapter.

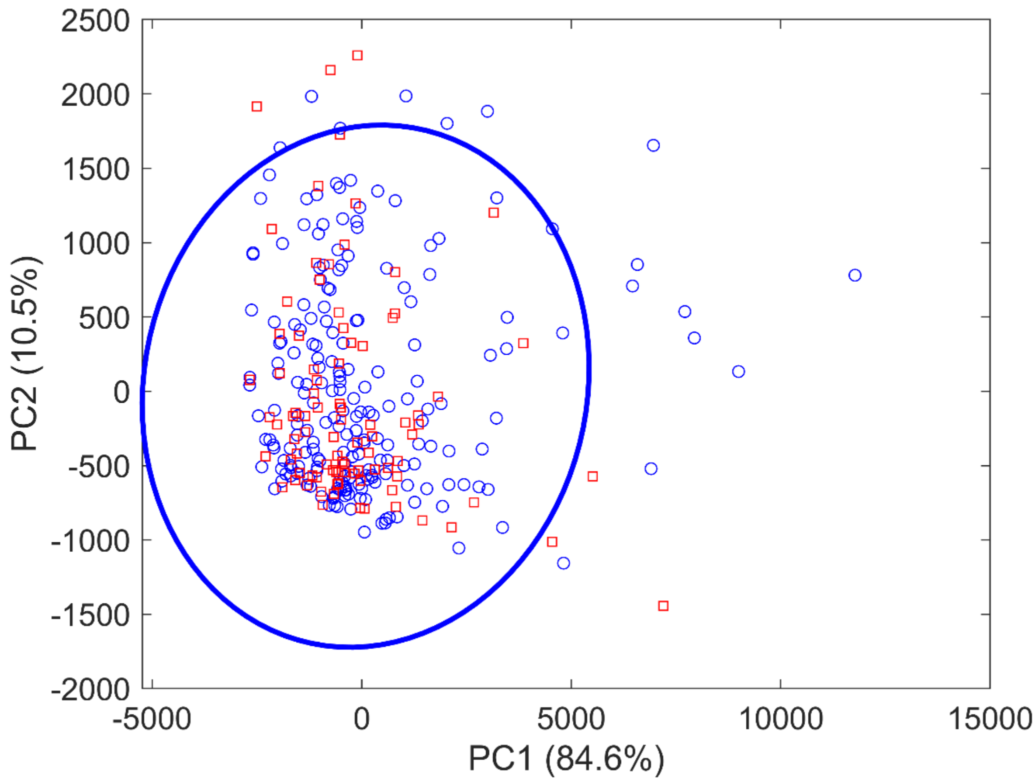



*Figure 8 Scatter plot of PCA scores for PC1 against PC2*

For PC3, the feature with the greatest weight was IBI_MIN, followed by IBI_STD, followed by IBI_MAX, followed by IBI_MEAN, followed by IBI_RMS, followed by BVP_MAX, and finally BVP_MEAN.

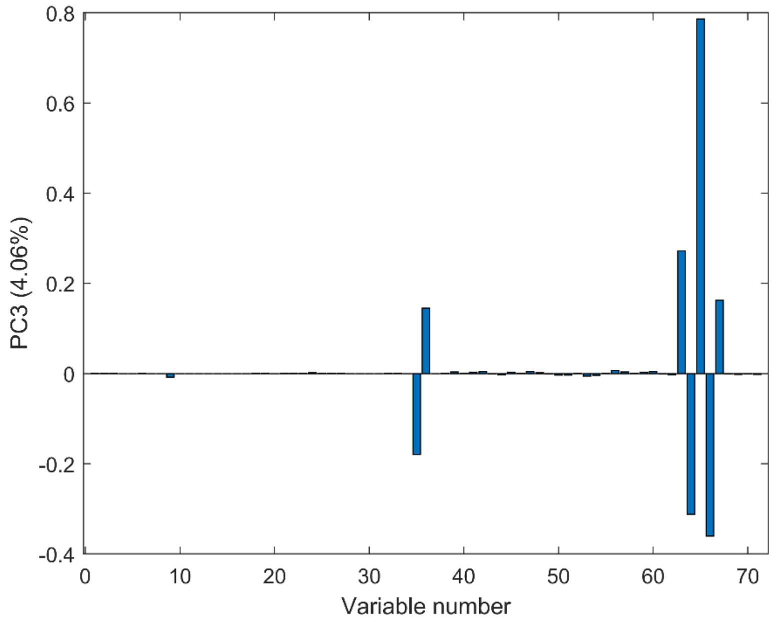



*Figure 9 Loading of features for PC3*

The fact that IBI and BVP related features are commonly found across the PCs as the features with the greatest weight could imply that measures of heart rate and cardiovascular activity are the best explainers of variance of all the physiological features tracked as part of this study. This is supported by the finding earlier in this chapter that there was statistically significant correlation found between a number of heart rate and BVP related features and ordinal stress.

Furthermore, another scatter graph was now produced to plot the PCA scores of each instance for PC1, PC2, and PC3.

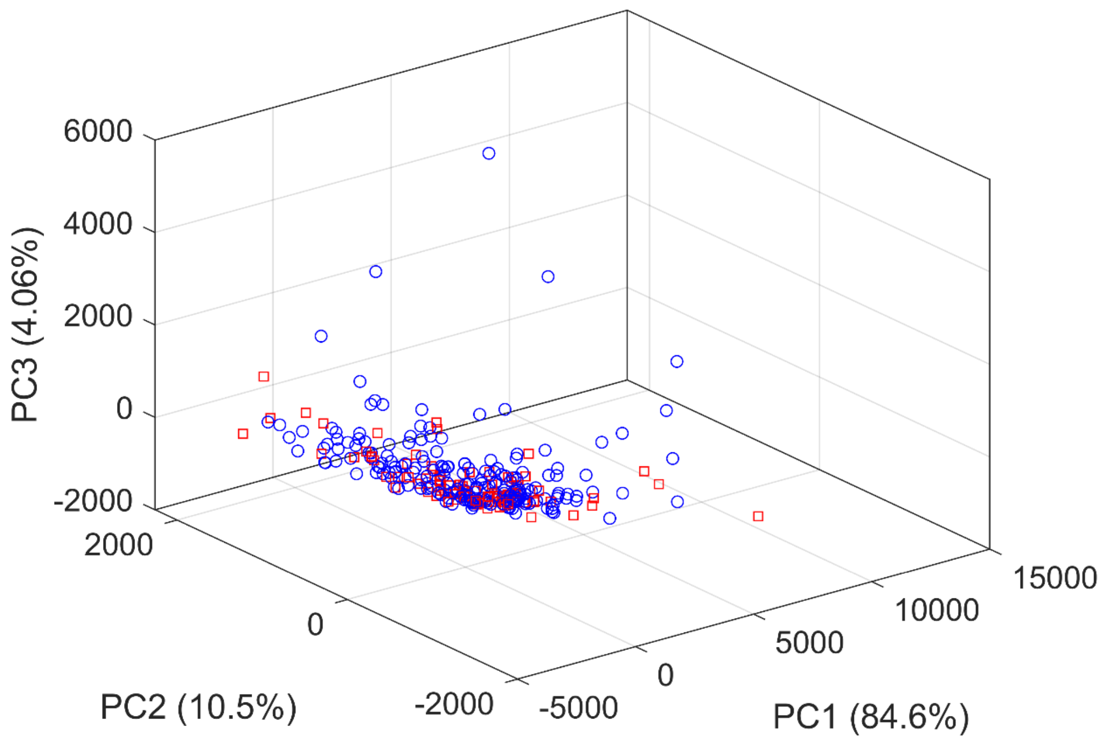



*Figure 10 PCA scores for each instance plotted by their value for PC1, PC2, and PC3*

Another principal component analysis was now performed on just the features identified as having significant weighting for PC1, PC2, and PC3. 85% of variance was accounted for by PC1 while 10.3% of variance was explained by PC2. When the PCA results were plotted on a scatter graph, there were once again a number of outliers identified. The majority of the outlying sessions were found to be from participant 8 and participant 9, with the majority of the sessions they recorded being outliers. The sessions which were these outliers were now removed from the dataset to be used for further analysis, including all of participant 8 and participant 9's session. With the features showing the greatest weighting across the first three principal components identified, with those components accounting for approximately 99% of the variance, and the outlying cases removed, it was decided to once more train the previously used machine-learning model based only on those identified features. A logistic regression model was also trained as it was suspected by the researcher it may achieve a high accuracy on the data at hand. The validation method chosen was 10-fold cross validation and 20 percent of the data was withheld as a testing set. The accuracy for the trained models can be found in table 35.

*Table 35 Validation and test accuracy of the machine learning models retrained on only the features with the greatest weightings*

| Model | Validation accuracy (%) | Test accuracy (%) |
|---|---|---|
| Logistic regression | 68.2 | 66.7 |
| Fine tree | 60.0 | 66.7 |
| Bagged tree | 69.1 | 61.1 |

The highest validation and training accuracy was achieved by the logistic regression model, with a test accuracy of 68.5%. This is a relatively good test accuracy compared to the models previously trained on the dataset and explained earlier in this chapter, however there are still issues with the model. The main issue is demonstrated in figure 11, showing the true positive and false negative rates of prediction for each class, in which one can clearly see that high stress sessions- given the label "2"- are misclassified at a relatively high rate. One explanation for this is the dataset is imbalanced with considerably more low stress instances than high stress, meaning the model is trained to have a bias towards classifying instances as low class.

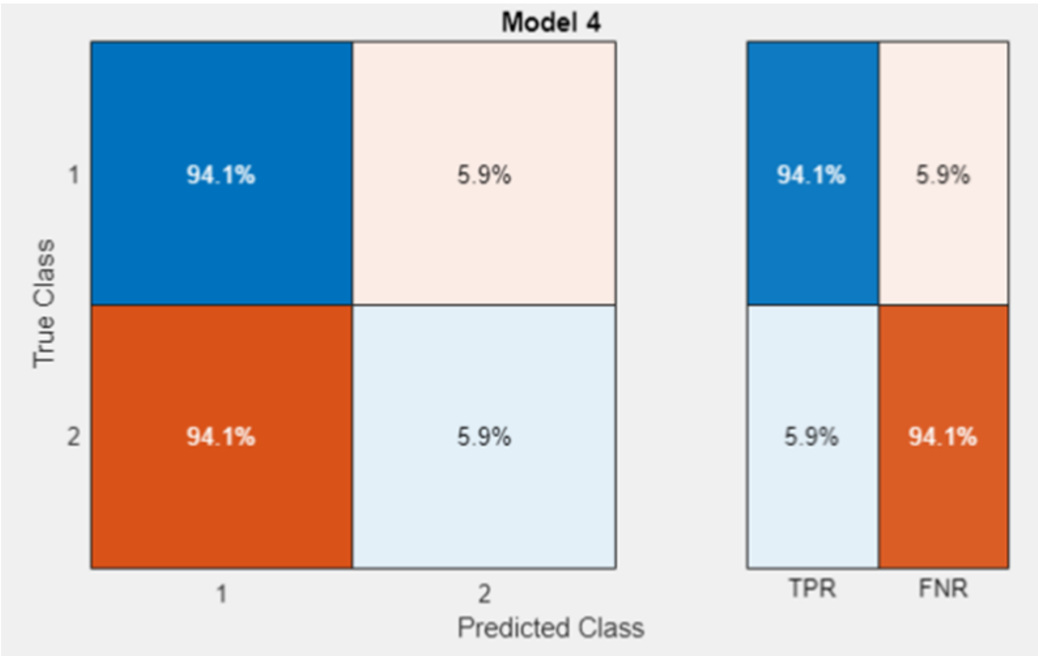



*Figure 11 True positive rate and false negative rate for the logistic regression model.*

In order to explore if the imbalance was the cause of the inaccuracy, and potentially increase the accuracy of the model's classification, SMOTE was used to artificially increase the number of instances of high stress observations. The result was a 126% increase in high stress instances, meaning there was an equal number of high stress and low stress instances. The fine tree, bagged tree, and logistic regression models were then retrain using the newly balanced dataset, again using 10-fold cross validation, and withholding 20% (three participants' worth as in the k-left out method described earlier) of the data as a test set. The validation and testing accuracy of each model can be shown in table 36.

