

# A Study of Methods in Computational Psychophysiology for Incorporating Implicit Affective Feedback in Intelligent Environments

Deba Pratim Saha

Doctoral Dissertation submitted to the Faculty of the  
Virginia Polytechnic Institute and State University  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
in  
Computer Engineering

Richard B. Knapp, Co-Chair  
Thomas L. Martin, Co-Chair  
Denis Gracanin  
Steven R. Harrison  
Joseph L. Gabbard

May 04, 2018  
Blacksburg, Virginia

Keywords: Affective Computing, Computational Psychophysiology, Intelligent  
Environments, Context-Awareness, Signal Processing

Copyright© 2018, Deba P. Saha

A Study of Methods in Computational Psychophysiology for Incorporating Implicit  
Affective Feedback in Intelligent Environments

Deba Pratim Saha

**Abstract**

Technological advancements in sensor miniaturization, processing power and faster networks has broadened the scope of our contemporary compute-infrastructure to an extent that Context-Aware Intelligent Environment (CAIE)—physical spaces with computing systems embedded in it—are increasingly commonplace. With the widespread adoption of intelligent personal agents proliferating as close to us as our living rooms, there is a need to rethink the human-computer interface to accommodate some of their inherent properties such as multiple focus of interaction with a dynamic set of devices and limitations such as lack of a continuous coherent medium of interaction. A CAIE provides context-aware services to aid in achieving user’s goals by inferring their instantaneous context. However, often due to lack of complete understanding of a user’s context and goals, these services may be inappropriate or at times even pose hindrance in achieving user’s goals. Determining service appropriateness is a critical step in implementing a reliable and robust CAIE. *Explicitly* querying the user to gather such feedback comes at the cost of user’s cognitive resources in addition to defeating the purpose of designing a CAIE to provide automated services.

The CAIE may, however, infer this appropriateness *implicitly* from the user, by observing and sensing various behavioral cues and affective reactions from the user, thereby seamlessly gathering such user-feedback.

In this dissertation, we have studied the design space for incorporating users affective reactions to the intelligent services, as a mode of implicit communication between the user and the CAIE. As a result, we have introduced a framework named CAFFEINE, acronym for Context-aware Affective Feedback in Engineering Intelligent Naturalistic Environments. The CAFFEINE framework encompasses models, methods and algorithms establishing the validity of the idea of using a physiological-signal based affective feedback loop in conveying service appropriateness in a CAIE. In doing so, we have identified methods of learning ground-truth about an individual users affective reactions as well as introducing a novel algorithm of estimating a physiological signal based quality-metric for our inferences. To evaluate the models and methods presented in the CAFFEINE framework, we have designed a set of experiments in laboratory-mockups and virtual-reality setup, providing context aware services to the users, while collecting their physiological signals from wearable sensors. Our results provide empirical validation for our CAFFEINE framework, as well as point towards certain guidelines for conducting future research extending this novel idea. Overall, this dissertation contributes by highlighting the symbiotic nature of the subfields of Affective Computing and Context-aware Computing and by identifying models, proposing methods and designing algorithms that may help accentuate this relationship making future intelligent environments more human-centric.

A Study of Methods in Computational Psychophysiology for Incorporating Implicit  
Affective Feedback in Intelligent Environments

Deba Pratim Saha

**GENERAL AUDIENCE ABSTRACT**

Physical spaces containing intelligent computing agents have become an increasingly commonplace concept. These systems when populating a physical space, provides intelligent services by inferring user's immediate needs, they are called intelligent environments. With this widespread adoption of intelligent systems, there is a need to design computer interfaces that focuses on the human user's responses. In order for this service-delivery interaction to feel natural, these interfaces need to sense a user's disapproval of a wrong service, without the user actively indicating so. It is imperative that *implicitly* inferring a user's disapproval of a service by observing and sensing various behavioral cues from the user, will help in making the computing system cognitively disappear into the background.

In this dissertation, we have studied the design space for incorporating user's affective reactions to the intelligent services, as a mode of implicit communication between the user and the intelligent system. As a result, we have introduced an interaction framework named CAFFEINE, acronym for Context-aware Affective Feedback in Engineering Intelligent Naturalistic Environments. The CAFFEINE framework encompasses models, methods and algo-

rithms exploring the validity of the idea of using physiological signal based affective feedback in intelligent environments. To evaluate the models and algorithms, we have designed a set of experimental protocols and conducted user studies in virtual-reality setup. The results from these user studies demonstrate the feasibility of this novel idea, in addition to proposing new methods of evaluating the quality of underlying physiological signals. Overall, this dissertation contributes by highlighting the symbiotic nature of the subfields of Affective Computing and Context-aware Computing and by identifying models, proposing methods and designing algorithms that may help accentuate this relationship making future intelligent environments more human-centric.

... *dedicated to all my teachers, mentors,*  
ma *and* baba ...

# Acknowledgments

I would like to take this opportunity to extend my heartfelt gratitude to my mentors Dr. Ben Knapp and Dr. Tom Martin for their continued encouragement and guidance for completing this dissertation work. Their influence in shaping this work, and therefore me, has been akin to parents, albeit academic parents. Discussion sessions with them have shown me a process for clear and structured thinking, even in informal settings. These sessions have constantly inspired me to explore unconventional ideas, wander in novel uncharted domains, while also anchoring the thoughts to never lose sight of the bigger picture. It has been my utmost privilege to work with both of them.

To Dr. Martin, it has been wonderful being in your constant guidance for the past seven years, right from my first day at Virginia Tech in your Wearables class, till the day I graduated on May the Fourth Two-thousand Eighteen! Thank you for always encouraging me to explore new perspectives, helping me organize my approach towards research, and for constantly working with me and inspire me to strive for the better.

To Ben, thank you for always helping me articulate my thoughts clearly. Your words of encouragement inspiring me to push a little more has helped me revisit my work every-time with a renewed perspective and arrive at meaningful results. I would most certainly miss the interactions we have had over the past several years, it was fun sailing with you on the ICAT ship.

To my dissertation committee Dr. Gracanin, Dr. Harrison and Dr. Gabbard, thank you for your probing questions and guidance throughout this work which has helped me arrive at a complete picture of a framework presented in this work. Thank you Dr. Gracanin for being available for quick discussions in Sandbox which has helped me collect valuable feedback regarding this work and helped me improve the final outcome.

I owe a great deal to the Institute for Creativity, Arts and Technology (ICAT) at Virginia Tech, for providing me with numerous unique opportunities for collaborating across disciplines as varied as architecture, psychology, visual-arts, music, science and engineering, while also showcasing our projects at the Smithsonian Museum. It was a pleasure working with you Melissa Wyers, Holly Williams, Tanner Upthegrove, Zach Duer, Adam Soccolich, Phyllis Newbill, Lisa Jansen, Liz Scharman, Aki Ishida, Brook Kennedy; especially thank you Melissa and Tanner for accommodating all the scheduling requests. It was wonderful sharing

the ever-evolving ‘Sandbox’ and ‘Create’ spaces and bumping into amazing colleagues Reza Tasooji, Run Yu, Jason Forsyth, Manpreet Hora, Kari Zacharias, Michael Stewart, Rabih Younes to name a few. Thank you Zach and Reza for your kind help with my VR experiment setup in the Cube. A shout-out goes to Brennon Bortz for the numerous technical discussions, and collaborations on arts and technology projects while working at ICAT. I would also like to thank Brennon for sharing his python implementation of Ledalab tool with our research group, which was used to compile results in two of my experiments. My heartfelt thanks to ICAT for financially supporting me throughout my doctoral studies.

To my wonderful friends - Prithwish, Saurav, Arijit, Rajni, Sayantan, Abhishek, Sagar, Sanchit, Anupam, Almas, Shubhajit, Sidhartha, Jitish, Devin, Kushal, Anant, Debopam, and many many more, you guys are awesome! Your presence has helped me sail through this entire period, thank you all for being around me and lending your patient ears whenever I needed you all.

Thanking your family is never enough. To *Ma* and *Baba*, this work would not have been possible without your constant encouragement, support, and blessings. *Loho pronaam*. To *Dada* and *Boudi*, thank you for your support and understanding. To *Sonu*, I have missed your precious childhood, but I promise you will find a playful uncle in me.

To Sanchita, you are incredible. All I can say is I couldn’t have done this alone without your indomitable patience and belief in our shared dreams. I am looking forward to exploring the future and getting old together.

Deba Pratim Saha,  
Blacksburg, Virginia.  
May the 4th, 2018.



# Contents

<b>Contents</b>	<b>ix</b>
<b>List of Figures</b>	<b>xiv</b>
<b>List of Tables</b>	<b>xvi</b>
<b>List of Relevant Publications</b>	<b>xviii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	4
1.1.1 Context-aware Intelligent Environments . . . . .	5
1.1.2 Role of Affective Computing . . . . .	6
1.1.3 Technostress in Computer Interaction . . . . .	7
1.1.4 Scheme for Human-Centric Intelligent Environment . . . . .	9
1.2 Research Questions and Contribution . . . . .	12
1.2.1 Research Questions (RQ) . . . . .	12
1.2.2 Contributions . . . . .	14
1.3 Document Organization . . . . .	15
<b>2 Background and Related Research</b>	<b>16</b>
2.1 Interfaces for Intelligent Environments . . . . .	17
2.1.1 Naturalistic Interactions in Intelligent Environments . . . . .	18
2.1.2 Affective Human-Computer Interaction . . . . .	20

2.1.3	Affective Interfaces for Intelligent Environments . . . . .	22
2.2	Context Definition: Role of Affect and Intentions . . . . .	25
2.2.1	Relevance to our Research . . . . .	26
2.2.2	Definition and Components of Context . . . . .	27
2.2.3	User Intentions in the Context Model . . . . .	28
2.2.4	Interdependence of Affect and Intentions . . . . .	30
	2.2.4.1 Process of Cognitive Appraisal . . . . .	31
	2.2.4.2 Process of Coping . . . . .	32
2.3	Technostress in Computer Interactions . . . . .	34
2.3.1	Neurobiology of Stress . . . . .	36
2.3.2	Psychophysiological Indicators of Technostress . . . . .	41
2.3.3	Technostress as Negative Feedback Signal . . . . .	42
2.3.4	Individuality of Stress Reaction . . . . .	43
2.3.5	Moderators of Technostress Reactions in Humans . . . . .	44
	2.3.5.1 Gender and Stress Reactions . . . . .	44
	2.3.5.2 Role of Personality in Technostress . . . . .	46
2.4	Computational Psychophysiology . . . . .	47
2.4.1	Electro-Dermal Activity (EDA) . . . . .	48
	2.4.1.1 Physiological Basis of EDA . . . . .	50
	2.4.1.2 Algorithmic Analysis of EDA . . . . .	52
2.4.2	Heart-Rate and Heart-Rate Variability . . . . .	54
	2.4.2.1 Physiology of Cardiovascular System . . . . .	55
	2.4.2.2 Algorithmic Analysis of HRV . . . . .	58
2.4.3	Computational Recognition of Technostress . . . . .	62
	2.4.3.1 Notes on Dimensionality Reduction . . . . .	63
	2.4.3.1.1 Principal Component Analysis . . . . .	64
	2.4.3.2 Notes on Machine Recognition of Stressed States . . . . .	64
	2.4.3.2.1 Support-Vector Machines (SVM) . . . . .	65

2.4.3.3	Notes on Transfer Learning (TL) Paradigm . . . . .	66
2.4.3.3.1	TrAdaboost: Adaboost for Transfer Learning . . . . .	67
2.5	Implicit Feedback in Human-Building Interaction . . . . .	68
2.5.1	Perceived Comfort in Built Environments . . . . .	69
2.5.2	Frustration and Lack of Agency . . . . .	70
2.6	Conclusion . . . . .	71
<b>3</b>	<b>The CAFFEINE Framework</b>	<b>73</b>
3.1	Components of the CAFFEINE Framework . . . . .	74
3.1.1	User Feedback in a CAIE . . . . .	74
3.1.2	Why Technostress? Hint: <i>Explicit vs Implicit</i> Feedback . . . . .	76
3.1.3	Implicit Feedback using Technostress . . . . .	78
3.1.3.1	Rationale for Physiology based Implicit Feedback . . . . .	80
3.1.4	Interaction Schematic Overview . . . . .	81
3.2	Psychophysiological Profile Learning . . . . .	83
3.2.1	Paced-Stroop Test (PST) . . . . .	85
3.3	Physiological Signal based Quality-Metric . . . . .	87
3.3.1	EDA Signal Analysis and Impulse Response Function . . . . .	87
3.3.2	Towards Defining an EDA Quality Measure . . . . .	88
3.4	CAFFEINE Framework: Bringing It All Together . . . . .	90
3.5	Conclusion . . . . .	92
<b>4</b>	<b>Experimental Setup and Results</b>	<b>93</b>
4.1	Experiment One: Warehouse Picker Experiment . . . . .	96
4.1.1	Experimental Protocol . . . . .	96
4.1.1.1	Order Picking Experiment (OPE) Phase . . . . .	97
4.1.1.2	Paced-Stroop Test (PST) Phase . . . . .	98
4.1.1.3	Picker Experiment Hardware Setup . . . . .	98
4.1.2	Picker Experiment Data Analysis . . . . .	99

4.1.2.1	Data Normalization and Preprocessing . . . . .	99
4.1.2.2	Feature Extraction and Feature Reduction . . . . .	99
4.1.2.3	Support Vector Machines (SVM) . . . . .	100
4.1.3	Picker Experiment Results . . . . .	101
4.1.3.1	SVM Kernel Selection . . . . .	102
4.1.3.2	Performance Evaluation Criteria . . . . .	103
4.1.3.3	Model Performance Evaluation . . . . .	105
4.1.4	Picker Experiment Discussion . . . . .	106
4.1.5	Picker Experiment Conclusion . . . . .	107
4.2	Experiment Two: Virtual SuperMarket with Navigation Assist Experiment .	108
4.2.1	Intelligent Supermarket in Virtual Reality Setup . . . . .	109
4.2.1.1	Experimental Protocol in VR Setup . . . . .	109
4.2.2	Hardware Setup . . . . .	111
4.2.3	Virtual Supermarket Experiment Data Analysis . . . . .	112
4.2.3.1	Electrodermal Activity (EDA) Analysis . . . . .	112
4.2.3.2	Integrated Phasic Response (IPR) Analysis . . . . .	112
4.2.4	EDA Analysis Results . . . . .	113
4.2.5	Virtual Supermarket Experiment Discussion . . . . .	115
4.2.6	Virtual Supermarket Experiment Conclusion . . . . .	116
4.3	Experiment Three: Defining an Inference Quality Metric using Physiological Signals . . . . .	117
4.3.1	Modified Experimental Protocol . . . . .	118
4.3.1.1	Baseline Sonic-Impulse Phase . . . . .	118
4.3.1.2	VR SuperMarket Phase . . . . .	118
4.3.2	Modified Hardware Setup . . . . .	120
4.3.3	Quality Metric Methods and Analysis . . . . .	120
4.3.3.1	EDA Signal based Quality Metric . . . . .	120
4.3.3.1.1	Point-wise Error Computation Algorithm . . . . .	121
4.3.3.1.2	Parametric Error Computation Algorithm . . . . .	122

4.3.4	Quality Metric Results and Discussion . . . . .	122
4.3.4.1	EDA Signal based Parametric Quality Metric Results . . . . .	122
4.3.4.2	Interpreting the $Q_{eda}$ Quality Score . . . . .	124
4.3.5	Quality Metric Discussion . . . . .	126
4.3.6	Quality Metric Experiment Conclusion . . . . .	127
<b>5</b>	<b>Summary, Conclusions and Reflections</b>	<b>128</b>
5.1	Discussion on Research Questions . . . . .	130
5.1.1	Validating CAFFEINE Interaction Scheme . . . . .	130
5.1.2	Identifying Technostress Physiology . . . . .	131
5.1.3	Deriving A Quality Metric . . . . .	132
5.2	Observations and Reflections . . . . .	133
5.2.1	New Sensing Modalities and Instruments . . . . .	133
5.2.2	New Computational methods . . . . .	135
5.3	Future Applications for CAFFEINE Framework . . . . .	136
5.3.1	Novel Systems for Human-Building Interaction . . . . .	136
5.4	Concluding Remarks . . . . .	138
	<b>Bibliography</b>	<b>139</b>

# List of Figures

1.1	Unintended Interaction with an Automatic Door . . . . .	10
1.2	Unintended Interaction with a Smart Home . . . . .	11
2.1	Primary and Secondary Appraisal Processes . . . . .	35
2.2	SAM and HPA Axes Activation in Response to Stressors . . . . .	37
2.3	Chronology of Stress-Reaction . . . . .	40
2.4	Graphical Representation of Phasic SCR . . . . .	52
2.5	QRS-Complex waveform in an ECG plot . . . . .	56
2.6	Example of frequency spectrum of an HRV time-series . . . . .	59
2.7	Example of a Poincaré Plot of an HRV Spectrum . . . . .	60
3.1	Intelligent services provided by a CAIE . . . . .	78
3.2	Affect generation in a user on appraisal of CAIE events . . . . .	79
3.3	Interaction-process Model of Service Delivery . . . . .	82
3.4	Screenshot of Incongruent Phase of Paced Stroop Test . . . . .	85
3.5	Schematic of $Q_{eda}$ Computation Process . . . . .	90
3.6	Affective Feedback in an Intelligent Environment . . . . .	91
4.1	Order Picking Experiment Setup . . . . .	97
4.2	User-wise $F_{\beta}$ -score Comparison For Case I, II and III . . . . .	106
4.3	Smart Supermarket with Navigation-Assist Service Setup . . . . .	110
4.4	EDA Decomposition using Ledalab . . . . .	114
4.5	Samples of normalized EDA responses to sonic impulse . . . . .	119

4.6	EDA Signal based quality metric computation . . . . .	121
4.7	Groupwise Per-Event IPR and Per-Event NSPP Analysis Plots . . . . .	125

# List of Tables

2.1	Generalized Positive and Negative Emotions in OCC Model . . . . .	32
2.2	Effects of Neural Activation . . . . .	38
2.3	Effects of SAM Activation . . . . .	39
2.4	Effects of HPAC Activation . . . . .	39
2.5	EDA Measures and Typical Values . . . . .	54
2.6	Conventional HRV Features . . . . .	58
2.7	Comparison of Various Transfer Learning approaches . . . . .	67
4.1	Comparison Table for Polynomial, Gaussian and Sigmoid Kernel Functions .	103
4.2	User-wise Confusion Matrix, Picker Experiment Case I . . . . .	104
4.3	User-wise Confusion Matrix, Picker Experiment Case II . . . . .	104
4.4	User-wise Confusion Matrix, Picker Experiment Case III . . . . .	105
4.5	User-wise integrated phasic response (IPR) . . . . .	113
4.6	User-wise $Q_{eda}$ measure, Cumulative IPR and Number of Phasic Peaks . . .	123
4.7	Per-Event IPR and Per-Event NSPP . . . . .	125



---

## Relevant Publication

---

- [1] **Saha, Deba Pratim**, Thomas L. Martin, and R. Benjamin Knapp. Towards Defining a Quality-Metric for Affective Feedback in an Intelligent Environment. In *IEEE International Conference on Pervasive Computing and Communications Workshops, 2018. PerCom Workshops'18*, Athens, Greece, March 2018.
- [2] **Saha, Deba Pratim**, R. Benjamin Knapp, and Thomas L. Martin. Affective Feedback in a Virtual Reality Based Intelligent Supermarket. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, UbiComp '17*, pages 646–653, New York, NY, USA, 2017. ACM.
- [3] **Saha, Deba Pratim**, Brennon C. Bortz, Wei Huang, Thomas L. Martin, and R. Benjamin Knapp. Affect-Aware Intelligent Environment Using Musical Cues as an Emotion Learning Framework. In *2016 12th International Conference on Intelligent Environments (IE)*, pages 178–181, September 2016.
- [4] **Saha, Deba Pratim**, Thomas L. Martin, and R. Benjamin Knapp. Towards incorporating affective feedback into context-aware intelligent environments. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 49–55, September 2015.
- [5] Brennon Bortz, Isabel Bradburn, Angela Scarpa, Martha Bell, **Saha, Deba Pratim**, Deanna Swain, and R. Benjamin Knapp. Emotion Regulation in Children with Autism through Music-based Biofeedback. In *4th International Conference on Music and Emotion (ICME4)*, Geneva, Switzerland, 2015.



# Chapter 1

## Introduction

---

*“In theory, there is no difference between theory and practice. But, in practice, there is.”*

— Jan L. A. van de Snepscheut

---

We spend the majority of our days immersed inside built-environments infused with networks of connected devices used for sensing the state of the system and often providing automated services. Modern buildings are functionally akin to large-scale interactive machines [1], and can be seen as extensions of computing devices and services that are progressively moving

away from the limited desktop scenario into our living environment, ever closer to the activities of our daily lives. Technological advancements in hardware sensor miniaturization, processing power, faster networks and their ever shrinking costs have pushed the contemporary computational infrastructure into our living rooms, closely following Mark Weiser’s overarching vision of “invisible, everywhere computing” [2]. Mark Weiser had sketched out a broad vision for Ubiquitous Computing as a human-centered computing platform in an information-rich and computationally-augmented reality [3, 4], envisioning computers to become part of the “*woodwork everywhere*”.

The collective term describing such a ubiquitous spread of computational infrastructure is *Context-aware Intelligent Environments (CAIE)*, which is a space where the ambient intelligent system “has contextual-awareness of a user’s current state and is capable of maintaining a consistent, coherent interaction across a number of heterogeneous smart devices” [5]. A CAIE needs to be sensible to identify situations *when* it should provide a service to the user and *how* to help them achieve their current goals while also preserving their privacy, safety and agency [6]. In the ubiquitous computing paradigm, similar to an intelligent building environment, a user interacts with a dynamic set of sensors and devices in the environment, possibly cohabitated by other fellow users [5]. From the human-computer interaction (HCI) perspective, for these pervasive technologies to become part of the “*woodwork everywhere*”, and thereby cognitively disappear completely, designers still need to address the nuances of human-environment interfaces to facilitate a seamless communication channel between the interacting agents, i.e. the user and the intelligent environment. In this article, we present CAfFEINE—Context-aware Affective Feedback in Engineering Intelligent Naturalistic Environments—a framework for incorporating implicit feedback from users enabling natural seamless interactions in ubiquitous computing infrastructure.

For a pervasive computational system to disappear from a user’s cognitive front, one of the

requirements is to acquire the ability to assimilate into their temporal workflow at all possible times, without being a hindrance to them. To achieve this, the environment may need to continuously and consistently infer a user's goals, intentions and instantaneous context by collecting sensory inputs, usually from the instrumented environment surrounding the user. One key aspect of an inference-based system is the probabilistic nature of *recognizing* a user's context and thereafter delivering *suitable* services. As a result, such an intelligent computational system may render itself to situations where it is delivering services which may not align properly with a user's current needs for achieving his goals. This lack of proper understanding of a user's instantaneous expectations from the system may arise from various reasons ranging from dynamic user preferences to improper modeling of user's context, goals or intentions (may be due to the lack of complete information). Although this may be alleviated by provisioning for some ways to *explicitly* ask the user about the appropriateness of the current provided service. In a way, however, this would defeat the purpose of having an inference-based context-aware pervasive computational system in the first place, if it has to stop and ask the user each time after delivering the services. If, on the other hand, the pervasive computational system could infer a user's *(dis)approval* about the service's (in)appropriateness from their behavioral or physical *cues*, the need for stopping and asking the user could be completely eliminated.

This dissertation intends to explore the various design parameters influencing decisions for incorporating such an *implicit* feedback loop in a pervasive computational system. Specifically, we will present our interaction scheme that a user is envisioned to follow in our system, and present a battery of experiments to validate various aspects of this interaction scheme. In the next section, we will present the overall motivation driving our research.

## 1.1 Motivation

During communication, more than 90% of messages are conveyed by non-verbal implicit modes [7, 8]. Thus, for attaining seamless communication, the implicit channel must be employed as it enables *convergence* in communication between interacting agents [9]. Affective Computing provides techniques to implicitly infer psychological states such as technostress from user's physiological signals such as electrodermal activity (EDA), heart-rate variability (HRV), skin-temperature (ST), electromyography (EMG) etc. [10]. Similar ideas of implicitly learning a user's disapproval of an intelligent service in a multi-turn interaction, has been widely used in information-retrieval domain to model search relevance [11, 12]. Specifically, search-relevance is modeled by predicting searcher frustration[13] and creating personalized models of search satisfaction [14]. Physiological signals have recently been explored for modeling search relevance [15, 16]. An implementation for using physiological signals in mobile search improvement has recently been patented by Google [17]. Thus, we see that service-relevance feedback by modeling user's frustration in personal computing scenario has been extensively studied and implemented. However, it has not been tried in pervasive computing scenario, specifically a CAIE, which is the main goal of our work.

In this section, we will briefly describe some concepts of a pervasive computational system embedded with various sensors to infer a user's context, as well as some nuances of interaction with such an expanse of distributed computational infrastructure. We will then shortly discuss the role of a relatively new branch of computational intelligence, namely Affective Computing (AC), in inferring a user's (dis)approval about (in)appropriateness of a service. Finally we will discuss how, if at all, the inference from an AC system can be used to incorporate the *implicit* feedback loop which we alluded-to in the introductory paragraphs.

### 1.1.1 Context-aware Intelligent Environments

An *Intelligent Environment (IE)* has been defined by Augusto et al. as a space in which “the actions of numerous networked controllers, each controlling different aspects of an environment, is orchestrated by self-programming pre-emptive processes (e.g., intelligent software agents) in such a way as to create an interactive holistic functionality that enhances occupant’s experiences” [6]. In this article, Augusto et al. argue that an IE should be sensible to identify situations *when* it should provide a service to the user and *how* to help them achieve their current goals while also preserving their privacy, safety and autonomous behavior. Situational information relevant to the interaction between a user and an application has been defined as *context* [18]. A *context-aware intelligent environment (CAIE)*, thus, is a space where the ambient intelligent system “has *contextual*-awareness of a user’s current state and is capable of maintaining a consistent, coherent interaction across a number of heterogeneous smart devices” [5].

In the *ubiquitous computing* paradigm, “a user maintains an ongoing interaction with a dynamic set of devices in the environment, including many of which he may not even be aware” [5]. Some of these devices are usually instrumented with sensors to provide a situational awareness snapshot of the interacting agent (i.e., the user), often a plurality of them in the same environment, enabling richer forms of interaction. However, interaction design for this paradigm is significantly challenging compared to traditional computing infrastructure (such as desktop computers, tabs, pads and boards [19]) due to their inherent differences such as (i) the lack of single focal point of interaction, (ii) dynamic set of interaction devices, (iii) potential of multiple simultaneous users each engaged in diverse set of modalities, etcetera [5]. We argue that these pose a pressing need to address these issues, while affective computing technologies (see next section) presents some promising opportunities.

### 1.1.2 Role of Affective Computing

As per the seminal article by Cowie et al. [9], *convergence* is a feature of human communication wherein the agents who are in sympathy (or intend to portray so), converge vocally or via other modes of communication. Cowie et al. categorize communication into *explicit* and *implicit* channels, wherein they hypothesize that the implicit channel imparts meaning and/or provides context to what is explicitly conveyed via words or actions. In this categorization, the *explicit* channel is surmised to consist of spoken words, actions and enacted facial expressions, whereas *implicit* channel may be thought as voice tonality, body-gestures, social gestures arising from underlying psychological states of the user. Mehrabian proposed a 7%-38%-55% rule for non-verbal communication which states that only 7% of messages are conveyed by the spoken words, the rest (more than 90%) are by voice tonality and body language [7, 8]. For technology to go into the “*woodwork everywhere*”, a seamless communication needs to be established between the interacting agents viz. the human user and the IE. This demands attaining the state of *convergence* in interaction between the agents by establishing an understanding of the *implicit channel* of communication.

Over the past decade, a new class of intelligent computing systems has emerged that strives to recognize, understand and actively influence human emotions and are collectively termed as *Affective Computing* (AC) systems. In a recent work, Thompson et al. note that “although computers will not actually experience emotions in the same way as humans would, the quality of interaction has been shown to improve even if the system *appears* to do so” [8]. AC systems/agents showing empathy have been successfully shown to improve overall usability of a system [8], sometimes helping reduce user’s experienced stress even when performing frustrating tasks, making the system more human-centric [20, 8]. As Thompson et al. point out, the ability to detect/measure stress due to difficulty caused by a system may allow



developers to pinpoint problems and improve the system by enabling it to respond in a more natural or realistic way [8]. Such AC systems have been successfully implemented in diverse applications such as robotic personas [21], learning companions [22], affective tutors [23, 24, 25], affective games [26, 27, 28, 29], psychotherapy (such as in autism) using wearable devices [30, 31, 32] to name a few. These systems have shown promising results in computationally understanding human psychological states arising from complex interactions, with various agents. In a recent thematic issue on “Affect-aware Ubiquitous Computing”, Doctor et al. state that integrating such sensing modalities into IE will result in a richer set of interaction, influencing how we interact with the IE system [33]. We hypothesize that such AC systems can be used to implement the service-relevance feedback to the IE, employing *implicit channel* of communication.

### 1.1.3 Technostress in Computer Interaction

Technostress has been defined by psychologist Craig Brod as “a modern disease of adaptation caused by an inability to cope with computer technologies” [34, 35]. In simpler terms, technostress is the perception of hassles due to system failures e.g. computer crash, response delay or a demanding learning-curve of new modalities during interaction with technology. It is a psychological as well as biological stressor, which results in activation of many biological subsystems. A leading cause of technostress is “achievement stress” which is observed to be heightened in system failures during time-pressured tasks, i.e., tasks having hard-deadlines associated with them [35, 36]. Following our on going discussion, technostress generated due to inappropriate service delivered from a CAIE can be used as a means of detecting user’s reaction. Thus, detecting and inferring technostressed states, could fit immaculately in the *implicit* feedback mechanism discussed earlier. A more detailed discussion on technostress and its physiological correlates will be presented in Chapter 2.

Design of environments with computational intelligence embedded in them, has hitherto been focused towards automating services (automated services) and reacting to a subset of users context (smart services), as discussed in details in Section 2.1.3. From the design perspective, the above reasoning demands that the design of intelligent environments also needs to react to users affect (intelligent service). Affect and emotions are integral parts of human intelligence, and to attain convergence in communication between agents, the services need to understand and reciprocate human emotions. To elaborate this point further, imagine a smart service such as an automatic door, it automatically opens up sensing just the presence of a user, not knowing if the user actually wanted to pass through it. When these systems produce such an unwanted behavior (say due to, limitations in their sensing or wrong inferences on context), the user first reacts emotionally to the error by getting stressed or angry. As a system designer however, we are ignoring/losing a very rich signal i.e. users emotion, knowing which may inform the intelligent environment about the relevance of its service. Being able to understand and react to such emotions may also help personalise each environment to its dwellers, thereby enhancing convergence in communication and immersivity in interaction with the user.

In this dissertation, we will present our experiments designed to elicit technostress arising due to the mismatch between a user's expectation and the actual services provided by a CAIE. We will use physiological signals to capture these user-specific response patterns and use computational methods to recognize a specific aspect of this technostressed state, namely sympathetic activation. In our experiments, we will demonstrate that intelligent services that do not match a user's expectations, indeed produce higher sympathetic activation. We will also present an individualized EDA signal-shape based quality-metric computation method, that will be used to independently assess the usefulness of this technostress inference drawn from features depicting a high sympathetic activation.

### 1.1.4 Scheme for Human-Centric Intelligent Environment

Formally speaking, a direct outcome of the discussion above is a scheme of a system wherein a *service provider* CAIE, which has over time built a *knowledge profile* for a particular user, is capable of providing intelligent services by recognizing the user's present context consisting of their physical, environmental, social situations. The *implicit channel* of interaction will be provided by the *affect recognition* component of this system, which continuously monitors the user's reactions and predicts a service-relevance score. This prediction is used by the CAIE as a service-relevance feedback to model user's satisfaction with the services, and in turn will help to reconfigure the CAIE services to better suit the user's immediate needs.

The schema described above can be illustrated by two example scenarios described below, depicting the need for creating an affect-aware implicit feedback loop in a CAIE. *Scenario One* depicted in Figure 1.1 shows a simple interaction scenario, whereas *Scenario Two* depicted in Figure 1.2 shows a complex interaction between a user and a CAIE. These scenarios presented here are just for illustration purposes. Experiments described later-on in this dissertation, that are used to validate our interaction framework, are modelled on practical CAIE scenarios.

Human-centered computing (HCC) is a set of methodologies used in any field that uses computing systems that are intended to directly interact with humans [37]. HCC aims to integrate human-sciences (social, cognitive and affective) into notions of human-computer interaction (HCI). Thompson et al. [8] noted that for human-centered design “the main focus is that user's needs should inform system design” [38]. In this respect, our approach is a holistic human-centric approach as the emphasis is on understanding the user's non-verbal modes of communication in addition to sensing their physical and environmental context, to complete the service-relevance discovery loop thereby improving the system's usability.