*Table 36 Validation and test accuracy of the machine learning models trained with the balanced dataset and features with greatest weighting from PCA.*

| Model | Validation accuracy (%) | Test accuracy (%) |
|---|---|---|
| Bagged tree | 71.4 | 73.7 |
| Logistic regression | 66.1 | 67.1 |
| Fine tree | 66.8 | 61.8 |

The increased accuracies give some support to the notion that the imbalanced dataset was negatively affecting the performance of the models trained previously in this chapter. Further support can be attained by looking at the true positive and false negative rates of both the logistic regression model and the bagged tree model (figure 12 and figure 13, respectively).

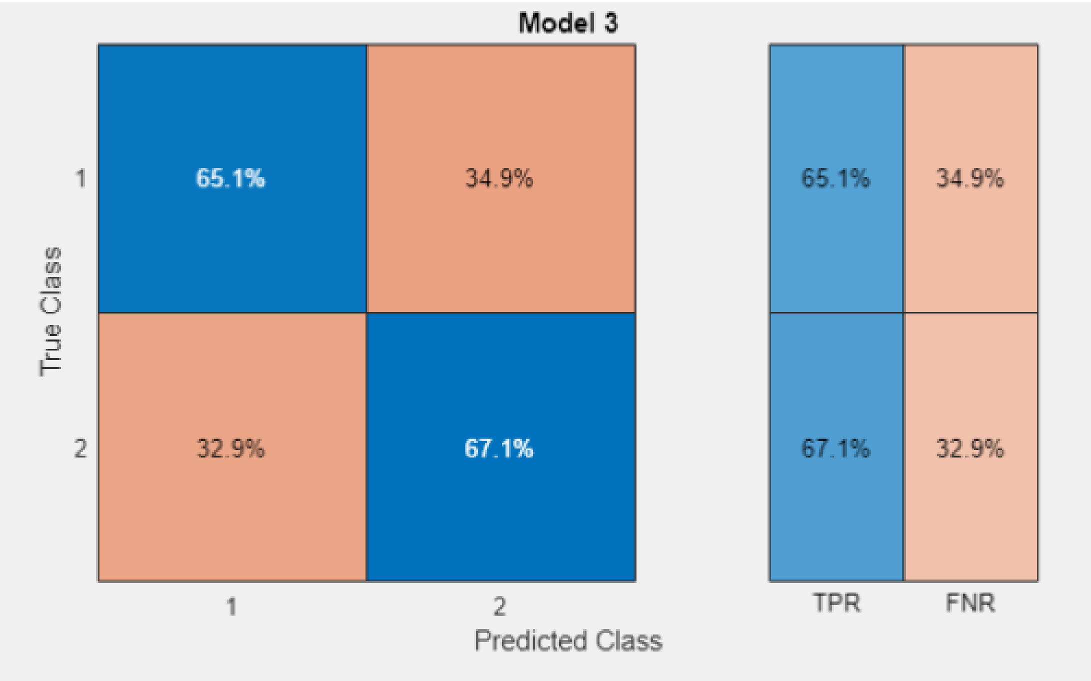



*Figure 12 True positive rate and false negative rate of the logistic regression model trained using SMOTE data*

As one can see, the logistic regression model has a significantly reduced misclassification rate for high stress instances, however the misclassification rate for low stress instances has also increased significantly. The bagged tree model has an approximately 10% higher true positive prediction rate for high stress instances than the logistic regression model, but an identical true positive prediction rate for low stress. This along with the overall higher test classification accuracy of 73.7% implies that the best model for predicting stress from the dataset collected for this study is the Bagged tree model.

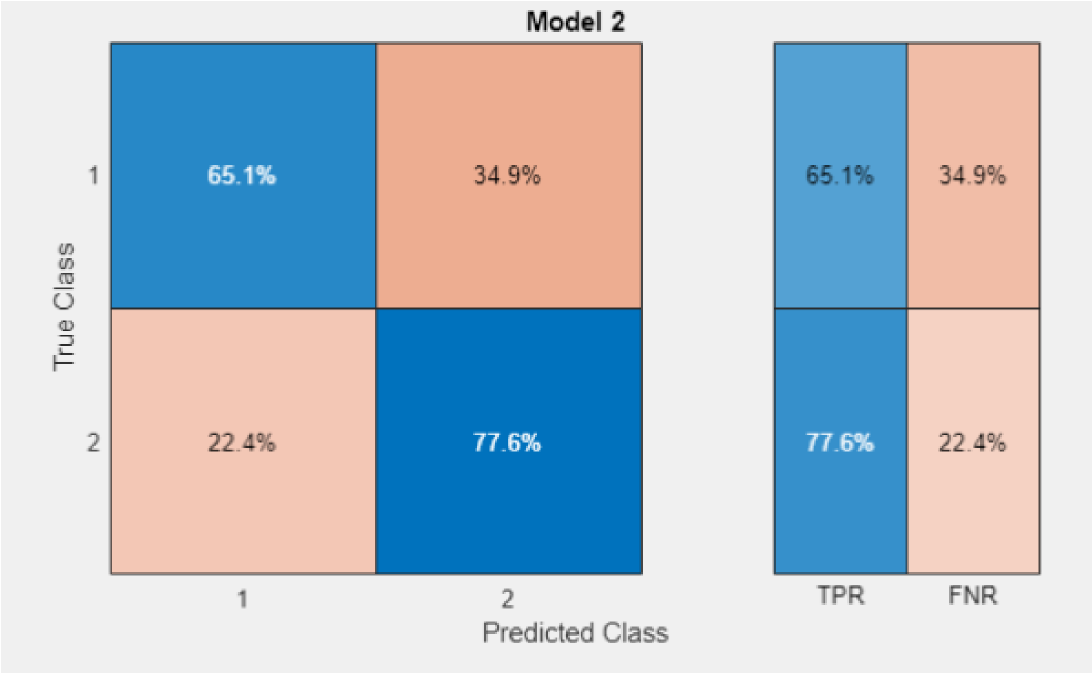



*Figure 13 True positive rate and false negative rate of the Bagged tree model*

## 5.4. Discussion

### 5.4.1 Relationships between stress and other attributes

Significant correlations were found between several features extracted from the physiological attributes and stress, with different features being significantly related to stress in home and office environments. For example, in the home environment, most of the features extracted from the EDA attribute correlate with stress, while only minimum EDA correlating with stress in the office environment. This difference could be explained by temperature differences in the two environments, as environmental temperatures can create noise in the EDA signal by increasing perspiration. Thus, if office environments used in the study were warmer than home environments, relatively high participant perspiration in the former may have negatively affected the accuracy of the EDA data. Another trend found in the data is that the features extracted from the attributes using the Empatica E4s PPG sensor tend to correlate significantly with stress more in the office environment than the home environment, with only median BVP correlating with stress in the home environment. On the other hand, maximum HR, HR variance, HR range, maximum BVP, minimum BVP, BVP variance, minimum IBI, NN50, NN25, pNN50, and pNN25 all correlate negatively with stress in office environments. The negative correlation between stress and these features could be considered surprising, as literature suggests that the correlation between heart rate and blood volume pulse, and stress is positive. However, the negative correlation between features extracted from IBI and stress is in line with the results of existing literature. The notion of a correlation between cardiovascular activity and stress is further supported by the results of the PCA which showed blood volume pulse and inter-beat interval related features had the greatest weight of any of the features in predicting binary stress level. In order to fully understand the differences in values of the physiological attributes and their relationship to environment it is important that future research collects environment and other contextual data which will allow for these variables' effects on the device's sensors to be fully evaluated.

Significant correlation between stress and productivity was found to exist for 2 participants in the study and not for others. Participant 2 was found to have a highly significant negative correlation between the two attributes, which is similar to the results of Galanti et al. (2021) who observed a clear negative correlation between stress and work probability in individuals working remotely during the COVID-19 pandemic [71]. This contrasts with Wolor et al. (2021), which reported that stress can be positively correlated with productivity, a phenomenon which is also exhibit by participant 5 of the current study [147, 148]. However, other than the two forementioned participants, no significant correlation was found between stress and productivity in the current study. This is in contrast to van Woerkom & Meyers (2015) who found that an employee productively using their strengths in work are more likely to be happy and less stressed than employees not engaging their strengths, suggesting a negative correlation between stress and productivity [149].

The lack of a significant correlation between environment and stress level in this project is not the expected result based on previous literature. Moretti et al. (2020) found no significant difference in individuals working from home rather than office environments, with 27.5% reporting no change, 39.2% reporting a reduction in stress and approximately 33% reporting increased stress [74]. This supports the findings of the current study; however considerable literature exists which contradicts the notion of no correlation between stress levels and environment. A study by Eurofound and the International Labour Office (2017) found that participants working from home as opposed to office reported increased stress due to reduced work-life barrier, longer hours of work and family conflict [150]. As such, 41% of home workers felt stressed compared with 25% of those working in an office, implying a strong relationship between working environment and work stress levels. Furthermore, Biron & van Veldhoven (2016) demonstrated that in most cases, office working was associated with higher levels of stress than home working, with the exception being employees with high levels of worktime control and job demand who have equal stress in both settings [151]. Indeed, the authors found that inter-person variations are important to understanding the effects of remote working. Likewise, Shao et al. (2021) found that the severity of the effect of stressors in different work environments depends upon the individual participants and their preferences or susceptibility to certain stressors [70]. Similarly, Galanti et al. (2021) found that social isolation and work-family conflict can increase stress levels in home environments, however working from home can also decrease other stressors related to commuting and the pandemic [71].

However, the results of the current study do suggest a relationship between stress and environment when accounting for the day of the week on which the work is occurring, with participants experiencing lower stress working from home from Monday to Wednesday and in the office on Thursday and Friday. The idea of different stress levels in different environments on different days of the week is supported by Tsai (2019), who found that the day of the week can have a considerable effect on an individual's feelings of happiness and stress, depending upon demographic information such as culture, religion, and place of residence [152]. In addition, Song & Gao (2020) also found that day of the week can be an important contributor to one's stress and affect *levels*, however they found that generally taking work home on weekdays was associated with more negative affect than doing that work in the office, however working from home for the full day was not associated with negative affect [75]. Moreover, they identified a greater prevalence of stress in participants working on the weekend. The findings of Freisthler et al. (2021) further support the idea, with their findings also indicating a reduction in stress of weekends as opposed to weekdays, which is not supported by the results of the current study [153]. However, a significant limitation of this study is that relatively few of the sessions (less than 5%) were collected on the weekend, limiting the reliability of any conclusions drawn about weekend working. Despite this, the current study does support the findings of existing literature regarding the notion that environment can affect stress levels differently on different days of the week.

Moreover, there is some evidence in the current study that there is a relationship between environment and stress level when considering the activity that the participant is performing, with "Data analysis and statistics" being associated with greater levels of stress in the office environment than the home environment, and the reverse being true for the activities "Meeting" and "Reading or writing emails". These results could be explained by the findings of Shao et al. (2021) who found that technology stressors are more prevalent at home. “Meeting” remotely tends to require more technology than meeting in person in an office environment, assuming meetings in the office environment aren’t remote, with the same being true for communicating by email and teaching labs or tutorials, both activities which by most utilised metrics are associated with higher stress in the home environment as opposed to the office environment [70]. Furthermore, data analysis and “Editing an electronic document” were found to be associated with higher stress, each by at least 2 utilised metrics, which could be explained by Prasad et al. (2020), who found employees completing such tasks may be more stressed in home environments due to lack of suggestion or guidance, or lack of work-life boundaries which may lead to working too much [154]. An alternative explanation could be that the difference in stress related to activity is due to participants feeling more capable of completing certain activities than others or are more comfortable completing certain activities in either the home or work environment. This would be supported by Woerkom & Meyers (2015) who state that employees engaging in activities they consider themselves proficient or capable of doing are more likely to be happier and less stress than those not engaging their perceived strengths [149]. However, this is speculation and in order to explain the differences in stress experienced in each environment related to activity, future iterations of this research would have to collect more information regarding the specific stressors experienced by participants. Nevertheless, there does appear to be some support in the literature and in the current study for different stress levels to be affected differently based on the work environment in which certain activities are performed.