**Scenario One**

*It's a cold winter day and Dave is walking in a hallway, past a proximity sensing automatic door. Even though Dave did not intend to go outside, the door opened up as it sensed his presence in front of the door. This was not an expected behavior of the system, although it is a very commonly occurring mis-triggering of the system. If the environment had an implicit feedback loop, it might have inferred Dave's intention and stopped the door from opening all the way.*

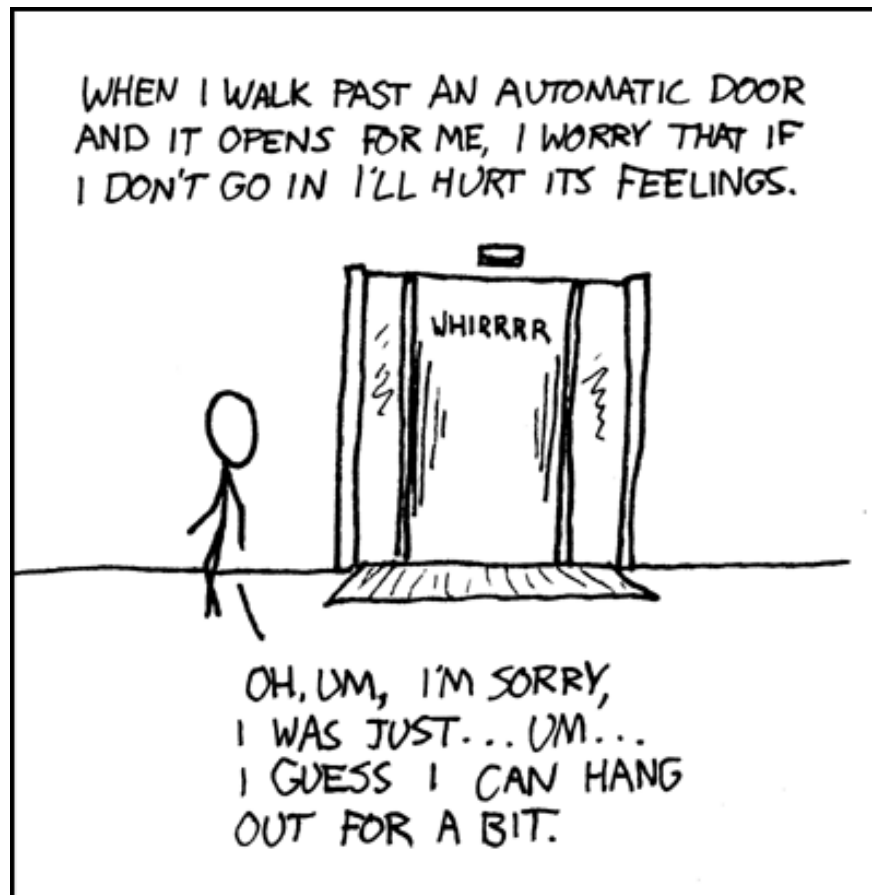


Figure 1.1: An example of an unintended interaction with an automatic door. For a humorous take, please visit <https://xkcd.com/175/>

**Scenario Two**

Brad and Tina are in an context-aware smart building, which has access to his personal planner and various other personal profiles. Brad is getting ready for one of his meetings. It is winter, and he picks up a jacket but is unsure of the weather outside. He vaguely asks the smart-building if the jacket would be enough for the planned outing. The pervasive computational intelligence decodes the semantic meaning of the query, infers that the user is asking about weather outside, checks his planner to see how long does he plan to be outside and replies in affirmative as well as cautions him to take an extra as the weather is going to get bad by the time he returns indoor after several other meetings. However, in doing so the building may have revealed too much information, which Brad may not have wanted his co-worker Tina to know. The smart-building was totally unaware of this social context. If the building was designed to be affect-aware, the computational system could have inferred Brad's preferences from his instantaneous affective reactions and may have stopped accordingly.

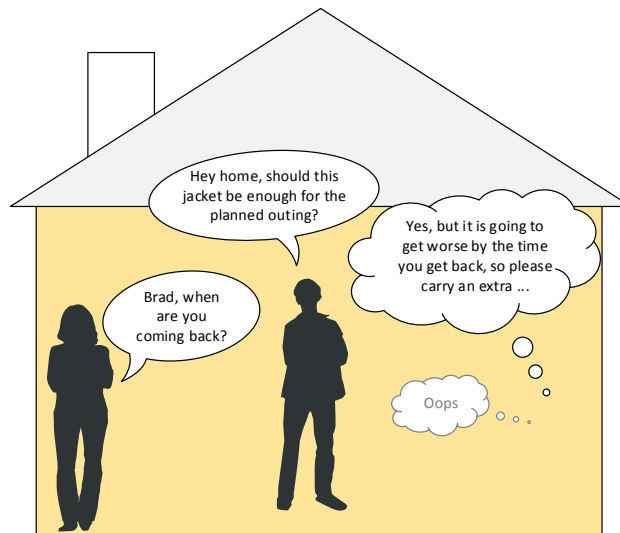


Figure 1.2: An example interaction within an IE, depicting the need for an implicit feedback.

## 1.2 Research Questions and Contribution

The discussion thus far can be summarized into an overarching design question, namely, *what are the methods in which affective computing techniques can be used to design an automatic service-appropriateness feedback in a real-life human-centric intelligent environment.* To answer this broad design question, we have broken it down into a set of research questions focusing on defining the model of interaction that supports such a service-relevance feedback, feasibility of such a feedback loop in a real-life CAIE, physiological underpinnings of such an affective feedback loop.

### 1.2.1 Research Questions (RQ)

***RQ1 : What is the interaction schema to effectively incorporate an implicit-feedback loop in a context-aware intelligent environment?***

Chapter 3 defines the interaction schematic and the interplay between various elements of the CAIE. For instance, we answer such questions as, how can we exploit a user's physiological responses to the system's services to create an *implicit* feedback loop, which modality of data collection used in AC is well suited for designing a dynamic human-centric interface in a CAIE and what are the trade-offs for choosing these modalities. In addition, what is the planned interaction schema and what is the chronology of the user's interaction under this new schema? How will the intelligent system learn emergent user behavior, or variants to the same behaviour? Which methods of ground truth collection for physiological profile learning are well-suited for a naturalistic CAIE?

***RQ2 : What are the basic design parameters, signal-features and methods for evaluation of the interaction scheme of a physiological-signal based affective***

***feedback loop in a CAIE.***

Chapter 4 experimentally validates the idea of incorporating an affective feedback loop in real life CAIE. Our hypothesis is that a computing system providing services which are not as desired by the user and is perceived to be a hindrance in achieving their time-critical goal, will induce *technostress* in them [39][35]. Our first experimental setup is an Order-Picking Experiment, mimicing the order-picking tasks used in large warehouses to fulfill customer orders (details in Chapter 4). We trained a classifier to detect changes in a user's physiological response whenever the system was made to malfunction. The results from this study empirically validates the idea that closing the service-relevance feedback loop is a worthy idea [40].

***RQ3 : Which parameters of the physiological signals are critical in performance improvement of a technostress based service-relevance feedback loop?***

*Technostress* is said to induce physiological and biological reactions, similar to cognitive stress [39, 35]. *Technostress* is a biological phenomenon, which produces heightened secretion of stress-hormone (e.g. cortisol, adrenaline) and detectable changes in patterns of physiological signals. Our next focus is to understand the underlying physiology of technostress as well as explore and design various computational methods to detect such a state in naturalistic intelligent environments. Towards this end, we conducted a second experiment in a Virtual Reality based Grocery Store setup and identified a few informative features from the Electrodermal Activity datastream that can be used to identify a group of services inducing technostress from another group of correct services [41]. This is described in Chapter 4.

***RQ4 : How to estimate the quality of the inference drawn from the physiological signals?***

Chapter 4 describes a method to estimate the quality of inference based on the comparison

between two impulse response function (IRF) shapes of Electrodermal Activity (EDA) signal: one recorded directly from the user while they listen to a sonic impulse and the other derived from the EDA decomposition framework. An empirical threshold can be set to define the quality of inference from the EDA measurement based on this score (see [42]). Such a quality-metric can be helpful in real-life implementation of the CAFFEINE framework.

### 1.2.2 Contributions

This dissertation makes two contributions to the field. First, it advances the discussion on the symbiotic relationship between *Affective Computing* and *Ubiquitous Computing* as posited in [33, 43]. We present a framework that deals with incorporating a physiological signal based implicit feedback loop for determining service-appropriateness in an context-aware intelligent environment. We believe that the interaction schematic presented in this work, is a main contribution to the field. Specifically, to the best of our knowledge, physiological signal based implicit feedback in a pervasive computing scenario used to ascertain service appropriateness has not been tried before.

Second, we apply a systematic approach in exploring the design space for employing a physiology-based inference of technostress, as a means of incorporating an implicit feedback loop as described above. We identified methods and algorithms that could be used for implementing such a framework. Additionally, we introduced a method to identify the quality of an inference based on certain features derived from an individual's physiological signals. We believe the experiments, the evaluation methods and the algorithmic approaches explored as a part of this work, will function as a baseline for further studies in this field.

In summary, this dissertation systematically explores the design space for incorporating a technostress-based implicit feedback loop and identifies methods for implementing them.



The algorithmic approaches used in our exploration may form a baseline method for future studies in this field.

### 1.3 Document Organization

In this current chapter, we have presented our overall motivation and an overview of the research questions that has been addressed in this dissertation. Chapter 2 will describe the state-of-the-art in affect-aware human-centric interfaces for CAIE research, psychophysiology of *technostress* and computational theories of emotion recognition. In this chapter we will also discuss the algorithmic approaches for a physiological profile learning framework. Chapter 3 will provide a brief description of our research methodology in relation to each of the research questions posed in Chapter 1. In Chapter 4, we will discuss our experimental setup for validation of the interaction scheme as well as the results. Chapter 5 presents a discussion on the conclusions and some possible avenues of future work.

## Chapter 2

# Background and Related Research

---

*“Information is a source of learning. But unless it is organized, processed, and available to the right people in a format for decision making, it is a burden, not a benefit.”*

— William Pollard

---

In the previous chapter, we have presented a short introduction to the motivation behind creating an implicit feedback loop for an intelligent environment. We have also discussed

the research questions that are addressed in this dissertation through which we have taken a systematic approach to define a computational framework for creating such a feedback loop. In this chapter, we will present the state of the current research in human-computer interfaces for CAIE. It is a widely held view that affective computing paradigm holds tremendous promise for transforming HCI design for CAIE [43]. We will start by briefly discussing various interfaces and types of interactions in an intelligent environment and what role can an affective computing system play in such interactions. Then, we will briefly discuss the envisioned role of *technostress* in interaction design for IE. Later on we will describe how computational models of affective agents based on motivational/appraisal theories of affective computing may inform our research on affective interfaces for CAIE. Lastly, we will discuss various computational methods used for feature extraction and learning models to be used in our research.

## 2.1 Interfaces for Intelligent Environments

Context-aware intelligent environments (CAIE) can be conceptualized as living spaces embedded with computational intelligence that provides companionship in our daily lives while cognitively disappears from our lives. Designing such a CAIE has been the persistent vision of computer scientists for nearly two decades, following Mark Weiser’s seminal article envisioning ubiquitous computing infrastructure that is characterized by “invisible, everywhere computing that does not live on a personal device of any sort, but is in the woodwork everywhere” [19]. The explosive growth of interconnected embedded systems, general purpose computers and intelligent devices—commonly termed as Internet of Things (IoT)—is an enabler of richer context awareness and tighter integration with users in their daily lives,

effectively shortening the service delivery pathways. Even with this tremendous progress, the model for interaction between the user and the computing infrastructure has always been *reactive*. However, an ideal context-aware system residing in the woodworks everywhere should *proactively* assist a user in their tasks where there is no need for the user to fetch appropriate assistance. In addition to not serving the user's needs, this passive service-delivery approach of the reactive interface creates potentially awkward social situations, thus, sometimes rendering itself as a liability for the user [44].

The present scope of context-aware applications leaves a lot to be desired, as evident from the opinion paper by Erickson [44] which highlights that the root of the problem lies in the fact that context awareness shown by humans is on a radically different paradigm as compared to that shown by computational machines. To quote him “people notice and integrate a vast range of cues, both *obvious* and *subtle*, and interpret them in the light of their previous experience to define their context. In contrast, current context-aware systems detect a very small set of cues, typically quantitative variations of the dimensions for which they have sensors.” The *obvious* and *subtle* cues that Erickson is talking about, are very similar to Cowie's *explicit* and *implicit* channels of human interactions, as previously described in Section 1.1.2 [9]. Cowie's claim that the implicit channel has been relatively less studied in the human-computer interactions (HCI) community, is evinced by Erickson's distinction between the paradigms of context-awareness used by humans and machines.

### 2.1.1 Naturalistic Interactions in Intelligent Environments

The pervasiveness of a CAIE not only poses challenges for user interface (UI) development due to factors described in Section 1.1.1, they also open up enormous possibilities of using a variety of naturalistic interfaces in the wild to design interaction. The possibility of physically

embodied interactions in a CAIE differentiates them from traditional personal computing. In a CAIE, the usual handling of a *thing* or a *prop* can provide a context of usage, thereby triggering a service. For example, when you sit in a chair to read a book, a computer vision system might detect the context and the lights might be automatically adjusted to your most preferred intensity setting based on your daily time of sleep, in order to aid you in your circadian rhythm [45]. Here, the chair and book became input devices, even though they are not directly used to control a computer. Not only as input, the embodied interaction via the *prop* may also take the form of a CAIE's smart output, such as flashing of the reading light to indicate some alarm [5]. Ishii et al. coined a term *phicons* (physical icons) [46], to describe a new class of connected *props* for CAIE, which may be used as input [46], output [46] or storage UI [47].

In the survey on interaction issues for IE, Shafer infers that an “automatic user behavior-learning interface”, akin to the *The Neural Network House* proposed by Mozer [48], is arguably the best UI for CAIE from the design point of view [5]. Such automatic behaviors can be programmed by the user or the manufacturer, taught by demonstration or learned by observation. As discussed briefly in the introductory paragraphs in Chapter 1, these learning-based context discovery systems have a probabilistic nature, and thus are prone to erroneous inferences. Hence, no matter what kind of interaction is designed into the CAIE, there must be a mechanism for the user to iterate with the service delivery system using an appropriateness feedback (which will tell the system if the service last received was as expected by the user), until their expectations are met vis-a-vis the service. Such an immediate feedback and provision of an *undo* mechanism for the user is considered as the first remedial measure for interaction issues in a CAIE by Shafer [5]. For an “automatic user behavior-learning interface”, which is arguably an ideal naturalistic UI, an explicit service-appropriateness feedback (e.g. in the form of say a switch to inform the system about the

appropriateness of services) will eventually defeat the purpose of the self-learning interface itself. A feedback loop employing the implicit channel of communication, (such as, using affective computing technologies) to understand a user's natural responses to *likes* or *dislikes* for the service, may provide a very meaningful and valuable method of assessing user's (dis)approval in a naturalistic way. In the following section, we present a brief discussion on the relevance and importance of affective interfaces for CAIE.

### 2.1.2 Affective Human-Computer Interaction

Affective computing systems have been defined as a new class of intelligent systems that understand and influence human emotions. As Doctor et al. note in the introduction to a recent thematic issue in “Affect-Aware Ubiquitous Computing”, the projected number of Internet of Things (IoT) devices may cross 16 billion units by the year 2020 [33], which will open up new avenues for human-centric computing. The emergence of affective-computing paradigm has helped in developing a symbiotic interaction and transference between the human and digital environment where the overall quality of interaction is expected to benefit [33]. As already discussed in Section 1.1.2, incorporating affective computing (AC) technologies improves system usability, the quality of interaction from the design point of view as well as reduces the overall perception of stress during computer interaction [8]. In addition to improved interaction, AC technologies have implications for user's performance at the interface. As pointed out by Brave et al., interfaces which lack the capacity to understand and reciprocate emotions, can dramatically impede user's performance [49]. Such AC systems have been successfully implemented in diverse applications such as robotic personas [21], learning companions [22], affective tutors [23, 24, 25], affective games [26, 27, 28, 29], psychotherapy (such as in autism) using wearable devices [30, 31, 32, 50] to name a few. Not only that, as we will see later on in this section, it is also interesting to note that with

an increased coupling of computing infrastructure in human lives, the quality of interface design has immense effect on human performance on information manipulation and user's overall motivation to adopt these enabling technologies.

**How important is Affect for Human-Computer Interaction?** : Following initial lacklustre response for affective computing research [51], there is a renewed focus in affective interfaces following continued research findings in psychology, physiology and computational behavioral analysis which demonstrate the interdependence among these fields [52, 8]. Relationship between human affect and cognition is bidirectional. Riesberg et al., Ohman, Anderson and Vuilleumier have independently shown that human affect has proven impact on the speed of information processing [53], to the extent of determining whether the information will be attended to [54, 55], or will it be registered in memory [56, 8]. Partala et al. have shown that simple quirks of human-computer interaction such as mouse response delay has bearings on human performance due to perceptions of stress, and that affective interventions help regain performance [52]. As will be highlighted later in Section 2.2, human affect has a symbiotic relation with motivation, goal-setting and thereby personality of a human, which is also helping researchers arrive at computational models of affective intelligence and methods of incorporating them in computers.

In Section 1.1.2, we stated that even if a computer gives an impression of understanding human affect, the quality of interaction as well as the usability of the system improves. In a study, Klein et al. deliberately elicited frustration in a user by incorporating random delays or unresponsive behavior by a computer inhibiting user's progress towards a goal—a phenomenon christened *technostress* [35, 39]—while also provisioned for methods to vent frustration (that is, a form of affective support)[57]. Findings demonstrate that users receiving this affective support persisted with the frustrating task for a longer duration, compared to users who did not receive support [8]. Prendinger et al. showed similar trends by providing

direct affective support using empathic agents [20, 8].

Thompson et al. point out that “measuring the stress or difficulty caused by a system may also allow developers to pinpoint problems, or simply allow the system to be improved by being able to respond in a more natural and realistic way” [8]. Thus, understanding and reciprocating human affect is indeed necessary for an effective design of the human-computer interface. With the closer coupling of computing technologies in our daily lives in the form of intelligent environments, affective understanding and support will go a long way in realizing the vision of computing in the “*woodworks everywhere*”.

### 2.1.3 Affective Interfaces for Intelligent Environments

Earlier we have discussed the prolific use of affective computing technology in designing interfaces for traditional desktop computers used for applications such as learning companions, affective tutors, affective games, psychotherapy, affective prosthesis and robotic companions (as shown in Section 1.1.2). We classify these as *active-agent* approaches in HCI design, wherein the user can either directly interact with the affective agent or can see the changes caused by the agent on the computing screen. However, the evidence of affective interfaces as a *passive-agent*, where the affective component doesn’t directly influence the interaction but rather modulates and reconfigures the services in an intelligent environment scenario, is limited. We have seen in Section 1.1, such *passive-agent* approach has been widely studied and implemented in personal computing scenarios, where affective computing techniques have been used in information retrieval domain to create an implicit user feedback loop by developing a personalized model of user-satisfaction from their frustrated states [11, 12, 13, 14, 15, 16, 17]. However, there is a gap in literature for using such techniques in pervasive computing scenarios.



Affective interfaces for intelligent environments are being envisioned as an extension of their desktop counter-part. Sanchez et al. have used animated avatars on display screens for portraying prototypical affective states to users in an intelligent environment [58]. Carnegie Mellon University researchers designed ComSlipper project which augments traditional slippers with sensory actuators to communicate prototypical emotions via tactile information such as single or rhythmic taps to another connected slipper which responds by vibrations and producing warmth [59]. Although this work is novel, however there is no inherently natural mapping of the signalling rhythms and users need to learn it. Some groups of researchers use smartphone usage, game based therapy and questionnaires to investigate early signs of social loneliness and cognitive impairment for elderly [60, 61, 62]. These are interesting works which lean more towards behavior modeling and affective intervention rather than designing an affective interface for intelligent environments.

Doulamis proposed a facial emotion recognition based adaptive pervasive computing framework which updates its performance by adapting to individual users based on their current context including social context (surrounding people such as friends, family and acquaintance) [63]. The results are promising and visual recognition modality is interesting. However, it also has some technical and social drawbacks such as expression suppression, visual occlusion and privacy, which may defeat its long term adoption. Moncrieff et al. in their work, use audio input from a smart environment to detect contextually anxious situations to infer hazard for elderly people to provide active intervention [64]. The initial results are encouraging, however the audio interface also has some technical and social limitations quite similar to visual interface. In addition, if the person in the environment consciously suppresses his vocal reactions due to the social context, the system is practically blind (or *deaf*) to their affect [65].

The Affect and Belief Adaptive Interface System (ABAIS) was designed to compensate for

performance-bias caused by user's affective states and active beliefs [66]. ABAIS presented a graphical user interface (GUI) reconfiguration system in the context of Air-Force combat task, where the performance-bias prediction (threat perception in view of ambiguous radar data, severity of action under threat etc.) is based on empirical findings from affective research, and the anxiety assessment module is modeled as a knowledge based system. The ABAIS system identifies the potential impact of the inferred affective-states and active beliefs on user's performance, and finally selects a mitigation strategy (such as redirecting focus on salient cues by redirecting attention to important cues or presenting additional/reconfigured information to reduce ambiguity and aid decision making). This system is very interesting and quite relevant to our work. However, the problem with rule-based and knowledge-based systems like ABAIS is, they are quite inflexible and follow the programmed rules *as-is* [67].

Benta et al. recently presented a facial expression based affective-feedback in affect-aware smart homes, where the system takes *positive* expression as approval and *negative* expression as disapproval of an intelligent system's decision/service [67]. Broekens has used a similar approach of facial expression based affective feedback to *teach* desired behavior in human-robot interaction, employing reinforcement learning [68]. The robot *learns* desired behavior from positive reinforcement (reward) by positive expressions, and *rejects* undesirable ones by negative reinforcement (punishment). We see two principle drawbacks with using facial expression and audio based sensing mechanisms, namely (a) lack of a continuous and pervasive sensing interface in a CAIE, for capturing such expressions that are very sensitive to sensing angle and occlusion, and (b) both these modalities capture fully formed emotions, which are susceptible to conscious suppression of outward display of emotions due to social context [69, pp 155-194] [70].

Although our work is, in principle, similar to many of these prior works, it differs on the implementation methodology. We are employing the physiological bio-sensing paradigm

using wearable on-body sensors for measuring physiological variables related with human psychological states, such as electro-dermal activity, heart-beat, blood volume pulse and skin temperature. We will delve deeper into the neurophysiology and psychophysiology of human affective states in Section 2.3.1. Wearable on-body sensing alleviates a major issue unique to a CAIE, arising from the lack of a continuous and consistent sensing modality for tracking user's reactions in the wild, as we will see in Section 3.1. Research on emotion suppression has shown that although users may consciously suppress some expressions, but that doesn't alter their experience of the emotion [69], which suggests that physiological variables may not be affected by emotion suppression. This is a distinct advantage of physiological sensing of a user's affective states, that is, it arguably captures pre-conscious processes in the human body providing an unaltered window into nascent emotional states [71], and thus has the least chance of being suppressed due to social context or any other such considerations. In Section 2.4, we will delve deeper into this phenomenon of Autonomic Nervous System (ANS) activation of end-organs. Thus, both of the above mentioned arguments address the problems of affective state recognition using variations in facial expression or voice prosody, as used by prior works of Benta et al. [67], Doulamis [63], Broekens [68] and Moncrieff [64]. In the next section we will delve deeper into the definition of a user's context, encompassing affect, goals and intentions. Here we will also present a brief overview of models of affect in relation to designing computational systems.

## 2.2 Context Definition: Role of Affect and Intentions

In this section, we will discuss certain definitions and constituents of user context that are well suited for the computational intelligence community. Maat et al. have presented a set of six questions [72], answers to which will help in providing an outline of the model of user

context, needed for designing effective service-delivery in a CAIE. We argue through these set of questions that human affect is an important part of user-context which should not be ignored when designing a user interface for CAIE. Later on, we discuss the various underlying cognitive processes of appraisal and coping that precede affect generation in users. We argue that a comprehensive knowledge of these underlying processes is extremely necessary to make sense of a complete model of user context.

### 2.2.1 Relevance to our Research

Discussion presented in this section will show that appraisal and coping, two important cognitive subprocesses of affect generation, are important in meaningful interpretation of a person's actions in a given context. In Section 2.3.5 we will see that personality traits play immense role in determining if an event is appraised as a stressed state, what coping strategies a person chooses and whether such a strategy will prove to be effective. Applications developed for continuous, real-time ambulatory monitoring of stress in-the-wild using non-invasive techniques—such as ours—remains an open challenge to date [73]. In addition, it is well known that physiological indicators of stress, such as heart-rate variability, electrodermal activity, respiratory rate, thermal imaging, display individual person-specific differences [74, 75, 73, 28, 76, 77]. In view of the discussion presented in the previous sections, personality traits present a plausible framework to explain such differences. A recent work by Chittaranjan et al. used mobile phone usage statistics to gauge dominant personality traits in a human being, such as “Big-Five Traits” (term coined by Lew Goldberg [78]) [79, 80]. Such innovative methods of ascertaining individualized patterns may hold key to designing personalized affective interfaces for context-aware intelligent environment. Thus, it is imperative to incorporate all these factors in our study.

## 2.2.2 Definition and Components of Context

Schilit et al. in their 1994 paper [81] introduced the term Context-Aware Computing in literature, wherein the authors present the idea of computing systems that determine and in-effect react to a user's ever-changing context in order to help the user accomplish their intended tasks. Schilit et al. define user-context as the immediate proximate environment of the user—people and things that the user is interacting with, in a particular place at a time, and the spatio-temporal history of their changes. Environmental variables such as ambient noise, temperature, available communication bandwidth etc. are also considered to constitute user context. Various other definitions of context have emerged ever since, as noted in a widely cited survey by Baldauf [82]. However, as mentioned in Section 1.1.1, the widely accepted definition of context, as given by Dey et al. is — “any information that can be used to characterize the situation of entities (i.e. whether a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves” [83, 18].

An interesting aspect about Dey's definition of context in [83] is the inclusion of user's emotional state and focus of attention as its constituents. Prekop et al. classify context along two dimensions--*external* and *internal* [84] while Hofer classifies context along *physical* and *logical* dimensions [85, 82]. Although Hofer's definition of context is computer-application centric as opposed to Dey's and Prekop's definition being user centric, but his idea of logical dimensions being used to accentuate the essence of physical dimensions of context, resonates with the widely accepted view of internal and external user context. The distinction between these dimensions is that the external or physical dimension refers to the context data collected from environmental sensors like ambient light, sound, temperature, movements, location, time; whereas the internal or logical dimension captures the monitoring of user interactions,

user goals, tasks and emotional states.

Maat et al. in [72] presented an affective multi-modal human-computer interaction (AMM-HCI) system where they have modeled user's context by answering six questions related to the user's activities in the computer interface, namely:

- *Who* – User's own identity and identity of other humans in user's surroundings
- *Where* – Environmental parameters and Location of the user
- *What* – Current task or activity of the user
- *When* – Time log of the user's activities and affect
- *How* – How is the user enacting interactive cues i.e. user's affect.
- *Why* – Why is the user performing the observed activities and portraying certain affective states

Answers to these questions will give us an outline of the model of user context that a context-aware computational system must adhere to. To distinguish these questions among the dimensions of context, *Who*, *Where*, *When* and *What* constitute the *external* context, whereas the questions of *How* and *Why* form the internal context of the user. As rightly pointed out by Maat et al., the most difficult question among these context questions is answering the *why* context question.

### 2.2.3 User Intentions in the Context Model

We need to use models for inferring User Intentions from their goals and actions to answer this *why* context question. Burghardt et al. have modeled user intention with tuples of users Goal, Strategy and Action [86]. Sadri et al. define *intention* detection (i.e. inferring

user goals), as the task of recognizing the intentions of an agent by analyzing some or all of their actions and/or analyzing the changes in the environmental state resulting from their actions [87]. Sadri also defines *plan* or *strategy* recognition as the task of recognizing the sequence of actions (including future actions) that the observed agent is following in order to achieve his intentions. Burghardt argues that the user has a certain amount of commitment towards his goals, they perform certain actions to fulfill their goals as per their chosen plan/strategy and in effect leave a trail of changes on the environment which are picked up by the external context (ambient) sensor infrastructure. In addition to an explicit change in users goals, changes in the environmental context of the user may also prompt the user to change their goals and strategies and thus their actions. Thus, user intention must be an integral part of users context. Logic based approaches such as abductive reasoning, case-based reasoning, task networks, causal theories, and probabilistic models have been used to model user intention [86, 87].

An intention recognition system (IRS) must have the following components – 1. Set of intentions to choose from 2. Some knowledge about how plans achieve goals 3. Sensed sequence of actions by the agent 4. Prior belief about intention. There is an assumption that the agent is a rational agent and their actions are geared towards achieving a goal. Knowledge about relations of plans and goals is modeled using techniques listed in the last paragraph. In an instrumented intelligent environment, component 3 listed above is fulfilled using activity recognition techniques. The most popular model for computational implementation of IRS is “Belief-Desire-Intention” (BDI) Model as presented in [88].

## 2.2.4 Interdependence of Affect and Intentions

Environmental and social context plays a very critical role to disambiguate various cues which are input to data analysis in emotion recognition systems. A commonly used example goes like this - without the social context an isolated flush may signal either pleasure or anger. Hammal points out that without social context, even a human may misunderstand interaction cues [43]. Vinciarelli says that the contextual information cannot be overlooked in automatic emotion recognition, also adding that incorporating social and environmental cues may provide a different flavor to each social interaction [89]. Cowie et al. point out that complex mental states should be inferred from heterogeneous information sources that provide enough information to detect the explicit modes as well as help to infer the implicit cues of human interactions [9]. The *how* context question raised in Maat's work [72] addresses the affect recognition problem in human-computer interaction. Various computational models have been proposed such as the Ortony, Clore and Collins (OCC) model [90, 91] and Emotion and Adaptation (EMA) model [92, 93] to name a few.

Ramos et al. highlight that it is beneficial to incorporate affect, mood and personality models of the user at each level of ambient intelligent systems—sensing, reasoning and action [78]. Historically, there were fundamental differences among groups of researchers with characterizations of affect and human emotions, such as there was little agreement on the set of basic emotions, there is considerable difference in terms of defining emotions to be selection driven as opposed to culture driven. Cowie et al. in a systematic review [9], trace back these differences: (a) Rene Descartes introduced the idea of “*human emotional space constituting of a few basic emotions*” (b) Charles Darwin thought “*emotion as a biological phenomenon of selective behavior*” However, Ramos et al. point out in their survey [78], that recent years have seen a positive attitude of researchers cutting across disciplinary boundaries



such as psychology, neuroscience and philosophy towards the influence of affect in human decision making processes. The relationship between affect and social context runs both ways--the evaluation of task completion viz-a-viz goals elicit emotions using the process of “*cognitive appraisal*” [88]; these emotions, in effect color our views towards restructuring the decisions and strategies viz-a-viz our goals using the process of “*coping*” [94, 95]. We will see in next Section 2.3 that physiological pre-disposition, personality and social context play equal roles in these two reverse and complementary processes.

#### 2.2.4.1 Process of Cognitive Appraisal

Everly et al. define *cognitive appraisal* as “the process of cognitive interpretation, that is, the meanings that we assign to the world as it unfolds before us, and *affective integration* as “the blending and coloring of felt emotion into the cognitive interpretation” [96]. From a psychological perspective (mainly Lazarus’s theory of Appraisal [97], as also discussed briefly in Section 2.3), the process of cognitive appraisal comprises of the integration of sensory inputs and basic subjective evaluation of the present situation or sequence of events to first form emotions, and then to ascertain a plan of action required to react(if at all) to the event [97, 96]. This process is critical in our perception of stressors, i.e. in the determination of whether the psychosocial *stimuli* is a psychosocial *stressor* or not. Please refer to Figure 2.1, for a clearer picture of the primary and secondary appraisal processes. This is a widely accepted theory, influencing various computational theories of cognitive agents. Ortony, Clore and Collins defined a computational model of affective states—popularly known as the OCC Model—which defines human affect as “valenced reactions to three different kinds of stimulus—*objects*, *consequence of events* and *action of agents*”. It is the most popular model for implementing environments with intelligent agents [78]. The OCC Model proposes that affective states are attained as a result of three types of subjective *cognitive appraisal*:

1. Appraisal of pleasantness of events viz-a-viz agent’s goals 2. Appraisal of approval of agent’s behavior with respect to accepted standards of behavior 3. Appraisal of liking of objects with respect to attitudes of agents. An updated and simplified OCC model considers only two affect categories of valence—positive and negative [91], as shown in Table 2.1.