Likewise, though no significant correlation was found overall between stress level and the time of day the activities were performed, working in the office of a morning and afternoon is associated with lower modal and mean stress than working at home at these times of the day, with higher levels of stress being demonstrated in the office environment of an evening. One could argue that this is in contrast to Song & Gao (2021) who found working from home of an evening was associated with higher stress [75]. Furthermore, the results of the current study show that mean stress increases in both work environments as the day progresses, as opposed to Zawadzki et al. (2019) who found some evidence of stressors being more common of the morning than the afternoon, and more common in the early evening as opposed to late evening, however they do not report these results as significant or consistent [155]. As such, no definitive relationship between stress level and time of day can be established, however it is clear there are some differences in stress at different times of day and thus this aspect should not be disregarded wholly as a potential contributing factor to stress and negative affect levels.

Overall, this section provides the answer to two of the research questions of the project. The first is “What differences exist in stress levels between instances of home and office working?” and the answer is that there is some evidence of differences between the stress experienced in each environment based on the time of day the work is being completed, the day of the week, and the activity being performed. Also, “What, if any, correlation exists between subjective stress, subjective productivity, time of day, day of week, physiological features and work environment?” is similarly answered in this section. The answer is that though there is no clear, significant correlation between the mentioned attributes overall in the dataset, there does appear to be relationships implied between some of the attributes. For example, participants experiencing lower stress working from home from Monday to Wednesday and in the office on Thursday and Friday, implying a relationship between environment and stress when considering the day of the week.

### 5.4.2. Prediction of stress

The validation prediction accuracy (99.3%) of the bagged tree ensemble model from this project is relatively high compared to the results of related projects. Indikawati & Winiarti (2020) achieved accuracy of up to 99% accuracy distinguishing stress from other physiological conditions, and stress and non-stress, respectively, using a random forest model [80]. This accuracy is very high compared to the original work of Schmidt et al. (2018), however, Indikawati & Winiarti (2020) do not specify the statistical features they extract from the physiological data, limiting the comparability of their results [17]. The predicted states in their project are different than the predicted states in this project, and they had more participants (15 as opposed to 13 on this project), and thus the comparison is limited. Another limitation to the comparison is that their data was collected in a controlled environment, whereas the data from the current project was collected in real-world, non-controlled environments.

Alternatively, Can et al. (2019) achieved 97.92% for distinguishing between mild, moderate and high stress using random forest, with data collected in a mixture of controlled and non-controlled environments [81]. Betti et al. (2018) used just 15 features to achieve an accuracy of up to of 86% for discriminating between stress and relaxation [77]. This is a relatively high accuracy, higher than the 74.5% accuracy in binary classification achieved by Wisjman et al. (2013) [78]. However, Han et al. (2017) achieved 94% accuracy in binary classification, using a combination of SVM and random forest (RF) classifiers on features extracted from ECG and respiration data [79]. Similar to Betti et al. (2018), Han et al. (2017) utilise the MAST in a controlled setting to elicit stress responses from 39 participants while the participants complete work-related tasks. The highest prediction accuracy was achieved using a combination of SVM and RF classifiers, not only achieving the aforementioned 94% accuracy for binary classification. The binary classification accuracy of the trained model from this project is higher than the accuracy of the models in all of the literature mentioned [77, 79].

The training accuracy of the bagged tree model was 56%, with previously unseen data being misclassified at a high rate, especially high stress instances, which is considerably lower than the accuracies achieved in the previously referenced studies in literature. As such, the use of PCA and SMOTE was decided upon, as is done by Askari et al. (2022) [156]. Using PCA and SMOTE, the authors reduce the complexity of their dataset and balance their classes, allowing their recurrent Neural Network to achieve an accuracy of 92% for detecting acute physiological stress. There are some limitations with comparing the work of Askari et al. (2022) with the current study, including that they used a modified version of SMOTE called ADASYN, which uses density distribution as a guide for generating synthetic instances rather than simply achieving an equal number of instances of each class, as is done with traditional SMOTE. Furthermore, the accuracies they achieve with their model by utilising PCA and data balancing are considerably higher than the 73.7% achieved in this study, but that is potentially explainable by their much higher sample size of participants. As such, future work on the current study would require the collection of data from a much larger sample size. Moreover, Anusha et al. (2018) used the physiological data of 34 participants to train a number of machine learning models for predicting stress [157]. Their highest achieved accuracy of 90.91% was achieved by a KNN which was trained on data which had been balanced using SMOTE. Compared to the same model trained on the original, imbalanced dataset (which achieved an accuracy of 75.79%) this is an increase in accuracy of approximately 15%. This is comparable with the increase in accuracy achieved by the use of data balancing in this current study, in which the accuracy increased from 56% to 73.7%, an increase of 17.7%. Once again, the relatively small sample size achieved in the current study is likely the reason for the limited accuracy scores achieved by the Bagged Tree model, however this study supports the finding of Anusha et al. (2018) that SMOTE and data balancing can increase the accuracy of machine learning models by reducing training bias towards the majority class.

There are some limitations of the model from this project. Firstly, the number of participants is comparable with some other research in the field, though collecting data from more participants would allow for greater reliability in results. Furthermore, the subjective stress values used to label the physiological data for the supervised machine learning was collected using self-assessment questionnaires, which can often suffer from inaccuracy due to bias. Moreover, the lack of demographic data, such as gender or marital status means that the effects of the participants' demographics cannot be explored.

Another limitation of the study was that there was limited exploration of tuning the model hyperparameters, which potentially limits the performance of the models trained. Hyperparameters are parameters which control the learning process and resulting parameters of a machine learning model, and so optimising these hyperparameters allows us to optimise and potentially improve the learning rate, model parameters, and ultimately model performance. One such hyperparameter is learning rate, which controls how much change occurs to the weights of the model at each stage of the learning process, or in other words how quickly the learning process takes place. The value is set between 0 and 1, with a commonly used value being 0.1, which means that the weights in the model are updated by 10% of the estimated weight error each time the weights are updated. Setting the learning rate too low may lead to the model weights converging too slowly or never at all, or the model becoming stuck with suboptimal weights and solutions. Alternatively, setting the learning rate too high may lead to convergence which is too quick, similarly reducing model performance [158]. Another important hyperparameter is optimisation function, or method used for optimisation. There are three main varieties of optimisation approaches: searches (random & grid) where the hyperparameter values are altered until the optimal values are found; evolutionary optimisation which mimics natural selection processes in mutating the hyperparameters and changing them iteratively and recombining the choices found to be effective; and Bayesian optimisation in which sequences of hyperparameters are tested, with each iteration of hyperparameters being a refinement of the ones in the previous iteration [159].

Hyperparameters which are specific to tree models and tree-based ensemble models include maximum tree depth, minimum leaf size, and branch split criterion [160]. The maximum tree depth parameter sets a limit to the number of splits and nodes that can be generated in the model. A higher value will mean more splits and more splits tends to lead to better classification, however it can also lead to overfitting of the model to the training data and can take longer and be more computationally expensive than if a smaller value were used

[161]. The minimum leaf size parameter sets a limit for the minimum number of observations that each node is allowed to have, which also limits the number of splits that can occur, as if a split would result in a node with fewer observations than is allowed by the minimum leaf size parameter the split would not occur. Having too low a value will likely lead to the model being overfit to the training data, however too large a value will mean the model cannot create enough splits to achieve a high performance [162]. Finally, the branch split criterion is the criterion used to decide how nodes are split into child nodes. Examples of split criteria include: the Gini impurity, where the probability of the parent node giving a misclassification is measured and the split that would reduce that probability the most is made; information gain, where the split is made by calculating the difference between the entropy of the parent node and the average entropies of the child nodes for each possible split; and reduction in variance, where the split which most reduces the variance in a target variable is made to generate the child nodes [163-165]. Each of these split criteria has their advantages and disadvantages, and so an exploration of how the use each of them effects the accuracy of the models trained in this study should be explored in future work.

Hyper parameters can be set either manually or automatically, with a wide range of potential methods for automatic hyperparameter tuning being available [161]. The hyperparameters chosen in this study were set manually and different hyperparameter values were not explored, and so future work should explore different hyperparameters and compare different methods for automated hyperparameter tuning.

Another limitation of this study was the lack of a calibration process that would make a person specific model more accurate by factoring in their baseline heart rate. Resting or baseline heart rates can vary person from person, based upon factors such as their physical fitness levels, age, health conditions, and an array of other demographic factors [166, 167]. This inter-personal difference in resting heart rate presents a potential problem for the models trained in this study. Normalisation of the physiological data attributes and features was useful in this regard, as it meant that rather than comparing the raw values of the features- which could vary greatly between participants- the models could learn from data which was all normalised within the range of 0 to 1. This reduced the effects of the interpersonal differences as all participants data was measured on the same scale and had similar variance [15]. However, only one method of normalisation was explored, and perhaps other methods could be explored in future work to see if this has any effect on the accuracy of the models. Another method for addressing this limitation which could be explored in future work is the engineering of new features which measure the distance of the current value of an attribute from the baseline or resting value of that attribute for each participant. For example, one could collect a baseline reading of each participants resting attributes, in this instance we shall focus on heart rate, for a session of around 5 minutes, as has been done in previous literature [17]. An average heart rate value could then be calculated from this session, with this being considered the participant’s resting heart rate, and a new feature could be engineered from each current window of the data being collected- difference from baseline heart rate (HR_DFB). To calculate this feature, one would take the average value of the heart rate measured in the relevant window and from it minus the resting heart rate value. Thus, if a participant has a resting heart rate value of 80 beats per minute (bpm), and the average value for the relevant window is 124 bpm, the HR_DFB value for that window would be 44. This would help to calibrate the models and reduce misclassifications. To illustrate how, consider the following example: participant A has a current heart rate of 114 bpm and a resting heart rate of 105 bpm, and participant b has a current heart rate of 115 bpm and a resting heart rate of 82 bpm. Whereas a model may interpret the current heart rates of both participants to be the same and thus equally indicative of stress, the HR_DFB value of each participant (9 for participant A and 33 for participant B) demonstrates that participant’s B heart rate is far more indicative of stress than participant A’s. Similar features could be calculated for other physiological attributes, such as skin temperature.

Overall, however, this section does answer the research question “Can machine learning be used to identify and predict occurrences of stress in both work environments, based on the physiological attributes of the individual, and what is the best model for doing so?“. The answer is that the best model for classifying and predicting the occurrences of work stress based upon the physiological data attributes of individuals in this study is the bagged tree ensemble model, achieving a binary validation classification accuracy of 99.3% and a testing accuracy of 73.7%.

### 5.4.3. Potential stress mitigations

Mitigating and reducing stress is vital, as this will help in the reduction of negative health outcomes related with long-term stress, such as cardiovascular disease and sexual dysfunction [3, 4]. The mitigations can be split into two main categories: organisational, which refers to policies, rules or incentives organisations can put in place to influence or help their employees; and personal, which are things the individual employees can do to mitigate their own work stress [168].

The results of the data analysis in this chapter of the thesis indicate a relationship between stress and environment on certain days of the week. As such, one personal work stress mitigation method could be for employees to consider the impact working from home or office environments on each day will impact their stress. For example, if an individual employee has a large number of meetings with colleagues on one day, it may be less stressful for them to work in an office environment on that day, as this would reduce their exposure to technological stressors. One limitation of this mitigation method is it requires an employee to have an insight on the influence environment has on their work stress on each day of the week. To achieve this, individuals could keep a work stress diary to track their feelings of work stress over time in each environment [169, 170]. Institutions could help individuals understand their fluctuations in work stress by encouraging them to keep a work stress study, or providing them with a system, like the machine learning model developed in this project, which can detect when the employee is experiencing high work stress. Institutions should also foster generous flexible working plans which allow employees the freedom to work in their preferred environment whenever possible [171].

Furthermore, the results of the analysis of the dataset in this thesis chapter show that working from the office of an evening- colloquially “staying late at the office”- should be discouraged as it is associated with higher average stress than office work at any other time of day, and higher average stress than working in the home environment at the same time of day.  Institutions could impose policies should as closing the offices at 5pm each evening, with employees having to request permission to stay later [172]. Similarly, institutions could provide incentive for employees to finish work before the evening begins on days associated with high average office environment work stress- Thursday and Friday in this study- by offering extra-curricular activities or social events. On the individual level, employees should aim to enforce a strict work-life boundary and ensure that they only work late in the office when absolutely necessary [73].