	Positive Reactions	Negative Reactions
Undifferentiated	... because something good happens ( <b>joy</b> )	... because something bad happens ( <b>distress</b> )
Goal-based	... about the possibility of something good happening ( <b>hope</b> )	... about the possibility of something good happening ( <b>fear</b> )
	... because a feared bad thing didn’t happen ( <b>relief</b> )	... because a hoped-for good thing didn’t happen ( <b>disappointment</b> )
Standard-based	... about a self-initiated praiseworthy act ( <b>pride</b> )	... about a self-initiated blameworthy act ( <b>remorse</b> )
	... about an other-initiated praiseworthy act ( <b>gratitude</b> )	... about an other-initiated blameworthy act ( <b>anger</b> )
Taste-based	... because one finds someone/something attractive ( <b>like</b> )	... because one finds someone/something unattractive ( <b>dislike</b> )

Table 2.1: Generalized *positive* and *negative* emotions, per OCC Model. Table from [91]

#### 2.2.4.2 Process of Coping

From a neurobiological perspective, coping is the process of reattaining the bodily state of *homeostasis* after the initial reaction from cognitively appraising an event [98, 36, 96]. From a psychological perspective, coping is the process of dealing with emotions, either externally (*problem-focused-coping*)—by forming intentions to act upon; or internally (*emotion-focused coping*)—by acting on self [97, 99]. As briefly described in Section 2.3.1, the activation of SAM axis is interpreted as *active-coping* and activation of HPAC axis as *passive-coping*. Problem focused coping refers to the efforts by the agent to improve troubled person-environment relation by acting on the exact reasons of the trouble, whereas emotion-focused coping is the process of altering beliefs, realigning goals and restructuring strategies to achieve new goals

in presence of strong emotions (which are generated due to appraisal of favorable/unfavorable actions viz-a-viz old goals) [94, 99].

These two processes of appraisal and coping interact with each other in a cyclic process of *appraisal*→*coping*→*reappraisal* and give temporal characteristics to human affect and intentions. This cycle forms the basis of emotional dynamics which says that an agent's appraisal of events create affective states, which color its further evaluations, ultimately changing its own emotional state at a future time. Two very important factors influence this emotional dynamic process—social context and personality traits of the subject [95]. Delongis et al. used Daily Process Method to investigate the effect of personality traits, nature of stressor and social contexts on the selection of coping strategies by several individuals in varied environments [95]. Key aspects of social context under study were personality of agents in the support system of the subject and individual's satisfaction with the support system. Delongis et al. found that choice of coping strategy and support system form a vicious cycle (that is, poor results from one feeds into poor performance of the other). This work also presents an extensive study of the Big-Five Traits viz-a-viz stressor types and found that personality interacts with stressor types to elicit coping strategies.

Thus, having described affective computing technologies and their role in defining a user's context, we are in a position to point out their symbiotic relation. From the previous two sections, we see that the concept of an inference based affective-feedback to determine the appropriateness of a intelligent service is a seamless fit into the scheme of a context-aware intelligent naturalistic environment. A user's context is fundamentally an inferred entity. Thus, the lack of an implicit feedback loop for a CAIE, prompting the user to explicitly undo an inappropriate service, defeats the idea of an automated service from the CAIE. Although, a few recent works demonstrate the use of implicit feedback in interface design in intelligent environments using the voice and visual modalities, we believe (a) the visual modality used

in these works has a basic drawback of the possibility of occlusion in a naturalistic intelligent environment, which is not the case with a wearable on-body sensing system. (b) both the visual and auditory modalities express fully formed emotions, thus, have the possibility of affect-suppression [71]. Towards that end, we propose to use physiological signals captured using wearable devices, to infer user’s pre-conscious affective state of *technostress* and use it as a surrogate for user’s (dis)approval of (in)appropriate services. We posit that a wearable device is well-suited for a CAIE, which has a fundamentally different interaction paradigm spanning multitude of devices and lacking a single focal point of interaction (see Section 1.1.1 for detailed discussion). In the next section, we will discuss the phenomenon called *technostress*, what are the physiological signatures of this state and how we plan to use it in this work.

## 2.3 Technostress in Computer Interactions

The interface between technology and human plays a critical role in human performance. More so, when intelligent agents and environments are being adopted at a never-seen-before pace, thus getting more entwined in our daily lives. Stress perception in human resulting from their interactions with technology is real [35]. This phenomenon has been formally defined as *technostress* by psychologist Craig Brod as “a modern disease of adaptation to cope with new computer technologies in a healthy manner” [34, pp 16]. Psychological stress, on the other hand, has been defined by Lazarus and Folkman as “particular relationship between the person and the environment that is appraised by the person as taxing or exceeding his or her resources and endangering his or her well being” [97].

In Lazarus’s theory of stress, people are continuously appraising potential stressors in the environment. The cognitive appraisal process comprises of *primary* and *secondary* phases.

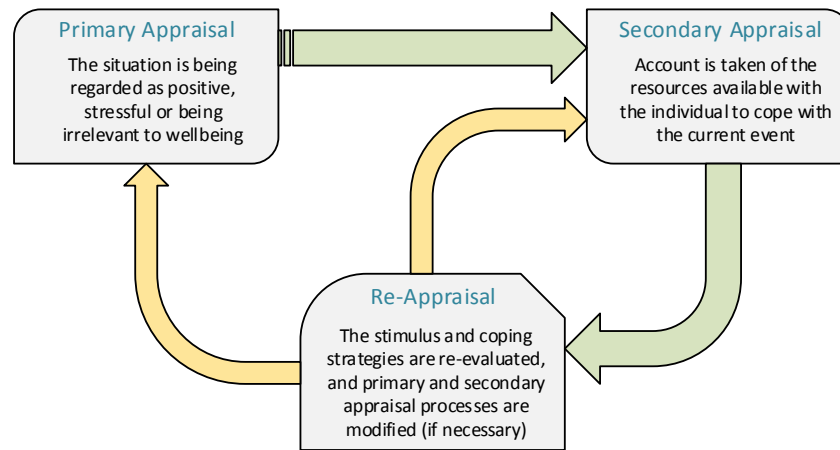


Figure 2.1: Primary and Secondary Appraisal Processes. Picture adapted from [98, pp 83]

The primary appraisal process determines the threatening agents posing challenges to the entity in the environment, whereas, the secondary process determines the ability of the entity to cope with the posed challenges. It is the secondary appraisal process which is critical in selecting  *coping*  strategies as well as affective physiological responses. Figure 2.1 shows the interplay between primary and secondary appraisal processes through a re-appraisal of the events at a subsequent time. A knowledge of the temporal dynamics of cognitive processes underlying stress appraisal is necessary as they are correlated with various physiological processes, as will be seen in details in the next section as well as Section 2.2.

Hudiberg studied and presented a comprehensive list of events that are known to cause *tech-nostress* [100], such as computer slowdown or crash, screen-freeze, unexpected and continued error messages, program response delays, input-output device malfunction, incomprehensible instructions, typing errors, poorly designed user-interface and poorly written system manuals etcetera. This list is available as *Computer Technology Hassle Scale* designed by Hudiberg [100, 101]. Various prior studies have reported evidence of stress during such episodes of hassles with computer interface [52, 102, 103, 104, 105].

It is worth noting here that many of the responses to technostress, have been shown to be physiologically similar to that during psychological stress responses [39]. In addition, in order to design a real-time affective computing system, it is also necessary to have a detailed picture of the temporal dynamics of the various biological subsystems involved in producing the human stress reaction. This chronological order of activation of biological subsystems as a part of the bodily stress-reaction, is also important in order to ascertain the physiological signal-streams necessary and sufficient for real-time on-body sensing of stress. In the next section, we produce a discussion on the neurophysiology of stress-reaction in the human body. The details presented in the following section of neurobiology and the chronological order of stress responses in humans, is especially relevant to developing the systems model of interaction for to address RQ 1 (shown in Chapter 1), as well as forms the basis of computational analysis of psychophysiological variables (shown in Section 2.4).

### 2.3.1 Neurobiology of Stress

An event, real or imagined, in the vicinity of a person has to be sensed and cognitively and affective appraised to be a “threat” to trigger the physiological stress reactions. In [96], Everly et al. presented a meta-analysis of state-of-the-art research on physiological reactions to stress which shows that physiological reaction to psychosocial stress can be temporally sequenced into activation of three nervous systemic axes: (i) Most immediate term response through Neural Axis activation, (ii) Intermediate term response through Neuro-Endocrine Axis (also termed as Sympathetic-Adrenal-Medullary (SAM) axis) Activation and (iii) Endocrine Axes (most prominent among these is Hypothalamus-Pituitary-Adreno-Cortical (HPAC) axis) Activation [96] [98, pp 87-89]. Each individual axes is activated sequentially, upon continual re-appraisal of a persistent stressor.

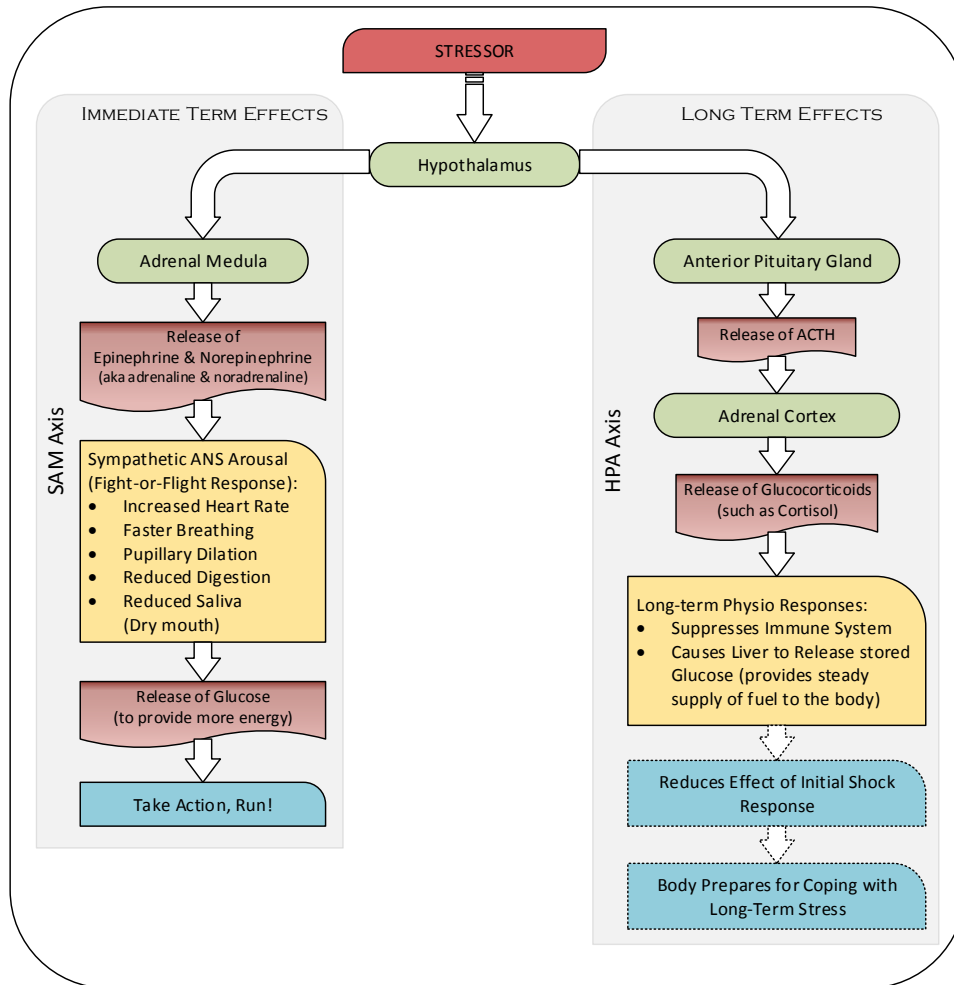


Figure 2.2: Sympatho-Adreno-Medullary Axis and Hypothalamus-Pituitary-Adrenal Axis Activation in Response to Stressors. Picture adapted from [98, pp 89]

**Neural Axis Activation :** The most immediate response to a stressful stimulus occurs via the direct neural innervations of end organs, via the activation of both the divisions of ANS i.e. Sympathetic Nervous System (SNS) and Para-Sympathetic Nervous System (PSNS). Since these are the end organs that are directly innervated by the ANS, measuring the responses from these end-organs gives the earliest indications of stress appraisal [96]. The neural impulses from the limbic system, which take part in appraisal of an event, trigger

posterior and anterior hypothalamus to activate SNS and PSNS, respectively.

The effect of SNS activation is generalized arousal of various body systems such as heart, sweat-glands, lungs etc, resulting in increased heart-rate, increased muscles stimulation, increased breathing to name a few. The effect of PSNS activation are inhibition and restorative functions of end organs. Although most common neural activation in humans is SNS activation, but PSNS activation has also been observed [96, pp 32]. This simultaneous activation of SNS and PSNS is counter-intuitive, although litera-

Table 2.2: Effects of Neural-axis Activation during Stress Appraisal [96]

Effect	Process
Secretes	Sweat Glands
Dilate	Eye Pupil
Decrease	Salivary Secretion
Increase	Heart Rate
Decrease	Bloodflow to Skin

ture provides enough evidence. We must note here, not all end organs are equally innervated by SNS and PSNS nerves, which is why there are specific systems which should be used for noninvasive detection SNS activity (See [96, Table 2.3] for list of organ innervation). Neural axis activation is the quickest response (and also the weakest, due to limited capability of SNS to continue secreting neurotransmitters [96, pp 32]), and is the main focus of wearable sensing in majority of prior research. The effects of this activation usually lasts till say 3 - 5 seconds [106, Ch 2], which explains the usual EDA processing window of 5 seconds following an event (as shown in Section 2.4).

**SAM Axis Activation :** To continue bodily responses in moderate to chronic stress conditions, the immediate neural axis activation is followed by the activation of adrenal-medulla gland (the neuroendocrinal axis), to trigger what is popularly known as “fight-or-flight” response. The neural impulses start at dorsomedial-amygdalar complex, travel through the spinal cord to adrenal-medulla situated on top of kidneys, which on activation secretes adrenal-medullary catecholamines (epinephrine—almost 80% and norepinephrine—the rest).



The effects of these catecholamines on end-organs are functionally identical to that of direct SNS activation, only these are an order of magnitude stronger (producing more pronounced responses) and require a delay of atleast 20 seconds to produce effect [96, pp 34] (see [96, Table 2.4] for organ effects). It is worth noting here that electrodermal activity (EDA) and bronchiole effects are not affected by these catecholamine release [96, pp 34]. Due to this similarity, the neural axes activation is often merged with SAM responses, as done in [98, pp 89] (See Figure 2.2). Researchers have called this activation as “Sympatho-Adreno-Medullar” (SAM) axis, and is generally regarded as an *active-coping* mechanism (to be discussed in Section 2.2.4) to bring the body back to *homeostasis*, after the initial shock response.

**HPA Axis Activation :** If the stressor (real or imagined) persists till a chronic stage, the endocrine axis is activated where the most prominent response is seen in “Hypothalamus-Pituitary-Adreno-Cortical” (HPAC) axis. The activation is initiated in septal-hippocampal complex, the neural impulses reach median eminence of hypothalamus and in turn secrete *corticotrophin release factor* (CRF) which on reaching anterior pituitary glands secrete *adrenocorticotrophic hormone* (ACTH) into the systemic circulation. ACTH reaches the adrenal-cortex glands situated on top of adrenal-medulla and releases glucocorticoids (e.g. cortisol—the “stress

Table 2.3: Effects of SAM-axis Activation during Stress Appraisal[96]

Effect	Process
Increase	Arterial Blood Pressure
Increase	Heart Rate
Increase	Cardiac Output
Increase	Muscle Stimulation
Decrease	Bloodflow to Kidneys
Decrease	Bloodflow to Skin

Table 2.4: Effects of HPAC-axis Activation during Stress Appraisal [96]

Effect	Process
Increase	Glucose Production
Increase	Gastric Irritation
Increase	Urea Production
Decrease	Apetite

hormone”) and mineralocorticoids into the systemic circulation. Effects of cortisol secretion are suppression of digestion, delay onset of sleep [36], increase in glucose production thereby heightened blood sugar, and suppression of immune system (see [96, Table 2.5] for complete list).

In addition, cortisol also acts as a negative feedback to catecholamine release hormones, thereby suppressing the effects of the initial “fight-or-flight” stress-responses of the SAM axes [36]. Activation of HPAC axis is the slowest occurring but the most chronic stress symptom, often associated with helplessness syndrome and is called *passive-coping* system (discussed in Section 2.2.4). Evidence of activation of the neural axis and the neuroendocrine (SAM) axis during technostress has been extensively reported in prior research. A recent study, however, also reports increase in the stress-hormone Cortisol due to computer breakdown [36], indicating long-term stress-response (see Figure 2.2 for detailed stress response timeline).