Similarly, there is a suggested relationship between the activity being performed in each environment and the stress experienced. One personal mitigation for this stress could be to ensure that certain activities are completed in the environment where they prevent the least potential stressors. For example, meeting in person instead of remotely may reduce technology stressors, however working in an office environment could increase other stressors. As such, each individual employee should be free to choose their own work environment where possible [171]. However, there will always be some instances in which employees are bound to experience high work stress, such as when they have an approaching work deadline, or they unavoidably have to perform a task they dislike. In these instances, individuals could utilise meditation or deep breathing exercises, as these methods have been shown to be highly effective at reduce stress frequency and intensity [173]. Furthermore, individuals and Institutions should aim to ensure that employees are given ample time after having to complete a stressful task to recover. In addition, organisation should ensure that all work being performed by employees is meaningful, useful, and where possible engages an employee's strengths, thus increasing their productivity [149].

## 5.5. Chapter summary

This chapter answered 4 of the research questions posited in this project. It was established in this chapter that there is some evidence of differences between the stress experienced in each environment based on the time of day the work is being completed, the day of the week, and the activity being performed. Also, though there is no clear, significant correlation between any of the attributes overall in the dataset, there does appear to be relationships implied between some of the attributes. For example, participants experiencing lower stress working from home from Monday to Wednesday and in the office on Thursday and Friday, implying a relationship between environment and stress when considering the day of the week. There is also an implication of relationship between stress and environment when considering the time of day, as working in the office of a morning and afternoon is associated with lower modal and mean stress than working at home at these times of the day, with higher levels of stress being demonstrated in the office environment of an evening. Furthermore, it is clearly established that the best machine learning model for classifying the stress level of participants on the current project is the bagged tree ensemble model, which achieved a binary classification accuracy of 73.7% when dimensionality reduction and data balancing techniques were applied.

# 6. Conclusion

## 6.1. Introduction

In this chapter, the conclusions of the proposed study are drawn. Section 6.2. reiterates all of the research questions and provides a concise summary of the answers to each provided in the previous chapters of the thesis. Section 6.3. then explains the futures work which should be conducted on this project, as well as in the dementia-related difficulties and work stress prediction domains in general.

## 6.2. Conclusion

There were 7 research questions which were posited in this research project. Here all 7 are reiterated and the answers provided to them in this thesis summarised.

### 6.2.1. What are the best physiological attributes and sensors for predicting the occurrence of dementia-related difficulties?

The best physiological attributes and sensing modalities for predicting dementia-related difficulties identified in the review detailed in chapter 3 were movement from accelerometers, coronary data attributes using PPG or ECG, with these attributes including BVP and HR, and EDA. All three of these modalities were found to be highly efficacious in identifying difficulties alone and in combination with other physiological attributes or sensing modalities. One other notable sensing modality is skin temperature, despite only being found to be an efficacious sensing modality in 1 of the included papers. However, it appears to be useful in combination with other attributes for predicting difficulties and so could be seen as a potentially efficacious sensing modality. More work should be done on evaluating a correlation between skin temperature and dementia-related agitation; however, it should be noted that no significant correlations between features extracted from skin temperature and stress in the data analysis outlined in chapter 5, not supporting the use of skin temperature for detecting stress-related difficulties.

### 6.2.2. What is the best wearable device which can be used to track the indicative physiological attributes of people with dementia, in a comfortable, unobtrusive and unobstructive manner?

The results of the device search outlined in chapter 3 of the thesis highlight that the Empatica E4 is the best wearable device which can be used to track the indicative physiological attributes of people with dementia, in a comfortable, unobtrusive and unobstructive manner. This device is designed for use in clinical trials, offering high quality, accurate and precise measurements of physiological data. It contains a PPG sensor, a tri-axial accelerometer, an EDA sensor, an infrared thermopile (for measuring skin temperature), and an internal real-time clock (allowing the readings of the other sensors to be accurately time referenced). It has an internal real time clock which allows for data to be timestamped and has a built-in event marker button so that events can be marked as they occur, allowing for easier matching of events to physiological data. It was found to be usable and comfortable for people with dementia in a number of previous studies in literature and contained all of the most commonly used and efficacious sensors for collecting physiological indicators of dementia-related agitation. However, the main drawback to this device is that it is highly expensive and may not be usable for some researchers due to financial limitations.

### 6.2.3. What is the best machine learning model for predicting the occurrences of dementia-related difficulties and the context in which they occur?

This question was not answered in the research project due to the pivot of the project to a different domain prior to the work on answering this question commencing. One could answer this question with a systematic literature review, as there is plenty of literature in which machine learning is used to identify dementia-related difficulties, though as identified in previous works, no comprehensive system for all dementia-related difficulties yet exists. As such, it would be best if this question were answered using a comprehensive dataset of physiological attributes from people with dementia being used to train novel comprehensive dementia-related difficulty prediction models.

### 6.2.4. What differences exist in stress levels between instances of home and office working?

Stress experienced in the office environment tends to be higher on Mondays, Tuesdays, and Wednesdays, and during the evenings as opposed to office working sessions at those times and on those days. Alternatively, home-based work sessions have higher average stress values on Thursdays and Fridays, and in the morning and afternoons. The differences in home and office work stress are also apparent when accounting for activity. For example, having meetings in the home environment is associated with higher average stress than in the office environment, with a similar but less obvious trend existing for email. This may be due to technology stressors, with added technological stressors being involved in remote meetings in home environments compared to office environments. On the other hand, data analysis and editing electronic documents have generally higher average stress in the office environment as opposed to the home environment. Furthermore, there is some difference in the physiological data features which significantly correlate with stress in home and office environments. For example, in the home environment, most of the features extracted from the EDA attribute correlate with stress, while only minimum EDA correlates with stress in the office environment. Overall, it is clear that there are differences in the stress that is experienced in each work environment, however, to fully understand those differences it may be important for future research to collect demographic information from participants, such as age and marital status.

### 6.2.5. What, if any, correlation exists between subjective stress, subjective productivity, time of day, day of week, physiological features, and work environment?

Significant correlations exist between stress and a number of physiological features, including many EDA, HR, BVP, and IBI features. The notion that BVP and IBI related features were useful in predicting stress was also supported by the PCA conducted which found BVP and IBI features tended to have the greatest weight in explaining the variance of the dataset. No significant correlation exists between stress and productivity in the dataset overall, however for one participant there was a negative correlation between stress and productivity, and for another participant that correlation was positive. Furthermore, though no clear correlation exists between stress and environment, a relationship is implied between stress and environment due to differences in stress in each environment when considering time of day, day of week, and activity. Likewise, no significant correlation was found between time of day and stress, day of the week and stress, and activity and stress.

### 6.2.6. Can machine learning be used to identify and predict occurrences of stress in both work environments, based on the physiological attributes of the individual, and what is the best model for doing so?

A novel bagged tree ensemble was developed for binary stress classification, achieving a validation accuracy of 99.3%, a relatively high precision and AUC, and few instances of misclassified observations. This accuracy is very high compared to models found in existing literature. This accuracy was reduced significantly when the model was tested with unseen data, however it was possible to increase the accuracy to 73.7% using PCA and SMOTE. The significantly increased accuracy after the PCA and SMOTE were employed highlights the importance of ensuring that relevant features are identified and utilised, and the importance of having a balanced dataset. However, it can be difficult to collect a balanced dataset, as in this study and a number of studies in literature, all the collected datasets were skewed towards having more low stress than high stress instances. As such, data balancing techniques such as SMOTE could be considered vital for stress detection. There are some limitations of the model from this project. Firstly, the number of participants is comparable with other research in the field, though collecting data from more participants would allow for greater reliability in results. Furthermore, the subjective stress values used to label the physiological data for the supervised machine learning was collected using self-assessment questionnaires, which can often suffer from inaccuracy due to bias. Moreover, the lack of demographic data, such as gender or marital status means that the effects of the participants' demographics cannot be explored. However, based on the analysis of data collected on this project, machine learning models trained using physiological data can be used to predict the levels of stress with high accuracy, and the best model found in this project was the bagged tree ensemble model.

### 6.2.7. What mitigations can be used to reduce or prevent work stress once it is detected, based on the causal environmental and personal factors identified in the answer to question

The potential mitigations to stress based on the results of the data analysis in chapter 5 of the thesis can be split into institutional and personal mitigation methods. The potential personal stressors including ensuring a work-life balance and avoiding working late at the office, understanding the impact environment has on one's work stress on each day of the week and time of day and choosing to work in the lowest stress environment where possible, and utilising methods such as meditation or deep breathing to reduce stress when stressors are unavoidable. Institutional stress mitigations could include policies to promote working from home on days when that is less stressful for certain employees (obviously requiring a study of their workforce's habits and stress), discouraging working in the office of an evening and doing overtime unless absolutely and unavoidably necessary, providing incentives for employees to finish work at a reasonable time such as social events and extra-curricular activities, and encouraging employees to be mindful of their stress levels and the impact their work environments can have on their stress. Overall, the work stress mitigations proposed in this thesis are relatively simple, common-sense ideas which should not be hard to implement in academic workplaces, or offices of many creative technology companies.

## 6.3. Discussion of Future Work

Future work on this project should focus upon ensuring the reliability of the classification results of the developed model. The main method for doing this should be conducting the data collection methodology with more participants, increasing the size of the dataset with which the models can be trained and evaluated. Furthermore, future iterations of the study should aim to collect demographic and environmental factors which could influence the differences in stress levels and physiological attribute values between the two work environments. Moreover, the transfer of the methods and lessons learnt back to the dementia-related difficulties detection domain should be conducted in the future, as the development of a comprehensive dementia-related difficulties support system remains highly important and topical.

In the dementia-related difficulties domain future work should focus on a dataset which contains a comprehensive range of physiological indicators of dementia-related difficulties. The dataset should ideally contain data from as many different people, with as many different types of dementia and symptoms as possible. Moreover, steps should be taken to ensure that the data is shareable with other researchers in the domain, such as anonymising or pseudonymising personalised data and obtaining informed consent from participants and their representatives for sharing the data in a controlled and ethical manner with researchers with valid reasons for accessing the data. Finally, the dataset should be used for developing a system which can predict the occurrence of a comprehensive range of dementia-related difficulties in a timely manner.

Future work in the work stress detection domain should focus on the development of more personalised models which focus on identifying the stress of individuals, as there are clear interpersonal differences in the way in which individuals exhibit and experience stress. Furthermore, more work should be done on comparing the physiological and subjective stress levels in home and office work environments, with an emphasis on how demographic and environmental factors influence the physiological attributes. Finally, more emphasis should be placed upon individual employees and their employers having an understanding and mindfulness of how choice of work environment can impact stress levels, with institutions encouraging the tracking of work stress by employees using a diary or a physiological data-based system for detecting stress, such as the model developed in this thesis.

# 7. References


[1] J. Ma and Y. Peng, "The performance costs of illegitimate tasks: The role of job identity and flexible role orientation," *Journal of Vocational Behavior,* vol. 110, pp. 144-154, 2019.

[2] S. J. Lupien, R.-P. Juster, C. Raymond, and M.-F. Marin, "The effects of chronic stress on the human brain: From neurotoxicity, to vulnerability, to opportunity," *Frontiers in neuroendocrinology,* vol. 49, pp. 91-105, 2018.

[3] K. S. Gawlik, B. M. Melnyk, and A. Tan, "Associations between stress and cardiovascular disease risk factors among million hearts priority populations," *American Journal of Health Promotion,* vol. 33, no. 7, pp. 1063-1066, 2019.

[4] M. Kivimäki and A. Steptoe, "Effects of stress on the development and progression of cardiovascular disease," *Nature Reviews Cardiology,* vol. 15, no. 4, pp. 215-229, 2018.

[5] E. S. van der Valk, M. Savas, and E. F. van Rossum, "Stress and obesity: are there more susceptible individuals?," *Current obesity reports,* vol. 7, no. 2, pp. 193-203, 2018.

[6] C. Delcea and A. SCAUNAȘ, "The impact of daily stress on sexual activity in stable couples in Romania," *International Journal of Advanced Studies in Sexology,* vol. 4, no. 1, pp. 67-74, 2022.

[7] J. D. O'Sullivan *et al.*, "The impact of perceived stress on the hair follicle: Towards solving a psychoneuroendocrine and neuroimmunological puzzle," *Frontiers in neuroendocrinology,* vol. 66, p. 101008, 2022.

[8] S. Wolvetang, J. M. van Dongen, E. Speklé, P. Coenen, and F. Schaafsma, "Sick Leave Due to Stress, What are the Costs for Dutch Employers?," *Journal of Occupational Rehabilitation,* vol. 32, no. 4, pp. 764-772, 2022.

[9] L. Bolliger, J. Lukan, M. Luštrek, D. De Bacquer, and E. Clays, "Protocol of the STRess at Work (STRAW) Project: How to Disentangle Day-to-Day Occupational Stress among Academics Based on EMA, Physiological Data, and Smartphone Sensor and Usage Data," *International Journal of Environmental Research and Public Health,* vol. 17, no. 23, p. 8835, 2020.

[10] F. Amato *et al.*, "CLONE: a Promising System for the Remote Monitoring of Alzheimer's Patients: An Experimentation with a Wearable Device in a Village for Alzheimer's Care," in *Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good*, 2018, pp. 255-260.

[11] L. Bolliger *et al.*, "The association between day-to-day stress experiences, recovery, and work engagement among office workers in academia–An Ecological Momentary Assessment study," *Plos one,* vol. 18, no. 2, p. e0281556, 2023.

[12] G. Vos, K. Trinh, Z. Sarnyai, and M. R. Azghadi, "Generalizable machine learning for stress monitoring from wearable devices: a systematic literature review," *International Journal of Medical Informatics,* p. 105026, 2023.

[13] M. Harper, F. Ghali, A. Hussain, and D. Al-Jumeily, "Review of Methods for Data Collection Experiments with People with Dementia and the Impact of COVID-19," in *International Conference on Intelligent Computing*, 2021: Springer, pp. 132-147.

[14] M. Harper and F. Ghali, "Roles of caregivers in physiological data collection experiments with people with dementia and mitigating the impacts of COVID-19," in *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, 2021: IEEE, pp. 149-155.

[15] F. G. M. Harper, W. Khan, "Comparison of subjective and physiological stress levels in home and office work environments, presented at the International Conference on Intelligent Computing 2022, Xi'an, China, 2022.

[16] Y.-W. Li *et al.*, "The autonomic nervous system: a potential link to the efficacy of acupuncture," *Frontiers in Neuroscience,* vol. 16, p. 1038945, 2022.

[17] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, 2018, pp. 400-408.

[18] C. Melander, J. Martinsson, and S. Gustafsson, "Measuring electrodermal activity to improve the identification of agitation in individuals with dementia," *Dementia and geriatric cognitive disorders extra,* vol. 7, no. 3, pp. 430-439, 2017.

[19] N. Ahmadi *et al.*, "Quantifying Occupational Stress in Intensive Care Unit Nurses: An Applied Naturalistic Study of Correlations Among Stress, Heart Rate, Electrodermal Activity, and Skin Temperature," *Human Factors,* p. 00187208211040889, 2021.

[20] D. Bredesen, *The end of Alzheimer's: The first program to prevent and reverse cognitive decline*. Penguin, 2017.

[21] J. Luo, C. R. Beam, and M. Gatz, "Is stress an overlooked risk factor for dementia? a systematic review from a lifespan developmental perspective," *Prevention Science,* pp. 1-14, 2022.

[22] M. Harper, J. Mustafina, A. J. AlJaaf, J. Lunn, S. Yasen, and F. Ghali, "Data science techniques to support prediction, diagnosis, and ReCODE treatment of Alzheimer's Disease," presented at the Developments in eSystems Engineering 2019, Kazan, Russia, 7th-10th October, 2019.

[23] C.-T. Yao, B.-O. Lee, H. Hong, and Y.-C. Su, "Evaluation of the music therapy program interventions on agitated behavior for people with dementia in Taiwan institutional care," *Educational Gerontology,* pp. 1-12, 2022.

[24] V. Puthusseryppady, S. Morrissey, H. Spiers, M. Patel, and M. Hornberger, "Predicting real world spatial disorientation in Alzheimer's disease patients using virtual reality navigation tests," *Scientific Reports,* vol. 12, no. 1, p. 13397, 2022.

[25] R. C. Hamdy *et al.*, "Patients with dementia are easy victims to predators," *Gerontology and geriatric medicine,* vol. 3, p. 2333721417734684, 2017.

[26] M. Reale *et al.*, "Network between cytokines, cortisol and occupational stress in gas and oilfield workers," *International journal of molecular sciences,* vol. 21, no. 3, p. 1118, 2020.

[27] G. Giorgi, L. I. Lecca, J. M. Leon-Perez, S. Pignata, G. Topa, and N. Mucci, "Emerging issues in occupational disease: mental health in the aging working population and cognitive impairment—a narrative review," *BioMed Research International,* vol. 2020, 2020.

[28] A. Giorgi *et al.*, "Wearable technologies for mental workload, stress, and emotional state assessment during working-like tasks: a comparison with laboratory technologies," *Sensors,* vol. 21, no. 7, p. 2332, 2021.

[29] B. M. Booth, H. Vrzakova, S. M. Mattingly, G. J. Martinez, L. Faust, and S. K. D'Mello, "Toward Robust Stress Prediction in the Age of Wearables: Modeling Perceived Stress in a Longitudinal Study with Information Workers," *IEEE Transactions on Affective Computing,* vol. 13, no. 4, pp. 2201-2217, 2022.

[30] M. Harper and F. Ghali, "A Systematic review of wearable devices for tracking physiological indicators of Dementia-related difficulties," presented at the Developments in E-Systems, Online, 2020.

[31] M. Harper, F. Ghali, A. Hussain, and D. Al-Jumeily, "Challenges in Data Capturing and Collection for Physiological Detection of Dementia-Related Difficulties and Proposed Solutions," in *International Conference on Intelligent Computing*, 2021: Springer, pp. 162-173.

[32] B. Caracciolo, K. Palmer, R. Monastero, B. Winblad, L. Bäckman, and L. Fratiglioni, "Occurrence of cognitive impairment and dementia in the community: a 9-year-long prospective study," *Neurology,* vol. 70, no. 19 Part 2, pp. 1778-1785, 2008.
[33] Q. Meng, M.-S. Lin, and I. Tzeng, "Relationship between exercise and Alzheimer's disease: A narrative literature review," *Frontiers in neuroscience,* vol. 14, p. 131, 2020.
[34] A. V. Galende, M. E. Ortiz, S. L. Velasco, M. L. Luque, C. L. d. S. de Miguel, and C. P. Jurczynska, "Report by the Spanish Foundation of the Brain on the social impact of Alzheimer disease and other types of dementia," *Neurología (English Edition),* vol. 36, no. 1, pp. 39-49, 2021.
[35] K. Pal, N. Mukadam, I. Petersen, and C. Cooper, "Mild cognitive impairment and progression to dementia in people with diabetes, prediabetes and metabolic syndrome: a systematic review and meta-analysis," *Social psychiatry and psychiatric epidemiology,* vol. 53, no. 11, pp. 1149-1160, 2018.
[36] B. Penke, M. Szűcs, and F. Bogár, "Oligomerization and conformational change turn monomeric β-amyloid and tau proteins toxic: Their role in Alzheimer's pathogenesis," *Molecules,* vol. 25, no. 7, p. 1659, 2020.
[37] T. Berman and A. Bayati, "What are neurodegenerative diseases and how do they affect the brain?," *Frontiers for Young Minds,* vol. 6, 2018.
[38] J. B. Tonga, J. Š. Benth, E. A. Arnevik, K. Werheid, M. S. Korsnes, and I. D. Ulstein, "Managing depressive symptoms in people with mild cognitive impairment and mild dementia with a multicomponent psychotherapy intervention: a randomized controlled trial," *International Psychogeriatrics,* vol. 33, no. 3, pp. 217-231, 2021.
[39] Y.-T. Wu, L. Clare, and F. E. Matthews, "Relationship between depressive symptoms and capability to live well in people with mild to moderate dementia and their carers: results from the improving the experience of dementia and enhancing active life (IDEAL) programme," *Aging & Mental Health,* vol. 25, no. 1, pp. 38-45, 2021.
[40] Y. Eisenmann, H. Golla, H. Schmidt, R. Voltz, and K. M. Perrar, "Palliative care in advanced dementia," *Frontiers in Psychiatry,* vol. 11, p. 699, 2020.
[41] R. N. Kalaria, "The pathology and pathophysiology of vascular dementia," *Neuropharmacology,* vol. 134, pp. 226-239, 2018.
[42] E.-S. Lee, J.-H. Yoon, J. Choi, F. R. Andika, T. Lee, and Y. Jeong, "A mouse model of subcortical vascular dementia reflecting degeneration of cerebral white matter and microcirculation," *Journal of Cerebral Blood Flow & Metabolism,* vol. 39, no. 1, pp. 44-57, 2019.
[43] X. Li *et al.*, "Mitochondrial Protection and Against Glutamate Neurotoxicity via Shh/Ptch1 Signaling Pathway to Ameliorate Cognitive Dysfunction by Kaixin San in Multi-Infarct Dementia Rats," *Oxidative Medicine and Cellular Longevity,* vol. 2021, 2021.
[44] W. J. Deardorff and G. T. Grossberg, "Behavioral and psychological symptoms in Alzheimer's dementia and vascular dementia," *Handbook of Clinical Neurology,* vol. 165, pp. 5-32, 2019.
[45] A. Surendranathan *et al.*, "Clinical diagnosis of Lewy body dementia," *BJPsych open,* vol. 6, no. 4, 2020.
[46] K. S. Chin, N. Yassi, L. Churilov, C. L. Masters, and R. Watson, "Prevalence and clinical associations of tau in Lewy body dementias: a systematic review and meta-analysis," *Parkinsonism & related disorders,* vol. 80, pp. 184-193, 2020.
[47] J.-P. Taylor *et al.*, "New evidence on the management of Lewy body dementia," *The Lancet Neurology,* vol. 19, no. 2, pp. 157-169, 2020.

[48] O. Piguet and F. Kumfor, "Frontotemporal dementias: main syndromes and underlying brain changes," *Current Opinion in Neurology,* vol. 33, no. 2, pp. 215-221, 2020.
[49] M. Montalbano *et al.*, "TDP-43 and tau oligomers in Alzheimer's disease, amyotrophic lateral sclerosis, and frontotemporal dementia," *Neurobiology of disease,* vol. 146, p. 105130, 2020.
[50] S. Ducharme *et al.*, "Recommendations to distinguish behavioural variant frontotemporal dementia from psychiatric disorders," *Brain,* vol. 143, no. 6, pp. 1632-1650, 2020.
[51] C. R. Marshall *et al.*, "Primary progressive aphasia: a clinical approach," *Journal of neurology,* vol. 265, no. 6, pp. 1474-1490, 2018.
[52] P. Masrori and P. Van Damme, "Amyotrophic lateral sclerosis: a clinical review," *European journal of neurology,* vol. 27, no. 10, pp. 1918-1929, 2020.
[53] K. A. Peterson, K. Patterson, and J. B. Rowe, "Language impairment in progressive supranuclear palsy and corticobasal syndrome," *Journal of neurology,* vol. 268, no. 3, pp. 796-809, 2021.
[54] M. Stamelou, G. Respondek, N. Giagkou, J. L. Whitwell, G. G. Kovacs, and G. U. Höglinger, "Evolving concepts in progressive supranuclear palsy and other 4-repeat tauopathies," *Nature Reviews Neurology,* vol. 17, no. 10, pp. 601-620, 2021.
[55] S. S. Khan *et al.*, "Agitation Detection in People Living with Dementia using Multimodal Sensors," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019: IEEE, pp. 3588-3591.
[56] S. S. Khan *et al.*, "Daad: A framework for detecting agitation and aggression in people living with dementia using a novel multi-modal sensor network," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2017: IEEE, pp. 703-710.
[57] R. Alam *et al.*, "Motion biomarkers for early detection of dementia-related agitation," in *Proceedings of the 1st Workshop on Digital Biomarkers*, 2017, pp. 15-20.
[58] R. Alam, M. Anderson, A. Bankole, and J. Lach, "Inferring physical agitation in dementia using smartwatch and sequential behavior models," in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 2018: IEEE, pp. 170-173.
[59] R. Alam, A. Bankole, M. Anderson, and J. Lach, "Multiple-Instance Learning for Sparse Behavior Modeling from Wearables: Toward Dementia-Related Agitation Prediction," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019: IEEE, pp. 1330-1333.
[60] J. S. Sefcik, M. Ersek, J. R. Libonati, S. C. Hartnett, N. A. Hodgson, and P. Z. Cacchione, "Heart rate of nursing home residents with advanced dementia and persistent vocalizations," *Health and Technology,* pp. 1-5, 2019.
[61] C. Nesbitt, A. Gupta, S. Jain, K. Maly, and H. R. Okhravi, "Reliability of wearable sensors to detect agitation in patients with dementia: A pilot study," in *Proceedings of the 2018 10th International Conference on Bioinformatics and Biomedical Technology*, 2018, pp. 73-77.
[62] T. Zhou, Z. Song, and K. Sundmacher, "Big data creates new opportunities for materials research: A review on methods and applications of machine learning for materials design," *Engineering,* vol. 5, no. 6, pp. 1017-1026, 2019.
[63] N. Huck, "Large data sets and machine learning: Applications to statistical arbitrage," *European Journal of Operational Research,* vol. 278, no. 1, pp. 330-342, 2019.
[64] I. Lee and Y. J. Shin, "Machine learning for enterprises: Applications, algorithm selection, and challenges," *Business Horizons,* vol. 63, no. 2, pp. 157-170, 2020.