The chronological progression of stress-reactions in bodily subsystems can be summarized pictorially as shown in Figure 2.3.

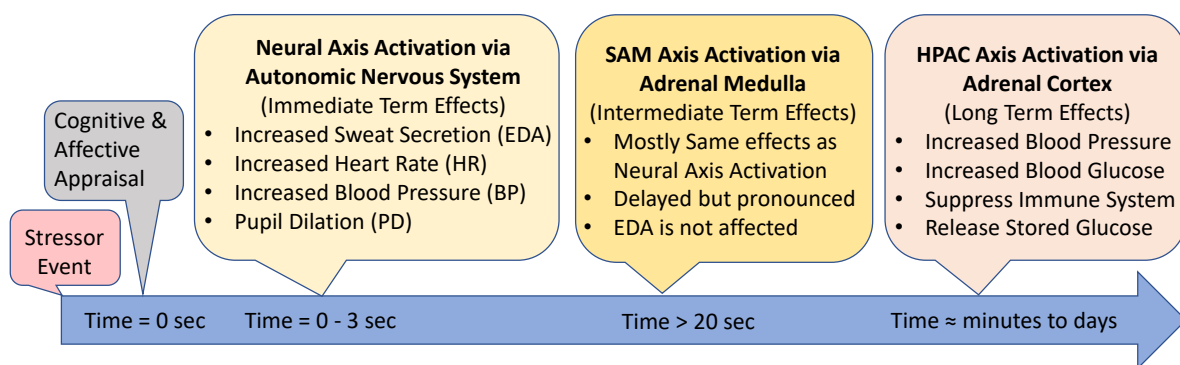


Figure 2.3: Chronology of stress-reaction showing activation of bodily subsystems on continued appraisal of a stressor.

### 2.3.2 Psychophysiological Indicators of Technostress

Technostress, as described in Section 2.3, is a biological and psychological phenomenon which is triggered due to a perception of hassle, such as variable response delay or system breakdown, during a user's interaction with computers. This perception of hassle may trigger activation of biological stress mechanisms (as described in Section 2.3.1) that encompass many physiological subsystems such as autonomic nervous systems, central nervous system and endocrine system [39]. As evidenced in the previous section, the neurobiological anatomy of the stress reaction finally ends in the activation of end-organs. The catecholamines secreted by adrenal medulla and glucocorticoid secreted by adrenal cortex glands, when pumped in the systemic circulation have various arousal effects on respiratory organs, skin, skeletal muscles and heart. Literature survey such as Reidl [39], show very similar physiological effects resulting from technostress.

Prior research shows physiological effects of technostress include heightened levels of adrenaline and noradrenaline (referred to as epinephrine and norepinephrine in U.S.A. [98, pp 88]) [102, 107], increased mean spontaneous skin-conductance responses (sympathetic nervous system activation) , heart rate and blood pressure (increased catecholamine secretion) [108, 109, 107], increased jaw muscle electromyograph (indicating clenching of teeth) [103, 110, 107]. A recent survey also reports the following physiological responses to technostress: (a) neuronal effects such as decreased P300 amplitude indicating fatigue (b) elevated stress hormones (adrenaline, cortisol) and other precursor stress enzymes (alpha-amylase, adrenocorticotrophic hormone) (c) elevated responses of sympathetic activation such as increased blood-pressure, skin conductance, heart-rate, muscle tension [39]. A detailed discussion on methods of computational analysis of these psychophysiological variables has been presented in Section 2.4.

### 2.3.3 Technostress as Negative Feedback Signal

We have seen in the previous section, technostress produces distinct physiological signatures. Prior research has primarily focused on identifying antecedents and consequences [36] of technostress as well as detecting technostressed-states [105, 104, 111] in order to develop ways to mitigate it by formulating and experimentally validating design guidelines for interface design. However, in our approach, the successful detection of technostressed states will be used as an implicit feedback from the user representing disapproval for a service, and to trigger a modification in the service in order to provide desirable output to the user.

In a way, instead of researching on methods to mitigate technostress, we conceptualize an adaptive interface that infers technostress during interaction as a *negative feedback* from the user about the service's desirability. The system then iteratively modifies the services, so as to minimize this technostressed state. We argue that this approach of responding to the user's innate feedback, conveyed via the *implicit* channel, will result in a more human-centric user interface as discussed in Section 1.1.4. We have already seen in our previous discussion in Section 2.1.2, that even if a computer gives an impression of understanding human affect, the quality of interaction as well as usability of the system improves, as also evinced by the experiment conducted by Klein et al. [57]. To the best of our knowledge, as also noted in our short survey on the current state of the art in affective interfaces design for intelligent environment (presented in Section 2.1.3), this scheme of using technostress in a CAIE to detect service relevance has never been tried before.

It is well known that physiological indicators of stress, such as heart-rate variability, electrodermal activity, respiratory rate and thermal imaging display individual person-specific differences [74, 75, 73, 28, 76, 77]. In the next section, we present a brief discussion on some of the factors that have considerable influence on the degree and pattern of end-organ

activation from stress-responses. We will highlight in more details, some of these individual differences related to each signal stream in Section 2.4. The information presented here is relevant in identifying various confounding factors for physiological sensing of technostress, and forms the basis of our investigations into designing a personalized physiological profile learning component in our framework.

### 2.3.4 Individuality of Stress Reaction

We have seen briefly in the previous section that the *cognitive appraisal* process is critical in the perception of stress and identification of a stimuli as a stressor (please refer Section 2.2.4 for a detailed discussion). Everly et al. have noted in [96], that such a perceptual process is uniquely individualized and dependent on biological predispositions such as personality patterns and available resources of coping. The physiological response to stress is dependent on response mechanism stereotypy (preferential pattern of organ system activation) and target-organ specificity (predisposition of the target organ to experience arousal), both of which are heavily determined by factors such as genetics [96].

Such individual differences are measurable from the human physiological reactions to stress, as evident from multiple studies relating electrodermal activity (EDA) lability to personality traits and predispositions [112] as well as tasks demanding vigilance and personal ability to allocate information processing capability to various stimuli [113] (see Section 2.4.1). For instance, Dawson present a survey of various studies suggesting the differences in EDA responses could reflect individual differences in higher central processes involved in information processing, which was consistent with some experimental findings showing “EDA labile children outperformed EDA stabiles on various tasks requiring perceptual speed and vigilance” [113]. This is very relevant to our study involving “Paced Stroop Test” (see Section 3.2.1)

wherein vigilance and information processing speed are critical. In addition, we hypothesize that these personal predispositions are also broadly related to the perception of technostressors and portraying typical reactions thereafter.

### 2.3.5 Moderators of Technostress Reactions in Humans

There is consensus in literature about the multidimensional character of technostress, having various antecedents and consequences. Moreover, there are multiple possible moderators that may affect the relationship between technostress and its antecedents and consequences. One such counter-intuitive moderator—cultural orientation—was reported in studies conducted by Wang and Tu on Chinese population. They report that unlike North America, technostress has no significant effect on employee productivity [36]. A discussion on a few major moderators is presented below.

#### 2.3.5.1 Gender and Stress Reactions

People use personal computers to accomplish specific tasks in various settings and contexts such as office, education or home. With the projected development in IoT, and further deployment of intelligent environments, people are expected to interact with these technologies to accomplish everyday tasks. The perception of hassle in such environments due to system delay or misinterpretation of user's commands will threaten the accomplishment of goals for the users. The expectation from the use of information and communication technologies (ICT) is to accelerate goal fulfillment, which only aggravates the negative impact of system malfunction because the user is expected to deliver better results faster when using ICT in all contexts [35]. Such complex human-computer interactions tasks that are also time-critical are said to impart *achievement stress* [35].

Prior research has shown the evidence of significant gender differences in physiological reactions to different types of psychological stress perceptions. Stroud et al. demonstrated gender sensitivity to different types of psychological stress by showing difference in cortisol secretion levels [114]. In the study, male and female participants were exposed to *achievement* stress (complex tasks) and *social rejection* stress (excluded from a conversation). The results showed that only male participants show significant increase in cortisol level while responding to *achievement* stress, whereas female participants show increased cortisol level on exposure to *social rejection* stress [114]. Taylor et al. show similar gender differences in behavioral responses to stress perception, arguing that behaviorally females show *tend-and-befriend* as opposed to males, who show *fight-or-flight* responses [115]. There are even more evidences which show performance failures in accomplishing time-pressured tasks (akin to ICT tasks) cause increased perception of stress in males [116], whereas, situations of interpersonal conflict (not similar to ICT tasks) cause increased perception of stress in females [117]. A plausible explanation for this heightened male reaction to *achievement* stress can be drawn from the field of evolutionary psychology (for details, see [115]), which argues that historically males acted as “hunters”, while females primarily acted as “gatherers” [35] and thus responded to stress with a *tend-and-befriend* behavior [115].

Riedl et al. hypothesize, citing this body of work, that there must be significant gender differences in technostress perceived in a typical time-pressured human-computer interface task [35]. They note that technostress research has continued to be gender-neutral, and list various reasons why it should not be so. Gender has been identified as an important factor to explain variances in activation of various biological subsystems in response to stress factors [118, 119]. A few studies also report gender based attitude differences towards technology acceptance (see [35] for details), whereas other studies have explicitly called for difference in interface design to account for the gender differences in perception of online

trust, color or preference to information processing [120, 121]. However, a few limited studies that strive to uncover these differences have reported conflicting results. Although two recent articles by Tarafdar et al. and Riedl et al. found men experience significantly more technostress compared to women in time-pressured ICT tasks [122, 35], a relatively old study by Elder et al. found exactly the opposite—that women experience more technostress [123]. This provides ample evidence which suggest that gender differences are critical factors in determining technostress responses. To the best of our knowledge, nobody has studied the detection of technostress in an intelligent environment, let alone studying gender differences.

### **2.3.5.2 Role of Personality in Technostress**

Personality traits are known to play significant roles in the appraisal and coping processes [97, 99, 95], both of which are integral processes in the trigger and mitigation of stress responses. Technostress may occur in organizational as well as personal context. Multiple studies examining the effects of personality on technostress focused on the general organizational context [124, 125], including specific cases such as banking [126] and library personnel [127]. These are largely focused on discovering correlation between personality traits with the choice of coping strategies during technostress. However, studies relating personality traits to differentiated perception of technostress could not be found.

Personality traits are known to have immense bearing on how a person chooses his goals, how he appraises a particular action viz-a-viz his goals that imparts certain emotions in them and what coping strategies they choose in a given situation [78]. Personality traits such as neuroticism are known to be strong predictors of perception/appraisal of events as psychosocial stressors [128, 39], and neurobiological and physiological reactions thereafter. Similarly, personality types have also been reported in research studies to significantly influence appraisal of events as stressful [98, Page 92]. Personality traits have also been indicative



of effectiveness of certain coping strategies in stressful situations [95]. Thus, we see ample evidence of personality traits and types influencing the psychosocial stress related appraisal process. Everly also notes that cognitive appraisal is an individualized process, vulnerable to personality patterns, learning history and available resources to cope [96]. We also see a research gap in finding evidence of reproducing similar effect for technostress in ICT settings. Our experimental setup of technostressors in an intelligent environment (to be discussed in details in later Chapter 4) may provide a unique opportunity to answer this question.

In the next section, we will present a survey of computational approaches of extracting statistical features of physiological signals, and methods for recognition of affective states such as technostress. As pointed out earlier, in this dissertation, our target for recognition of technostressed states is to identify heightened sympathetic activation from physiological signals. Computational recognition of these states can be used as a service-appropriateness feedback for our interaction framework as described in Chapter 3. These techniques are especially relevant to answering RQ 3 and RQ 4, as described in Chapter 1.

## 2.4 Computational Psychophysiology

Emotion studies were largely divorced from the brain-related studies, until recently, when the neuroscience community proposed new methods to understand neural correlates of emotional states. The emergence of the field of *Affective Neuroscience* has helped standardize the underlying neural circuitry for emotional experiences [129]. Psychological stress is one such affective state, which draws immense attention from various fields of computer sciences, engineering, psychology and their related branches. Everly defines human stress response as “a physiological response that serves as a mechanism of mediation linking any given stressor to its target-organ effect” [96].

Affective studies have paid considerable attention towards ANS activity, in part due to the ease of measuring its constituents non-invasively using wearable on-body sensors. For example, physiological correlates of ANS activity can be directly measured by observing changes in heart rate, blood pressure, blood pulse volume, respiration rate, pupil dilation, electro dermal activity, skin temperature and skeletal muscle tension [49]. Physiological sensing provides a reliable modality of non-invasively capturing responses directly from ANS thereby opening a window into signals that may reflect various processes that are beyond cognitive intent [71, 130]. ANS activity is known to be more pronounced during negative emotional states [131] (such as technostress). We might recall from Section Section 2.3.1 where we discussed the neurobiology of human stress response and temporal chronology of biological processes that *Neural axis* is activated immediately following the appraisal of a stressor. For detecting *technostress* from events in a CAIE using physiological signals, temporally the most ideal indicators of stress are those effected by the activation of *Neural axis* as noted in Table 2.2. Colomer et al. [132] in a comprehensive analysis of various indicators of ANS arousal show that features derived from electrodermal activity (EDA) and heart-rate-variability (HRV) are the most significant attributes, a result corroborated by other researchers such as Yoo et al [133] as well. Following such comprehensive analysis of affective computing literature as well as our own survey into neurobiology of stress, we have decided that for our research, we will be using the features derived from Electro-Dermal Activity (EDA) and Heart Rate (HR) signal streams. In the subsequent sections, we will describe their salient features relevant to our research.

### 2.4.1 Electro-Dermal Activity (EDA)

Electrodermal activity has often been described as “perhaps the most widely used index of activation” in the field of psychophysiology, being under active research for over 100 years.

One key reason for the popularity of this datastream among psychophysiology researchers is the direct and undiluted representation of sympathetic neural activity [113, 130]. Eccrine sweat glands are primarily responsible for thermoregulation in the human body. In a comprehensive work on EDA by Dawson et al. [113], the authors report earlier findings by Darrow et al. that “the function of secretory activity in the palms is primarily to provide pliable adhesive surfaces facilitating tactual activity and grip on object”. Thus, eccrine glands present in large concentration on the inner side of the palm and feet are thought to be aiding gripping and less responsible for thermal cooling, and hence more responsive to psychological stimuli. EDA is an umbrella term which defines the change in electrical properties of skin measured across specific active sites. It occurs due to sweat secretion from eccrine glands during aroused SNS activity [130].

EDA signal arguably is a unique channel which is innervated only by the SNS division, thus, making it a reliable marker of sympathetic activation [130]. Skin conductance (SC) is the most widely used measure to quantify EDA, which is composed of a slow changing background component called skin conductance level (SCL) and a rapidly changing component called skin conductance response (SCR). In [134], Darrow provided empirical proof for using change in skin conductance as a reliable measure of electrodermal activity. Per Dawson et al. [113], tonic SCL has been widely reported to be low during sleep and high during activated states such as anger or activity, whereas phasic SCR has been related to attention and noted that this response is sensitive to stimulus novelty, intensity and significance. EDA is an established measure of SNS arousal as it is arguably the only physiological variable that reflects the SNS activity uncontaminated by parasympathetic nervous system (PNS) activity [130]. Event related phasic SCR (ER.SCR) are quite informative and have shown wide variation in rise-time, decay-time, amplitude and latency based on the nature of stimulus applied. Electrical conductivity recordings are arguably very reliable indicators of SNS

activity [35, 131, 75, 77] which shows heightened activity during the perception of stress.

#### 2.4.1.1 Physiological Basis of EDA

The eccrine glands consists of a secretory portion which is in the form of a coiled duct, and an excretory portion which is a long duct that opens up on the skin surface as a small pore. To understand the physiological basis of EDA, it is convenient to model these long sweat ducts as sets of resistors connected in parallel [113]. Columns of sweat will rise in these ducts in varying amounts corresponding to the degree of activation of the SNS. As sweat fills the ducts, conductive paths are formed through the otherwise relatively insulating skin, thereby reducing the value of these parallel resistors and in turn resulting in observable change in EDA. Convincing experimental evidences have conclusively proved that eccrine glands have predominantly sympathetic innervations and there is a high degree of correlation between bursts of SNS activity and SCRs. For detailed description of the multiple complex neural pathways considered responsible for EDA generation and modulation, please see the seminal works by Dawson et al. [113] and Boucsein [135].