[65] J. M. Luis-Martínez, M. C. Martínez-Martínez, and I. A. García-Montalvo, "Physical activity: academic stress regulator in time of covid-19 pandemic. Covid-19 and academic stress: COVID-19 AND ACADEMIC STRESS," *Journal of Negative and No Positive Results,* vol. 6, no. 6, pp. 872-880, 2021.

[66] P. Odriozola-González, Á. Planchuelo-Gómez, M. J. Irurtia, and R. de Luis-García, "Psychological effects of the COVID-19 outbreak and lockdown among students and workers of a Spanish university," *Psychiatry research,* vol. 290, p. 113108, 2020.

[67] Y. Marcén-Román, A. Gasch-Gallen, I. I. Vela Martín de la Mota, E. Calatayud, I. Gómez-Soria, and B. Rodríguez-Roca, "Stress Perceived by University Health Sciences Students, 1 Year after COVID-19 Pandemic," *International Journal of Environmental Research and Public Health,* vol. 18, no. 10, p. 5233, 2021.

[68] M. Fornili *et al.*, "Psychological distress in the academic population and its association with socio-demographic and lifestyle characteristics during COVID-19 pandemic lockdown: Results from a large multicenter Italian study," *PloS one,* vol. 16, no. 3, p. e0248370, 2021.

[69] A. Salazar, J. Palomo-Osuna, H. de Sola, J. A. Moral-Munoz, M. Dueñas, and I. Failde, "Psychological Impact of the Lockdown Due to the COVID-19 Pandemic in University Workers: Factors Related to Stress, Anxiety, and Depression," *International journal of environmental research and public health,* vol. 18, no. 8, p. 4367, 2021.

[70] Y. Shao, Y. Fang, M. Wang, C.-H. D. Chang, and L. Wang, "Making daily decisions to work from home or to work in the office: The impacts of daily work-and COVID-related stressors on next-day work location," *Journal of Applied Psychology,* vol. 106, no. 6, p. 825, 2021.

[71] T. Galanti, G. Guidetti, E. Mazzei, S. Zappalà, and F. Toscano, "Work from home during the COVID-19 outbreak: The impact on employees' remote work productivity, engagement, and stress," *Journal of occupational and environmental medicine,* vol. 63, no. 7, p. e426, 2021.

[72] W. Tan *et al.*, "Is returning to work during the COVID-19 pandemic stressful? A study on immediate mental health status and psychoneuroimmunity prevention measures of Chinese workforce," *Brain, behavior, and immunity,* vol. 87, pp. 84-92, 2020.

[73] D. W. Irawanto, K. R. Novianti, and K. Roz, "Work from home: Measuring satisfaction between work–life balance and work stress during the COVID-19 pandemic in Indonesia," *Economies,* vol. 9, no. 3, p. 96, 2021.

[74] A. Moretti, F. Menna, M. Aulicino, M. Paoletta, S. Liguori, and G. Iolascon, "Characterization of home working population during COVID-19 emergency: a cross-sectional analysis," *International journal of environmental research and public health,* vol. 17, no. 17, p. 6284, 2020.

[75] Y. Song and J. Gao, "Does telework stress employees out? A study on working at home and subjective well-being for wage/salary workers," *Journal of Happiness Studies,* vol. 21, no. 7, pp. 2649-2668, 2020.

[76] V. L. O. Messenger. J, Gschwind. L, Boehmer. S, Vermeylen. G, Wilkens, M., "Working anytime, anywhere: The effects on the world of work," Eurofound & ILO, 2017.

[77] S. Betti *et al.*, "Evaluation of an integrated system of wearable physiological sensors for stress monitoring in working environments by using biological markers," *IEEE Transactions on Biomedical Engineering,* vol. 65, no. 8, pp. 1748-1758, 2017.

[78] J. Wijsman, B. Grundlehner, H. Liu, J. Penders, and H. Hermens, "Wearable physiological sensors reflect mental stress state in office-like situations," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013: IEEE, pp. 600-605.

[79] L. Han, Q. Zhang, X. Chen, Q. Zhan, T. Yang, and Z. Zhao, "Detecting work-related stress with a wearable device," *Computers in Industry,* vol. 90, pp. 42-49, 2017.

[80] F. I. Indikawati and S. Winiarti, "Stress detection from multimodal wearable sensor data," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 771, no. 1: IOP Publishing, p. 012028.

[81] Y. S. Can, N. Chalabianloo, D. Ekiz, and C. Ersoy, "Continuous stress detection using wearable sensors in real life: Algorithmic programming contest case study," *Sensors,* vol. 19, no. 8, p. 1849, 2019.

[82] P. Concheiro-Moscoso, B. Groba, S. Monteiro-Fonseca, N. Canosa, and C. Queirós, "SQoF-WEAR Project. The Use of Wearable Devices to Identify the Impact of Stress on Workers' Quality of Life," *Engineering Proceedings,* vol. 7, no. 1, p. 25, 2021.

[83] E. E. Kaczor, S. Carreiro, J. Stapp, B. Chapman, and P. Indic, "Objective measurement of physician stress in the emergency department using a wearable sensor," in *Proceedings of the... Annual Hawaii International Conference on System Sciences. Annual Hawaii International Conference on System Sciences*, 2020, vol. 2020: NIH Public Access, p. 3729.

[84] B. Kikhia *et al.*, "Utilizing ambient and wearable sensors to monitor sleep and stress for people with BPSD in nursing homes," *Journal of Ambient Intelligence and Humanized Computing,* vol. 9, no. 2, pp. 261-273, 2018.

[85] H. F. Posada-Quintero and K. H. Chon, "Innovations in electrodermal activity data collection and signal processing: A systematic review," *Sensors,* vol. 20, no. 2, p. 479, 2020.

[86] L. Giromini, R. Morese, A. Salatino, M. Di Girolamo, D. J. Viglione, and A. Zennaro, "Rorschach Performance Assessment System (R-PAS) and vulnerability to stress: A preliminary study on electrodermal activity during stress," *Psychiatry research,* vol. 246, pp. 166-172, 2016.

[87] C. Lai Kwan, Y. Mahdid, R. Motta Ochoa, K. Lee, M. Park, and S. Blain-Moraes, "Wearable Technology for Detecting Significant Moments in Individuals with Dementia," *BioMed research international,* vol. 2019, 2019.

[88] L. Valembois, C. Oasi, S. Pariel, W. Jarzebowski, C. Lafuente-Lafuente, and J. Belmin, "Wrist actigraphy: a simple way to record motor activity in elderly patients with dementia and apathy or aberrant motor behavior," *The journal of nutrition, health & aging,* vol. 19, no. 7, pp. 759-764, 2015.

[89] K. S. Thomas, W. Zhang, P. Y. Cornell, L. Smith, B. Kaskie, and P. C. Carder, "State variability in the prevalence and healthcare utilization of assisted living residents with dementia," *Journal of the American Geriatrics Society,* vol. 68, no. 7, pp. 1504-1511, 2020.

[90] C. Benson, A. Friz, S. Mullen, L. Block, and A. Gilmore-Bykovskyi, "Ethical and Methodological Considerations for Evaluating Participant Views on Alzheimer's and Dementia Research," *Journal of Empirical Research on Human Research Ethics,* p. 1556264620974898, 2020.

[91] M. Kaenampornpan, N. D. Khai, and K. Kawattikul, "Wearable Computing for Dementia Patients," in *International Conference on Computing and Information Technology*, 2020: Springer, pp. 21-30.

[92] E. Grober, D. Wakefield, A. R. Ehrlich, P. Mabie, and R. B. Lipton, "Identifying memory impairment and early dementia in primary care," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring,* vol. 6, pp. 188-195, 2017.

[93] L. McGarrigle, S. E. Howlett, H. Wong, J. Stanley, and K. Rockwood, "Characterizing the symptom of misplacing objects in people with dementia: findings from an online tracking tool," *International psychogeriatrics,* vol. 31, no. 11, pp. 1635-1641, 2019.

[94] M. H. Connors, K. Seeher, A. Teixeira-Pinto, M. Woodward, D. Ames, and H. Brodaty, "Dementia and caregiver burden: A three-year longitudinal study," *International journal of geriatric psychiatry,* vol. 35, no. 2, pp. 250-258, 2020.

[95] A. P. Allen *et al.*, "Informal caregiving for dementia patients: the contribution of patient characteristics and behaviours to caregiver burden," *Age and ageing,* vol. 49, no. 1, pp. 52-56, 2020.

[96] J.-A. Su and C.-C. Chang, "Association between family caregiver burden and affiliate stigma in the families of people with dementia," *International journal of environmental research and public health,* vol. 17, no. 8, p. 2772, 2020.

[97] B. S. Husebo, H. L. Heintz, L. I. Berge, P. Owoyemi, A. T. Rahman, and I. V. Vahia, "Sensing technology to facilitate behavioral and psychological symptoms and to monitor treatment response in people with dementia. A Systematic Review," *Frontiers in pharmacology,* vol. 10, p. 1699, 2020.

[98] W.-H. Lin, D. Wu, C. Li, H. Zhang, and Y.-T. Zhang, "Comparison of heart rate variability from PPG with that from ECG," in *The international conference on health informatics*, 2014: Springer, pp. 213-215.

[99] D. Goerss *et al.*, "Automated sensor-based detection of challenging behaviors in advanced stages of dementia in nursing homes," *Alzheimer's & Dementia,* 2019.

[100] D. L. Devoe and A. P. Pisano, "Surface micromachined piezoelectric accelerometers (PiXLs)," *Journal of Microelectromechanical Systems,* vol. 10, no. 2, pp. 180-186, 2001.

[101] R. Mukhiya *et al.*, "Design, modelling and system level simulations of DRIE-based MEMS differential capacitive accelerometer," *Microsystem technologies,* vol. 25, no. 9, pp. 3521-3532, 2019.

[102] F. Liu *et al.*, "Optimal design of high-g MEMS piezoresistive accelerometer based on Timoshenko beam theory," *Microsystem Technologies,* vol. 24, no. 2, pp. 855-867, 2018.

[103] J. Jussila, N. Venho, H. Salonius, J. Moilanen, J. Liukkonen, and M. Rinnetmäki, "Towards ecosystem for research and development of electrodermal activity applications," in *Proceedings of the 22nd International Academic Mindtrek Conference*, 2018, pp. 79-87.

[104] H. F. Posada–Quintero, Y. Kong, and K. H. Chon, "Objective pain stimulation intensity and pain sensation assessment using machine learning classification and regression based on electrodermal activity," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology,* vol. 321, no. 2, pp. R186-R196, 2021.

[105] M. Yu, G. Yu, and B. Dai, "Graphene fiber-based strain-insensitive wearable temperature sensor," *IEEE Sensors Letters,* vol. 4, no. 10, pp. 1-4, 2020.

[106] D. Nepi *et al.*, "Validation of the heart-rate signal provided by the Zephyr bioharness 3.0," in *2016 Computing in Cardiology Conference (CinC)*, 2016: IEEE, pp. 361-364.

[107] Zephyr, "BioHarness 3.0: User Manual," ed: ed: Zephyr Technology, 2012.

[108] Zephyr. "BioHarness 3 - Wireless Professional Heart Rate & Physiological Monitor with Bluetooth (S-M)." Amazon.com. https://www.empatica.com/research/e4/?utm_source=Google&utm_medium=cpc&utm_campaign=conversion&gclid=Cj0KCQiAvc_xBRCYARIsAC5QT9l9EUAmaWXCRxEQ17aUYewwgoMkYIn-xFsuJld5R-Ib_6wGuhYqKToaAsomEALw_wcB (accessed 30th January 2020).

[109] Empatica. "E4 Wristband." Empatica. https://www.empatica.com/en-gb/research/e4 (accessed 30th January, 2020).

[110] PhilipsElectronics. "DTI-2." FCCID. https://fccid.io/2AALC-DTI2/User-Manual/UserManual-pdf-2206443.pdf (accessed 30th January, 2020).

[111] T. G. Stavropoulos, G. Meditskos, S. Andreadis, K. Avgerinakis, K. Adam, and I. Kompatsiaris, "Semantic event fusion of computer vision and ambient sensor data for activity recognition to support dementia care," *Journal of Ambient Intelligence and Humanized Computing,* pp. 1-16, 2016.

[112] Pebble. "Pebble Smartwatch - Black." Amazon.co.uk. https://www.amazon.co.uk/Pebble-301BL-Smartwatch-Black/dp/B00BKEQBI0 (accessed 30th January, 2020).

[113] B. Ye *et al.*, "Challenges in collecting big data in a clinical environment with vulnerable population: Lessons learned from a study using a multi-modal sensors platform," *Science and engineering ethics,* vol. 25, no. 5, pp. 1447-1466, 2019.

[114] N. Vuong, S. Chan, C. T. Lau, S. Chan, P. L. K. Yap, and A. Chen, "Preliminary results of using inertial sensors to detect dementia-related wandering patterns," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015: IEEE, pp. 3703-3706.

[115] A. Karakostas, I. Lazarou, G. Meditskos, T. G. Stavropoulos, I. Kompatsiaris, and M. Tsolaki, "Sensor-based in-home monitoring of people with dementia using remote web technologies," in *2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*, 2015: IEEE, pp. 353-357.

[116] S. Spasojevic *et al.*, "A pilot study to detect agitation in people living with dementia using multi-modal sensors," *Journal of Healthcare Informatics Research,* pp. 1-17, 2021.

[117] S. Teipel *et al.*, "Multidimensional assessment of challenging behaviors in advanced stages of dementia in nursing homes—The insideDEM framework," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring,* vol. 8, pp. 36-44, 2017.

[118] NHS. "How to get a dementia diagnosis." NHS.uk. https://www.nhs.uk/conditions/dementia/diagnosis/ (accessed April 19th, 2020).

[119] MerseyCare. "Important information about changes to our services." NHS. https://www.merseycare.nhs.uk/about-us/news/coronavirus-changes-to-mersey-cares-services/ (accessed 15th March, 2021).

[120] M. Aveiro. "Rapid Response, Dementia patients: a vulnerable population during the COVID-19 Pandemic." BMJ. https://www.bmj.com/content/370/bmj.m3709/rr-6 (accessed 15th March, 2021).

[121] ONS. "Number of deaths in care homes notified to the Care Quality Commission, England." GOV.uk. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/numberofdeathsincarehomesnotifiedtothecarequalitycommissionengland (accessed 15th March, 2021).

[122] L. Cuffaro, F. Di Lorenzo, S. Bonavita, G. Tedeschi, L. Leocani, and L. Lavorgna, "Dementia care and COVID-19 pandemic: a necessary digital revolution," *Neurological Sciences,* vol. 41, no. 8, pp. 1977-1979, 2020.

[123] M. Canevelli *et al.*, "Facing Dementia During the COVID-19 Outbreak," *Journal of the American Geriatrics Society,* 2020.

[124] I. McCarthy *et al.*, "Infrastructureless pedestrian navigation to assess the response of Alzheimer's patients to visual cues," 2015.

[125] M. Kolakowski and B. Blachucki, "Monitoring Wandering Behavior of Persons Suffering from Dementia Using BLE Based Localization System," in *2019 27th Telecommunications Forum (TELFOR)*, 2019: IEEE, pp. 1-4.

[126] Y. Liu, B. Batrancourt, F. Marin, and R. Levy, "Evaluation of apathy by single 3D accelerometer in ecological condition: Case of patients with behavioral variant of fronto-temporal dementia," in *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 2018: IEEE, pp. 1-4.

[127] R. Radziszewski, H. K. Ngankam, V. Grégoire, D. Lorrain, H. Pigot, and S. Giroux, "Designing calm and non-intrusive ambient assisted living system for monitoring nighttime wanderings," *International Journal of Pervasive Computing and Communications,* 2017.

[128] J. Gong *et al.*, "Home wireless sensing system for monitoring nighttime agitation and incontinence in patients with alzheimer's disease," in *Proceedings of the conference on Wireless Health*, 2015, pp. 1-8.
[129] P. Koldrack, S. J. Teipel, and T. Kirste, "Sensing disorientation of persons with dementia in outdoor wayfinding tasks using wearable sensors to enable situation-aware navigation assistance," *Alzheimer's & Dementia: The Journal of the Alzheimer's Association,* vol. 12, no. 7, p. P160, 2016.
[130] M. Donaldson, "An assistive interface for people with dementia," in *Proceedings of the Australasian Computer Science Week Multiconference*, 2018, pp. 1-5.
[131] J. Kowalska, J. Mazurek, and J. Rymaszewska, "Analysis of the degree of acceptance of illness among older adults living in a nursing home undergoing rehabilitation–an observational study," *Clinical interventions in aging,* vol. 14, p. 925, 2019.
[132] L. Clare, C. Quinn, I. R. Jones, and R. T. Woods, "“I Don’t Think Of It As An Illness”: Illness Representations in Mild to Moderate Dementia," *Journal of Alzheimer's Disease,* vol. 51, no. 1, pp. 139-150, 2016.
[133] R. Alam, N. Homdee, S. Wolfe, J. Hayes, and J. Lach, "Besi: behavior learning and tracking with wearable and in-home sensors - a dementia case-study: poster abstract," presented at the Proceedings of the International Conference on Internet of Things Design and Implementation, Montreal, Quebec, Canada, 2019. [Online]. Available: https://doi.org/10.1145/3302505.3312595.
[134] R. Radziszewski, H. Ngankam, H. Pigot, V. Grégoire, D. Lorrain, and S. Giroux, "An ambient assisted living nighttime wandering system for elderly," presented at the Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services, Singapore, Singapore, 2016. [Online]. Available: https://doi.org/10.1145/3011141.3011171.
[135] A. Bieber, N. Nguyen, G. Meyer, and A. Stephan, "Influences on the access to and use of formal community care by people with dementia and their informal caregivers: a scoping review," *BMC health services research,* vol. 19, no. 1, p. 88, 2019.
[136] C. Reed *et al.*, "Factors associated with long-term impact on informal caregivers during Alzheimer’s disease dementia progression: 36-month results from GERAS," *International psychogeriatrics,* pp. 1-11, 2019.
[137] Á. Romero-Martínez, G. Hidalgo-Moreno, and L. Moya-Albiol, "Neuropsychological consequences of chronic stress: the case of informal caregivers," *Aging & mental health,* vol. 24, no. 2, pp. 259-271, 2020.
[138] V. C.-C. Cheng *et al.*, "The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2," *Journal of Infection,* vol. 81, no. 1, pp. 107-114, 2020.
[139] K. K. Cheng, T. H. Lam, and C. C. Leung, "Wearing face masks in the community during the COVID-19 pandemic: altruism and solidarity," *The Lancet,* 2020.
[140] M. A. U. Alam, N. Roy, S. Holmes, A. Gangopadhyay, and E. Galik, "Automated functional and behavioral health assessment of older adults with dementia," in *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, 2016: IEEE, pp. 140-149.
[141] A. Bankole *et al.*, "BESI: Behavioral and Environmental Sensing and Intervention for Dementia Caregiver Empowerment—Phases 1 and 2," *American Journal of Alzheimer's Disease & Other Dementias®,* vol. 35, p. 1533317520906686, 2020.
[142] A. Simons, T. Doyle, D. Musson, and J. Reilly, "Impact of Physiological Sensor Variance on Machine Learning Algorithms," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020: IEEE, pp. 241-247.

[143] P. Bobade and M. Vani, "Stress detection with machine learning and deep learning using multimodal physiological data," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020: IEEE, pp. 51-57.

[144] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system," 2017.

[145] P. Schmidt, A. Reiss, R. Dürichen, and K. V. Laerhoven, "Wearable-based affect recognition—A review," *Sensors,* vol. 19, no. 19, p. 4079, 2019.

[146] C. Kyriakou, S. E. Christodoulou, and L. Dimitriou, "Do vehicles sense, detect and locate speed bumps?," *Transportation research procedia,* vol. 52, pp. 203-210, 2021.

[147] C. W. Wolor, S. Dalimunthe, I. Febrilia, and S. Martono, "How to manage stress experienced by employees when working from home due to the Covid-19 virus outbreak," *International Journal of Advanced Science and Technology,* vol. 29, no. 5, pp. 8359-64, 2020.

[148] C. W. Wolor, A. Nurkhin, and Y. Citriadin, "Is Working from Home Good for Work-Life Balance, Stress, and Productivity, or does it Cause Problems?," *Humanities and Social Sciences Letters,* vol. 9, no. 3, pp. 237-249, 2021.

[149] M. van Woerkom and M. C. Meyers, "My strengths count! Effects of a strengths-based psychological climate on positive affect and job performance," *Human Resource Management,* vol. 54, no. 1, pp. 81-103, 2015.

[150] V. L. O. Messenger. J, Gschwind. L, Boehmer. S, Vermeylen. G, Wilkens, M., "Working anytime, anywhere: The effects on the world of work," Eurofound & ILO, 2017.

[151] M. Biron and M. Van Veldhoven, "When control becomes a liability rather than an asset: Comparing home and office days among part-time teleworkers," *Journal of Organizational Behavior,* vol. 37, no. 8, pp. 1317-1337, 2016.

[152] M.-C. Tsai, "The good, the bad, and the ordinary: The day-of-the-week effect on mood across the globe," *Journal of Happiness Studies,* vol. 20, no. 7, pp. 2101-2124, 2019.

[153] B. Freisthler, P. J. Gruenewald, E. Tebben, K. S. McCarthy, and J. P. Wolf, "Understanding at-the-moment stress for parents during COVID-19 stay-at-home restrictions," *Social Science & Medicine,* vol. 279, p. 114025, 2021.

[154] B. Muralidhar, D. K. Prasad, and M. Rao, "Association among Remote Working Concerns and Challenges on Employee Work-Life Balance: An Empirical Study Using Multiple Regression Analysis with Reference to International Agricultural Research Institute, Hyderabad," *International Journal of Advanced Research in Engineering and Technology,* vol. 11, no. 6, 2020.

[155] M. J. Zawadzki *et al.*, "Understanding stress reports in daily life: A coordinated analysis of factors associated with the frequency of reporting stress," *Journal of Behavioral Medicine,* vol. 42, pp. 545-560, 2019.

[156] M. R. Askari, M. Abdel-Latif, M. Rashid, M. Sevil, and A. Cinar, "Detection and classification of unannounced physical activities and acute psychological stress events for interventions in diabetes treatment," *Algorithms,* vol. 15, no. 10, p. 352, 2022.

[157] A. Anusha, J. Jose, S. Preejith, J. Jayaraj, and S. Mohanasankar, "Physiological signal based work stress detection using unobtrusive sensors," *Biomedical Physics & Engineering Express,* vol. 4, no. 6, p. 065001, 2018.

[158] S. Ray, "A quick review of machine learning algorithms," in *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 2019: IEEE, pp. 35-39.

[159] H. Song, I. Triguero, and E. Özcan, "A review on the self and dual interactions between machine learning and optimisation," *Progress in Artificial Intelligence,* vol. 8, no. 2, pp. 143-165, 2019.

[160] W. Alawad, M. Zohdy, and D. Debnath, "Tuning hyperparameters of decision tree classifiers using computationally efficient schemes," in *2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, 2018: IEEE, pp. 168-169.

[161] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing,* vol. 415, pp. 295-316, 2020.

[162] D. Bertsimas and J. Dunn, "Optimal classification trees," *Machine Learning,* vol. 106, pp. 1039-1082, 2017.