Postganglionic sudomotor fibers, directly connected with the eccrine sweat glands, are responsible for transmitting the nerve firing signal to eccrine sweat glands to start sweat discharge. Postganglionic sudomotor fibers which are slow fibers which have a conduction velocity of roughly 0.5 to 2m/s. Conduction time from central activation to the sweat glands of the fingertips (with a mean distance of 1.1 m) was estimated at 1.1s [136].

**Electrical Recording of EDA:** EDA measurement is carried out on the skin surface by passing a small current through a pair of electrodes placed in skin contact. The principle is of Ohm's Law, which states that resistance across the electrodes is equal to the the voltage (V) applied across the electrodes divided by current (I) being passed through the skin, i.e.  $R$

=  $V/I$ . Lykken et al. [137] and Boucsein et al. [138] strongly argued measuring SC directly by applying constant voltage, and most commercially available devices use this principle [113]. The preferred areas of placing EDA electrodes are palms of hands, soles of feet, medial and distal phalanges of the hand fingers [135, 113]. Although, experimental validations by Poh et al. [139] found that distal forearm recordings have good correlations with finger EDA, although diminished in amplitude. Van Dooren et al. [140] studied the correlation of with traditional sites of EDA recordings from fingertips with 16 other sites on the body that are better suited for long-term ambulatory monitoring, and found that recordings from foot and shoulders have good correlation with recordings from fingertips especially during emotional, physical and cognitive stress . Some important considerations to be mindful of during data collection are the size of the contact area [113], force applied to the electrodes [141], left hand–right hand laterality [142, 113], measurement site responsivity (distal vs. medial phalanges vs. wrist) [113], temperature, humidity and diurnal variations [113].

For our work, we have identified two commercially available devices for reliably recording EDA signal: (i) BioEmo Sensor from BioControl Systems<sup>1</sup> (ii) Empatica E4 Smartwatch from Empatica Inc.<sup>2</sup> Both of these devices are constant-voltage exosomatic measurement devices, which pass small amount of direct-current through the skin, proportional to the skin conductance across two electrical terminals placed in contact with the skin at the distal phalanges and the distal forearm (i.e. interior skin at the wrist) of the non-dominant hand, respectively. The amplified sensor reading is converted to digital format using analog-to-digital converters (ADCs), transmitted through Bluetooth to a nearby computer to be stored and further analysed.

**Individual Differences in EDA:** Individual differences in EDA are relatively more consistent, and have been shown to be reliably associated with behavioral differences or some

<sup>1</sup>[www.biocontrol.com](http://www.biocontrol.com), accessed 04/28/2018

<sup>2</sup>[www.empatica.com](http://www.empatica.com), accessed 04/28/2018

psychopathological differences [113]. Certain specific characteristics of EDA such as number of NS-SCRs and rate of SCR habituation are collectively termed as “Electrodermal lability”, and are known to be broadly consistent within groups of individuals portraying similar behavioral traits. *EDA labiles* are individuals demonstrating high rates of NS-SCRs and slow habituation, and *EDA stabiles* are those showing low rates of NS-SCRs and faster habituation. For instance, in [112], Crider report that “greater EDA lability is associated with undemonstrative and agreeable dispensation, whereas, greater EDA stability is associated with expressive and antagonistic dispensation”. EDA lability has reportedly been consistently correlated with personality traits, information processing abilities, vigilance and perceptual speed [113]. We hypothesize that this might play an important predictor of susceptibility to *technostress* and help us in characterizing groups of people for their expected responses for service failure in CAIE.

#### 2.4.1.2 Algorithmic Analysis of EDA

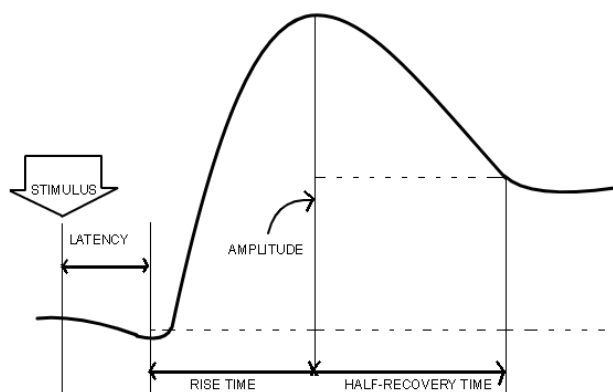


Figure 2.4: Graphical Representation of Phasic SCR. (Picture redrawn from [113, pp 165])

We have already discussed that skin-conductance signal can be thought of as a fast changing phasic SCR value superimposed on a slow changing tonic SCL. Tonic SCL generates a constantly changing baseline within an individual over time, and can differ considerably between different individuals. From this, Boucsein concludes that actual SCL level is of little consequence, nor is it easy to derive [135, 130]. In order to get

an acceptable estimation of the tonic arousal in an EDA recording, “at the very least, phasic

SCR amplitude must be subtracted from the tonic SCL” [130]. Phasic SCRs are superimposed small variations on the broader tidal drifts of SCL [137].

Presentation of novel unexpected stimuli is known to elicit “event-related” SCR (ER-SCRs), which is known to occur in a window of 1s-3s (values derived based on frequency distributions of observed SCRs) following the onset of the stimulus, however, effects have also been reported to be in longer windows of time [143]. All other SCRs outside this window are called “non-specific” SCR (NS-SCR). This window based segregation of ER-SCR and NS-SCR is widely practiced, as noted in [113, 136]. Apart from the relative amplitudes, another measure of background tonic EDA activity is the count of non-specific SCR peaks (typically 1-5 per minute during rest and close to 20 per minute during high arousal) [135, 130]. Additionally, amplitude and standard deviation of NS-SCR peaks are valuable indicators of underlying tonic arousal processes [130].

**Features of EDA Signal:** Various measures are derived from the SCR and SCL components of EDA signals for computational analysis as listed here. (i) Tonic SCL is known to vary widely between and within same subject based on different psychological states. Computing *log* transformation SCL can reduce skew and kurtosis significantly [113]. (ii) It is common for tonic SCL to decrease gradually while subject is at rest, increase when novel stimulus is presented and then decrease again. (iii) SCR amplitude, when found to be positively skewed, show kurtosis or problems with homogeneity of variance, *log* or *square root* (i.e.  $\sqrt{SCR}$ ) transformations have been found to alleviate the problem, however, it is not always necessary. (iv) There are various other measures of ER-SCR shape that are informative of the EDA characteristics such as amplitude, latency, rise time, half-recovery time etcetera, as shown in Figure 2.4 and Table 2.5.

To account for inter-individual differences, it is a common practice to normalize the EDA time-series data. No universally accepted method exists for normalization, some of them

Measure	Definitions	Typical Values
SCL	Tonic level of electrical conductivity of skin	2-20 $\mu$ S
Change in SCL	Gradual changes in SCL measured at two points in time	1-3 $\mu$ S
Frequency of NS-SCR	Number of SCR in absence of stimulus	1-3 per minute
SCR Amplitude	Phasic increase in amplitude shortly after stimulus	0.1-1 $\mu$ S
SCR Latency	Time difference between stimulus and SCR onset	1-3 seconds
SCR Rise-time	Time interval between SCR onset and peak	1-3 seconds
SCR Half-recovery time	Time interval between SCR peak and 50% amplitude fall	2-10 seconds
SCR Habituation	Number of stimulus presentation before no response	2 - 8 stimulus
SCR Habituation Slope	Rate of change of ER-SCR amplitude	0.01 - 0.5 $\mu$ S

Table 2.5: EDA Measures and Typical Values. (Table adapted from [113, pp 165])

are even controversial (such as range-correction due to the use of startle responses which are not similar to the experiment domain [130]). As will be seen in Chapter 4, we apply the recommended *z-normalization* which converts SCR values to Z-scores with *mean* of 0 and *standard-deviation* of 1 or to T-scores with *mean* of 50 and *standard-deviation* of 10 [135, 130].

## 2.4.2 Heart-Rate and Heart-Rate Variability

We might recall from the section on “Neurobiology of Stress” (refer Section 2.3.1) as well as in [96, Table 2.3], the cardiovascular system (consisting of HR, peripheral blood flow and blood pressure), is affected immediately following the appraisal of a stressor. Measurement of the heart-rate is the most commonly used method to monitor changes in the cardiovascular system. Heart-rate variability (HRV) is the variation in the interval of consecutive heartbeats (or in other words, oscillation in heart-rate calculated at each beat), and it is known to be a



good indication of mental effort and stress in adults [144]. However, it must be noted here that these temporal changes of beat-to-beat intervals have good correlation with respiration—the so called *respiratory-sinus-arrhythmia* (RSA)—and are a reflection of changes in cardiac autonomic activation [145]. RSA is known to vary with age and physical activity, which in turn modulates the autonomic activation of the heart [146]. Although concurring views allude the research community on the exact contributions of SNS and PSNS towards causing HRV, numerous time and frequency domain techniques have been studied over the years, and HRV has been related with emotional states, emotion regulation [147], mental workload [148] and cognitive stress and anxiety [149].

#### 2.4.2.1 Physiology of Cardiovascular System

The human heart is a mechanical pump for the blood, which receives electrical signals from autonomic innervations from both the sympathetic as well as parasympathetic divisions. These signals cause the heart muscles to contract and expand, following a rhythmic pattern. RSA is the observed increase in HR (short R-R intervals) during inhalation and decrease in HR (long R-R intervals) during exhalation. However, it must be noted here that HRV and RSA are not exactly the same, but are often used interchangeably [145]. Over the years, various physiological phenomena have been surmised to be causing HRV such as, central neural activation, reflex activation of lungs, mechanical changes in thoracic pressure during respiration [145]. However, with systematic experimental evaluation it is now clear that RSA at any given moment is a complex function of the activation of cardiac vagus nerve, SNS (increases the HR), PSNS (decreases the HR), mechanical as well pacemaker cells located in the sinoatrial node [145], although their exact roles are not yet conclusively agreed upon [150].

In response to psychosocial stress, direct sympathetic neural activation causes epinephrine

to be released in the blood stream which is detected by the ventricles of the heart, and respond by increased speed and force of ventricular contraction [151]. As a result of these, vasoconstriction follows, in effect reducing the blood-flow to the extremes of the body such as fingertips, forehead and toes. The decline in the blood flow results in drop in skin temperature, though it has been known to be not too reliable measure of peripheral blood-flow [151].

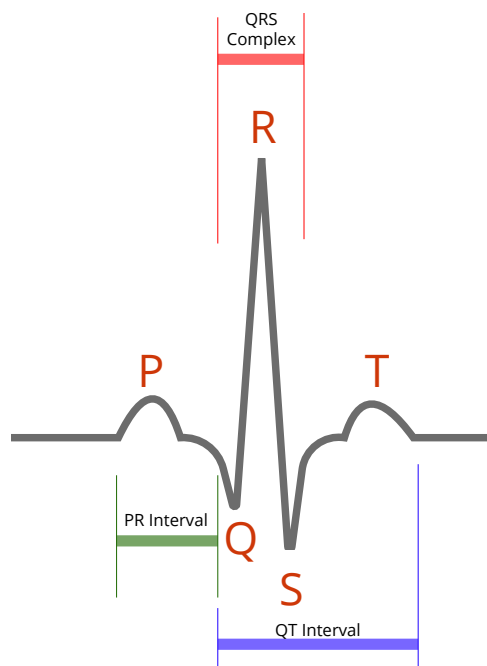


Figure 2.5: QRS-Complex waveform in an ECG plot<sup>3</sup>. Usual lengths: P-wave (0.08-0.10 s), QRS (0.06-0.10 s), PR-interval (0.12-0.20 s), and QT-interval ( $QT/\sqrt{RR} \leq 0.44s$ ) [144]

**Measurement of Heart Pulse:** Various methods exist in practice of precisely measuring the period of the cardiac cycle, for instance, phonocardiogram (PCG) measures the heart-beat sound, whereas echocardiogram produces a visual representation of the beating heart using ultrasound. The two most common methods of measuring heart-rate are electrocardiography (ECG) and photoplethysmography (PPG). While ECG measures depolarized electrical changes of muscular contraction associated with cardiovascular activity [144], PPG measures the blood flow at certain specific sites on the body such as fingertips, toes, calves and works on the principle of light absorption characteristics of the blood at different optical frequencies. While both of these methods are non-intrusive, ECG measurement is a bit more in-

involved with respect to access to measurement sites as well as device setup, compared to PPG measurement. For our current work, we have identified two commercially available

<sup>3</sup>Redrawn from [www.commonswiki.org/wiki/File:SinusRhythmLabels.svg](http://www.commonswiki.org/wiki/File:SinusRhythmLabels.svg)

systems for reliably recording ECG and PPG: (i) BioBeat Sensor from BioControl Systems <sup>4</sup> which is a chest-ECG measurement system and (ii) Empatica E4 Smartwatch from Empatica Inc <sup>5</sup> which is a PPG sensor in wristband form factor.

The wave depolarization that starts at the apex of the right-atrium and ends in the apex of right ventricle, the body being the conductor transmitting this electrical pulse. This electrocardiac signal can be measured at the surface of the body by taking the potential difference across two or more electrodes placed at either sides of the heart at specific sites on the chest, guided by Einthoven's triangle [152], which is a diagram showing on-body electrical fields during various cardiac cycles with the heart at the center. Heart pulse waveform obtained from ECG (see Figure 2.5) is quite detailed, however, Berntson et al. postulate that most of the requirements for HRV analysis should be fulfilled by a simple data stream showing large R-wave peaks [153], such as photoplethysmogram (PPG) as described below. ECG is known to be affected by movement artifacts (produces baseline changes) and power-line interference (50Hz or 60Hz) [152].

Plethysmography is a technique to measure the blood-flow under an organ of interest—such as the heart—by measuring the change in pressure in the blood-vessel walls caused by the blood ejected during cardiac cycles. Photoplethysmography (PPG) is a method that attains this by shining fixed wavelength infra-red light onto the skin surface at sites mentioned above, and detecting either the transmitted or the reflected light pulse, to finally measure the light absorption in the blood due to the oxygen saturation of the blood [152]. PPG, although is immune from electrical interferences, it is very sensitive to motion artifacts, and special algorithmic approaches are needed for data processing [154, 155, 156].

<sup>4</sup>[www.biocontrol.com](http://www.biocontrol.com), accessed 04/28/2018

<sup>5</sup>[www.empatica.com](http://www.empatica.com), accessed 04/28/2018

Variable	Unit	Definition
<i>Time Domain Features</i>		
SDNN	ms	SD of all normal R-R intervals
SDANN	ms	SD of the average normal R-R intervals calculated over short time periods (usually 5 mins)
RMSSD	ms	Square root of mean squared difference between adjacent normal R-R interval
SDNN Index	ms	Mean of SD of normal R-R intervals calculated over short time periods
NN50		Number of pairs of adjacent normal R-R intervals
pNN50	%	NN50 divided by total number of R-R intervals
HRV Triangular Index		Number of normal R-R intervals divided by height of the histogram of all normal R-R intervals measured on discrete scale with bins of 1/128 s
TINN	ms	Baseline width of minimum square difference of triangular interpolation of the highest peak of the histogram of all normal R-R intervals
<i>Frequency Domain Features</i>		
Total Power	$ms^2$	Area under the entire power spectral curve (<0.4 Hz)
ULF	$ms^2$	Power in ultra low frequency band (<0.003Hz)
VLF	$ms^2$	Power in very low frequency band (0.003 - 0.04 Hz)
LF	$ms^2$	Power in low frequency band (0.04 - 0.15 Hz)
HF	$ms^2$	Power in high frequency band (0.15 - 0.4 Hz)
LFnu	nu	Normalized low frequency power (LF/LF+HF)
HFnu	nu	Normalized high frequency power (HF/LF+HF)
LF/HF		Ratio of low- to high- frequency power

Table 2.6: Conventional Heart-Rate-Variability Features. Table reproduced from [145]

### 2.4.2.2 Algorithmic Analysis of HRV

ECG waveform consists of the QRS-complex waveform formed during the various phases of contraction of the ventricles, as shown in Figure 2.5. HRV is one of the oft-used methods of ECG analysis, wherein the R-R interval (or normal-to-normal i.e. N-N interval) of each consecutive beats are detected from the QRS-complex are determined. Detection of the QRS-complex (see Figure 2.5) from the ECG time-series is well-studied problem, and a standard Pan and Tompkins algorithm [157] is widely used for R-wave detection and subsequent HRV time-series generation. This HRV time-series is used to derive various time, frequency and non-linear features.