[163] O. Rahmati, M. Avand, P. Yariyan, J. P. Tiefenbacher, A. Azareh, and D. T. Bui, "Assessment of Gini-, entropy-and ratio-based classification trees for groundwater potential modelling and prediction," *Geocarto International,* vol. 37, no. 12, pp. 3397-3415, 2022.

[164] P. Guleria, N. Thakur, and M. Sood, "Predicting student performance using decision tree classifiers and information gain," in *2014 International conference on parallel, distributed and grid computing*, 2014: IEEE, pp. 126-129.

[165] M. T. Ramakrishna, V. K. Venkatesan, I. Izonin, M. Havryliuk, and C. R. Bhat, "Homogeneous Adaboost Ensemble Machine Learning Algorithms with Reduced Entropy on Balanced Data," *Entropy,* vol. 25, no. 2, p. 245, 2023.

[166] J. Alexander, M. Sovakova, and G. Rena, "Factors affecting resting heart rate in free-living healthy humans," *Digital Health,* vol. 8, p. 20552076221129075, 2022.

[167] J. Yamaguchi *et al.*, "Factors affecting home-measured resting heart rate in the general population: the Ohasama study," *American journal of hypertension,* vol. 18, no. 9, pp. 1218-1225, 2005.

[168] C. Maslach and M. P. Leiter, *The truth about burnout: How organizations cause personal stress and what to do about it*. John Wiley & Sons, 2008.

[169] G. P. Clarkson and G. P. Hodgkinson, "What can occupational stress diaries achieve that questionnaires can't?," *Personnel Review,* 2007.

[170] Y. S. Scharp, K. Breevaart, and A. B. Bakker, "Using playful work design to deal with hindrance job demands: A quantitative diary study," *Journal of Occupational Health Psychology,* vol. 26, no. 3, p. 175, 2021.

[171] A. Kaduk, K. Genadek, E. L. Kelly, and P. Moen, "Involuntary vs. voluntary flexible work: Insights for scholars and stakeholders," *Community, Work & Family,* vol. 22, no. 4, pp. 412-442, 2019.

[172] M. C. Bakhuys Roozeboom, R. Schelvis, I. L. Houtman, N. M. Wiezer, and P. M. Bongers, "Decreasing employees' work stress by a participatory, organizational level work stress prevention approach: a multiple-case study in primary education," *BMC Public Health,* vol. 20, no. 1, pp. 1-16, 2020.

[173] L. Toussaint *et al.*, "Effectiveness of progressive muscle relaxation, deep breathing, and guided imagery in promoting psychological and physiological states of relaxation," *Evidence-Based Complementary and Alternative Medicine,* vol. 2021, 2021.

# Appendix A- Email template for requesting data from other researchers

Dear [Name of corresponding author on relevant paper]

My name is Matthew Harper and I am a PhD student in the Department of Computer Science at Liverpool John Moores University. The aim of my project is to develop a wearable system which can detect agitation/difficulties in people with dementia by tracking physiological indicators, and trigger context-appropriate digital interventions (meaning the system will need to be context-aware).

In your paper {Name of the relevant paper], you [brief discussion of the relevant paper's methodology].

Would it be possible for you to share [description of the physiological data which is being requested]? I would of course reference and acknowledge you in any future publications for work in which I used and built upon your data.

Thanks

Matt Harper

## Appendix B- Table of attributes and extracted and engineered features

| Attribute | Feature name | Feature abbreviation | Description |
|---|---|---|---|
| DateTime | Time of day | TimeOfDay | The time at which the session began. |
| | Day of week | DayOfWeek | The day of the week the session occurred on |
| Environment | N/A | N/A | The environment in which the session occurred |
| Activity | N/A | N/A | The activity performed during the session |
| Stress | Stress level | Stress_Level | The level of stress, either high (above the median) or low (below the median), experienced during the session. |
| Productivity | N/A | N/A | How productive the participant felt during the session. |
| Length of session data files | Duration | N/A | The duration of the work session. |
| Electrodermal activity (EDA) | Mean EDA value | EDA_MEAN | The mean value of the EDA data. |
| | Maximum EDA value | EDA_MAX | The maximum value of the EDA data. |
| | Minimum EDA value | EDA_MIN | The minimum value of the EDA data |

| | | | |
|---|---|---|---|
| | Median EDA value | EDA_MED | The median value of the EDA data |
| | Variance of EDA | EDA_STD | The variance of the EDA data |
| | Range of EDA values | EDA_RANGE | The difference between the greatest EDA value and the lowest EDA value |
| EDA tonic component (TONIC) | Mean value of the tonic component | MEAN_TONIC | The mean value of the tonic component of the EDA |
| | Variance of the tonic component | STD_TONIC | The variance of the tonic component of the EDA |
| EDA phasic component (PHASIC) | Mean value of the phasic component | MEAN_PHASIC | The mean value of the phasic component of the EDA |
| | Variance of the phasic component | STD_PHASIC | The variance of the phasic component of the EDA |
| | Number of skin conductance responses per minute | NUM_SCRs_PERMIN | The number of peaks in the phasic component of the EDA which correspond to an activation of the ANS. |
| | Sum of the magnitude of skin conductance responses per minute | SCR_MAG_NORMALIZED | The summed magnitude of the peaks in the phasic component of the EDA which correspond to an activation of the ANS. |

| | | | |
|---|---|---|---|
| | Sum of the duration of skin conductance responses per minute | SCR_DUR_NORMALISED | The summed duration of peaks in the phasic component of the EDA which correspond to an activation of the ANS. |
| Heart rate (HR) | Mean HR value | HR_MEAN | The mean value of the HR data |
| | Maximum HR value | HR_MAX | The maximum value of the HR data |
| | Minimum HR value | HR_MIN | The minimum value of the HR data |
| | Variance of HR | HR_STD | The variance of the HR data |
| | Range of HR values | HR_RANGE | The difference between the greatest HR value and the lowest HR value |
| Skin temperature (ST) | Mean ST value | ST_MEAN | The mean value of the ST data |
| | Maximum ST value | ST_MAX | The maximum value of the ST data |
| | Minimum ST value | ST_MIN | The minimum value of the ST data |
| | Median ST value | ST_MED | The median value of the ST data |
| | Variance of ST | ST_STD | The variance of the ST data |
| | Range of ST values | ST_RANGE | The difference between the greatest ST value and the lowest ST value |

| | | | |
|---|---|---|---|
| Blood volume pulse (BVP) | Mean BVP value | BVP_MEAN | The mean value of the BVP data |
| | Maximum BVP value | BVP_MAX | The maximum value of the BVP data |
| | Minimum BVP value | BVP_MIN | The minimum value of the BVP data |
| | Median BVP value | BVP_MED | The median value of the BVP data |
| | Variance of BVP | BVP_STD | The variance of the BVP data |
| Accelerometer (ACC) | Mean ACC value from the X axis | ACC_X_MEAN | The mean value of the accelerometer X axis data |
| | Maximum ACC value from the X axis | ACC_X_MAX | The maximum value of the accelerometer X axis data |
| | Minimum ACC value from the X axis | ACC_X_MIN | The minimum value of the accelerometer X axis data |
| | Median ACC value from the X axis | ACC_X_MED | The median value of the accelerometer X axis data |
| | Variance of ACC from the X axis | ACC_X_STD | The variance of the accelerometer X axis data |
| | Range of ACC values from the X axis | ACC_X_RANGE | The difference between the greatest accelerometer X axis value and the lowest accelerometer X axis value |

| | | | |
|---|---|---|---|
| | Mean ACC value from the Y axis | ACC_Y_MEAN | The mean value of the accelerometer Y axis data |
| | Maximum ACC value from the Y axis | ACC_Y_MAX | The maximum value of the accelerometer Y axis data |
| | Minimum ACC value from the Y axis | ACC_Y_MIN | The minimum value of the accelerometer Y axis data |
| | Median ACC value from the Y axis | ACC_Y_MED | The median value of the accelerometer Y axis data |
| | Variance of ACC from the Y axis | ACC_Y_STD | The variance of the accelerometer Y axis data |
| | Range of ACC values from the Y axis | ACC_Y_RANGE | The difference between the greatest accelerometer Y axis value and the lowest accelerometer Y axis value |
| | Mean ACC value from the Z axis | ACC_Z_MEAN | The mean value of the accelerometer Z axis data |
| | Maximum ACC value from the Z axis | ACC_Z_MAX | The maximum value of the accelerometer Z axis data |
| | Minimum ACC value from the Z axis | ACC_Z_MIN | The minimum value of the accelerometer Z axis data |

| | | | |
|---|---|---|---|
| | Median ACC value from the Z axis | ACC_Z_MED | The median value of the accelerometer Z axis data |
| | Variance of ACC from the Z axis | ACC_Z_STD | The variance of the accelerometer Z axis data |
| | Range of ACC values from the Z axis | ACC_Z_RANGE | The difference between the greatest accelerometer Z axis value and the lowest accelerometer Z axis value |
| | Mean ACC value from the summed axes | ACC_SUM_MEAN | The mean value of the summed accelerometer axes data |
| | Maximum ACC value from the summed axes | ACC_SUM_MAX | The maximum value of the summed accelerometer axes data |
| | Minimum ACC value from the summed axes | ACC_SUM_MIN | The minimum value of the summed accelerometer axes data |
| | Median ACC value from the summed axes | ACC_SUM_MED | The median value of the summed accelerometer axes data |
| | Variance of ACC from the summed axes | ACC_SUM_STD | The variance of the summed accelerometer axes data |

| | | | |
|---|---|---|---|
| | Range of ACC values from the summed axes | ACC_SUM_RANGE | The difference between the greatest summed accelerometer axes value and the lowest summed accelerometer axes value |
| Interbeat interval (IBI) | Mean IBI value | IBI_MEAN | The mean value of the IBI data |
| | Maximum IBI value | IBI_MAX | The maximum value of the IBI data |
| | Minimum IBI value | IBI_MIN | The minimum value of the IBI data |
| | Variance of IBI | IBI_STD | The variance of the IBI data |
| | Root mean square of the IBI values | IBI_RMS | The square root of the sum of the squares of the successive intervals between heart beats |
| | Number of IBIs varying by more than 50 milliseconds | NN50 | The number of interbeat intervals which differ from the proceeding interval by more than 50 milliseconds |
| | Percentage of IBIs varying by more than 50 milliseconds | pNN50 | The percentage of interbeat intervals which differ from the proceeding interval by more than 50 milliseconds |

| | | | |
|---|---|---|---|
| | Number of IBIs varying by more than 25 milliseconds | NN25 | The number of interbeat intervals which differ from the proceeding interval by more than 25 milliseconds |
| | Percentage of IBIs varying by more than 25 milliseconds | pNN25 | The percentage of interbeat intervals which differ from the proceeding interval by more than 25 milliseconds |

# Appendix C- List of relevant publications by researcher

Harper, Matthew, et al. "Data Science Techniques to Support Prediction, Diagnosis and Recode Treatment of Alzheimer's Disease." *2019 12th International Conference on Developments in eSystems Engineering (DeSE)*. IEEE, 2019.

Harper, Matthew, and Fawaz Ghali. "A Systematic review of wearable devices for tracking physiological indicators of Dementia related difficulties." *2020 13th International Conference on Developments in eSystems Engineering (DeSE)*. IEEE, 2020.

Harper, Matthew, et al. "Challenges in data capturing and collection for physiological detection of dementia-related difficulties and proposed solutions." *Intelligent Computing Theories and Application: 17th International Conference, ICIC 2021, Shenzhen, China, August 12–15, 2021, Proceedings, Part III 17*. Springer International Publishing, 2021.

Harper, Matthew, et al. "Review of methods for data collection experiments with people with dementia and the impact of COVID-19." *Intelligent Computing Theories and Application: 17th International Conference, ICIC 2021, Shenzhen, China, August 12–15, 2021, Proceedings, Part III 17*. Springer International Publishing, 2021.

Harper, Matthew, and Fawaz Ghali. "Roles of caregivers in physiological data collection experiments with people with dementia and mitigating the impacts of COVID-19." *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*. IEEE, 2021.

Harper, Matthew, Fawaz Ghali, and Wasiq Khan. "Comparison of Subjective and Physiological Stress Levels in Home and Office Work Environments." *Intelligent Computing Methodologies: 18th International Conference, ICIC 2022, Xi'an, China, August 7–11, 2022, Proceedings, Part III*. Cham: Springer International Publishing, 2022.