**HRV Time-Domain Features :** Time domain features are commonly used statistical features such as mean, standard-deviation (SD) of N-N intervals (SDNN), the SD of the 1st difference of HRV time-series, count and percentage of total beats in a window having more than 50ms as N-N interval i.e. NN50 and pNN50, count and percentage of total beats in a window having more than 20ms as N-N interval i.e. NN20 and (pNN20). The SDNN feature measures the total variability arising from periodic and random sources (similar to power in specific spectral bands in frequency analysis) [145].

**HRV Spectral Features :** Spectral analysis of HRV time-series is performed by decomposing the power (or total variance) of a continuous series of heart beats into frequency-components [145] and then spectral power for a given frequency band can be determined by computing the area-under the power-spectral-density curve. Please refer to Figure 2.6 for an example of an HRV power spectrum plot.

Two common methods of spectral analysis are Fast-Fourier Transform (FFT) which assumes only deterministic components in the time-series, and Autoregressive (AR) modeling which has no such restrictions i.e. the data can be composed of deterministic as

well as stochastic components. For short duration recordings (2min range), three main peaks in the power spectrum are often identified: (i) Very-Low Frequency (VLF) band [ $< 0.04\text{Hz}$ ] (ii) Low-Frequency (LF) band [ $0.04\text{Hz}-0.15\text{Hz}$ ] (iii) High-Frequency (HF) band [ $0.15\text{Hz}-0.4\text{Hz}$ ]. Notice the peaks in each frequency bands in Figure 2.6. The power in LF and HF bands are computed in normalized units (nu), by dividing the power in each band

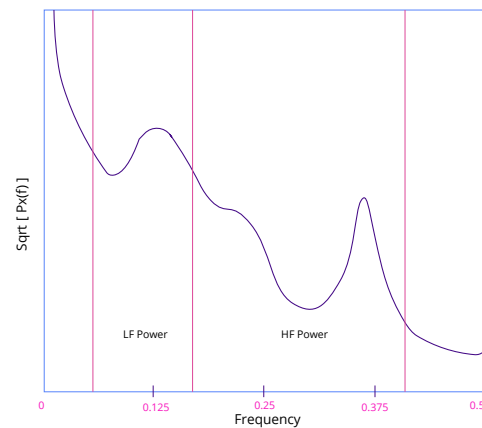


Figure 2.6: Example of frequency spectrum of an Heart-Rate Variability time-series, marked with LF and HF regions. Figure redrawn from [158]

by total power in LF + HF bands [145]. Finally the ratio of LF to HF power is significant as it is thought to depict the index of sympatho-vagal activation balance, although there are various studies pointing towards limitations of this ratio for inferring an index of sympathetic activation [150, 159].

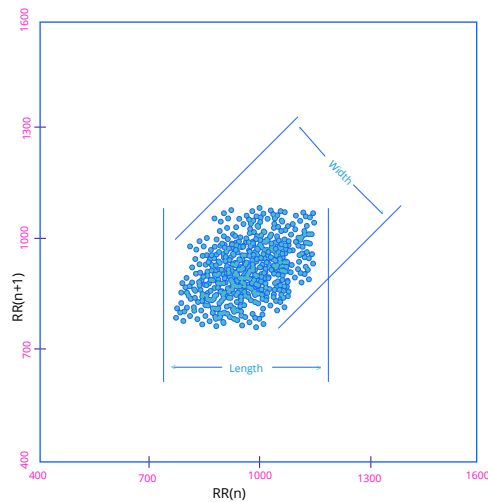


Figure 2.7: Example Poincaré Plot of a Heart-Rate Variability Spectrum similar to the one shown in Figure 2.6 (Figure redrawn from [158])

**HRV Geometric Features :** Poincaré geometry indices are increasingly being used to capture the dynamics of fluctuations in HRV interbeat intervals [144, 160, 161, 145], owing to their utility in characterization of complex non-linear organic systems. Poincaré plot is a scatter-plot representing the value of each pair of consecutive R-R interval plotted in a simplified phase-space or Cartesian plane, and an ellipse is fitted for quantitative analysis of the scatter of the system [160]. A series of these consecutive points on the Poincaré plot represent a curve showing a system's evolution. Some derived features include:

(i) minor axis of the ellipse or SD1, representing the

SD of the instantaneous changes in HRV. Physiologically it signifies the index of parasympathetic activation, as it is known that the vagal effect on sinus node supersedes the sympathetically mediated effects. (ii) major axis of the ellipse or SD2, representing the standard deviation of the long-term HRV. Physiologically it signifies both the sympathetic and parasympathetic tones (iii) the relation of minor axis to major axis or SD1/SD2 representing the index of parasympathetic activation compared to sympathetic activation.. This method essentially quantifies the temporal changes in vagal and sympathetic activation of the HRV time-series without the requirement of *stationarity* imposed on the data, which is rarely true

for a complex system such as the heart [160]. Another useful geometric analysis technique is finding the HRV Triangular Index, wherein the series of NN intervals are plotted as a geometric shape such as a triangle distribution and the measure of an interpolated shape such as the base of the triangle is used to signify the variance [145].

**HRV Non-Linear Features :** Deterministic *chaos* in biological systems promotes stability (variation within limits) and flexibility (multiple x-value for single y-value), properties that allows living organisms to maintain a stable internal environment, i.e. *homeostasis*, as it adapts to changes in environmental demands [145]. In the late 20th century, various evidences conjectured that biological processes in our cardiovascular systems do not follow regular periodic oscillation, but rather operate under non-linear dynamic behavior. Thus, linear statistical and spectral analysis may not provide the sensitivity needed to model subtle changes in the HRV time-series. Over the years, multiple researchers have used Chaos Theory and Fractal mathematics to describe HRV dynamics and complexity, for instance, heart rate frequency ( $f$ ) follows an inverse-power law relation ( $1/f$ )—a defining characteristic of fractals, Detrended Fluctuation Analysis which detects the existence of fractal-like properties in HRV series, Lyapunov exponent analysis, multiscale or approximate entropy measures to name a few [145, 160]. These non-linear indices of HRV series, often are better predictors of adverse cardiovascular events than traditional statistical methods [145].

Features such as Poincaré graphical indices and various non-linear HRV features as noted in Section 2.4.2.2, can be extracted with the aid of already existing tools such as Kubios HRV [162] analysis software<sup>6</sup> v2.2 (Kuopio University, Finland) and Chaos Data Analyzer<sup>7</sup> Professional (CDA Pro), v 2.2 (J.C. Sprott, University of Wisconsin, USA).

<sup>6</sup>[www.kubios.uef.fi](http://www.kubios.uef.fi), accessed 10/20/2016

<sup>7</sup>[www.sprott.physics.wisc.edu/cda.htm](http://www.sprott.physics.wisc.edu/cda.htm), accessed 20/20/2016

### 2.4.3 Computational Recognition of Technostress

In this section, we will present some computational methods that we have identified as suitable for modelling and inferring technostressed states from human physiological signals described in the previous two sections. We have already identified the most prolific statistical features as well as various other surrogates for these features, a small subset of which are mentioned in Table 2.5 and Table 2.6. Here, we will see details of a typical Machine Learning (ML) pipeline consisting of techniques in feature extraction and selection, known ML algorithms used in similar scenarios, as well as a subbranch of ML called *transfer learning* which is a good candidate approach for implementing physiological profile-learning component of our framework proposed in Section 3.2.

**Computational Pipeline for ML:** Computational pipelines for analysis of physiological signals elicited as a result of discrete-stimuli follow a windowing approach wherein the physiological signal in question (e.g. EDA, HRV, etc.) are only analyzed for a fixed window of time following the presentation of each stimuli. We have already seen in Section 2.4.1.2, that SCRs occurring within a window of *1second* - *5second* following a stimulus can be considered as ER-SCRs [113, 135, 136]. Recent works use an approach of a measurement window, that has to be long enough to capture most relevant ER-SCRs, while short enough to exclude capturing NS-SCRs as well as effects from subsequent stimuli [143]. According to a report by Figner et al. [143], a common window of interest may be considered to start from *1second* and end at *6second* after the stimulus onset.

Once the window for data processing has been defined, the above mentioned features are extracted and concatenated together to form the feature vector  $\mathbf{x}_i^\top \in \mathcal{R}^D$  for each observation, where  $D$  is the total number of features, and  $\top$  is the transpose operator. The data matrix consisting of  $\mathcal{N}$  such observations is represented by  $\mathcal{X} \in \mathcal{R}^{\mathcal{N} \times D}$ . When this data



matrix is a part of a supervised learning task having a total of  $\mathcal{C}$  classes,  $\mathcal{X}$  is combined with  $1 \times \mathcal{N}$ -dimensional class-label vector  $\mathcal{Y}$  such that (s.t.)  $y_i \in \{c_1, c_2, \dots, c_C\}$  to form the dataset  $\mathcal{S} = [\mathcal{X}, \mathcal{Y}]$ .

### 2.4.3.1 Notes on Dimensionality Reduction

Machine learning tasks can be collectively characterized as a quest for similarity searching [163, pp 485]. One of the basic assumptions about the data matrix  $\mathcal{X}$  for various data analysis algorithms is that the columns should be uncorrelated. However, the features (or dimensions) are derived from physiological signals produced from a complex set of highly-correlated biological processes (as described in Section 2.3.1), thereby producing correlated features [77]. In addition, there are various intrinsic computational problems related with high-dimensional data, for instance, the number of samples needed for accurately estimating an arbitrary function grows exponentially with the number of features that the samples comprise of [163, pp 486]. For similarity searching this translates to an exponential growth of the search space with number of dimensions. As Hastie et al. state in [164, ch 18], similar problems arise for a data matrix having more number of features compared to the number of observations, i.e.  $\mathcal{X}$  with  $D \gg N$ . Apart from these issues, searching in a low-dimensional feature sets is computationally tractable problem. Hence, it is necessary to find a low dimensional representation of the feature vector by eliminating the correlated information in the redundant features, while preserving the discriminatory information present in the data [165].

Dimensionality reduction (DR) and feature selection (FS) are two ways to obtain a reduced feature representation [166]. FS removes the redundant features from the dataset as they may impact the final classification accuracy, while DR produces a transformed combination of the original features. One disadvantage of DR methods is the lack of interpretability of

the new transformed dimensions [166]. These techniques can be supervised (i.e. dependent on the problem at hand, e.g. classification or regression task) or unsupervised (i.e. agnostic to the current problem, producing reduced feature-sets blindly based only on some intrinsic property of the data such as covariance). Depending on the type of problem and relation between the features, linear (LDR) or non-linear (NLDR) may be useful [165].

#### 2.4.3.1.1 Principal Component Analysis

The Karhunen-Loève transform (commonly known as Principal Component Analysis (PCA)) is a very popular method of unsupervised feature transformation that produces a linear combination of the original features to produce an orthogonal low-dimensional representation while preserving most of the covariance of the original data-set [167, pp 331]. The addition of a PCA based feature-preprocessing step is known to produce smaller error for various pattern-recognition algorithms such linear Support Vector Machines (SVM), SVM with radial basis function (RBF) kernel, k-Nearest Neighbor (k-NN) [166]. For a data matrix  $\mathcal{X} \in \mathcal{R}^{N \times D}$  with  $N$  as the number of observations having  $D$  features each, a standard PCA process chooses the first  $c$  eigenvalues ( $c < D$ ) of the ordered list of eigenvalues of matrix  $cov(\mathcal{X})$ , thereby reducing the dimensions of the  $D$ -dimensional dataset. Please see [165] for a step-by-step guide on PCA. NLDR methods such as kernel-PCA, neural network based NLDR [165] might be helpful.

#### 2.4.3.2 Notes on Machine Recognition of Stressed States

ML algorithms such as SVM, Decision Trees (D-tree), Random-Forests (RF), Multi-class Classifier (MCC) and Adaboost have been successfully used for computational recognition of stressed states such as frustration [132]. Prior work on physiology based stress-recognition

has demonstrated SVM outperforms various other ML algorithms [74, 75, 77, 76]. However, in a recent work, Colomer et al. compared the performance of these algorithms with a gamut of physiological features for emotion recognition and found that Adaboost with RF outperforms all of the above ML algorithms on their dataset on audiovisual stimuli [132]. In this section, we will present a short survey of these machine learning techniques which we plan to use in our work for recognition of *technostressed* states.

#### 2.4.3.2.1 Support-Vector Machines (SVM)

Support Vector Machine is a very widely used linear discriminative classification algorithm which, being data distribution independent, is known to successfully classify a wide variety of problems with good accuracy. SVM is predominantly a binary classifier where the primary objective is to come up with a maximum margin classifying high-dimensional hyperplane between the classes such that there are minimum number of support vectors inside the margin, where margin is defined as the minimum distance between the classifying hyperplane and a point in the dataset. So for a training dataset of labeled points  $\mathbf{S} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^n$  with  $y_i \in \{+1, -1\}$ , soft-margin SVM has the following dual formulation :

$$\begin{aligned} \text{Objective : } \max_{\alpha} L_{dual} &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{K}(x_i, x_j) \\ \text{Constraints : } &0 \leq \alpha_i \leq C \quad \forall i \in \mathbf{S} \quad \text{and} \quad \sum_{i=1}^n \alpha_i y_i = 0 \end{aligned} \quad (2.1)$$

where  $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)$  is the kernel function used to map data vectors to a more expressive feature space which aids in classification of non-linear datasets. For a detailed geometrical analysis on SVM, please see [168, Ch 21].

### 2.4.3.3 Notes on Transfer Learning (TL) Paradigm

The mission specified by Defense Advanced Research Projects Agency (DARPA) for Transfer Learning (TL) is “the ability of a system to recognize and apply knowledge and skills learned in previous tasks to novel tasks” [169]. Thus, TL aims to extract knowledge from a *source task* and applies the knowledge to a *target task*.

Transfer learning literature follows a convention as noted below:

- (i) A *domain*  $\mathbb{D} = \{\mathcal{X}, P(X)\}$  comprises of a feature space  $\mathcal{X}$  and a marginal probability distribution of the data  $P(X)$  where  $X = \{x_1, \dots, x_n\} \in \mathcal{X}$ .
- (ii) A *task*  $\mathbb{T} = \{\mathcal{Y}, f(\cdot)\}$  comprises of a label space  $\mathcal{Y}$  and a predictive function  $f(\cdot)$ , which is not observed and has to be learned from the training dataset  $\mathcal{S} = \{x_i, y_i\}$  where  $x_i \in \mathcal{X}$  and  $y_i \in \mathcal{Y}$
- (iii) Given a source domain  $\mathbb{D}_S$  and learning task  $\mathbb{T}_S$ , a target domain  $\mathbb{D}_T$  and learning task  $\mathbb{T}_T$ , a transfer learning approach aims to improve the learning performance on  $f(\cdot)$  in  $\mathbb{T}_T$ , where either  $\mathbb{D}_S \neq \mathbb{D}_T$  or  $\mathbb{T}_S \neq \mathbb{T}_T$ .

For this current work, the observations from Paced Stroop Test Paced Stroop Test (PST) or responses from other physiological learning stimuli form the  $\mathbb{D}_S$  and the observations from the user’s responses from CAIE form the  $\mathbb{D}_T$ , whereas,  $\mathbb{T}_S$  and  $\mathbb{T}_T$  are the cognitive stress and technostress learning tasks, respectively. Pan et al. [169] provide a detailed classification of the types of TL and when to use which approach (please see Table 2.7).

Following the Table 2.7, for this current work we hypothesize that both *inductive TL* and *transductive TL* approaches can be applied. This follows from the following reasoning:

- (a) *transductive TL* because— $\mathbb{D}_S$  and  $\mathbb{D}_T$  are “different but related” as the nature of stress generated from ground-truth collection techniques described in Section 3.2 as well as tech-

TL Setting		$\mathbb{D}_S$ and $\mathbb{D}_T$	$\mathbb{T}_S$ and $\mathbb{T}_T$	$\mathbb{D}_S$ Labels	$\mathbb{D}_T$ Labels
Traditional ML		Same	Same	Available	Unavailable
Inductive TL	Multi-task Learning	May or May not be Same	Different but Related	Available	Available
	Self-taught Learning			Unavailable	Available
Transductive TL	Domain-adaptation, Sample-selection Bias	Different but Related	Same	Available	Unavailable
Unsupervised ML		Different but Related	Different but Related	Unavailable	Unavailable

Table 2.7: Comparison of Various Transfer Learning approaches

nostressors may be different.  $\mathbb{T}_S$  and  $\mathbb{T}_T$  are “same” as the tasks remain the same i.e. the correct detection of stressed states. (b) *inductive TL* because— $\mathbb{D}_S$  and  $\mathbb{D}_T$  are not the same (as argued previously), in addition,  $\mathbb{T}_S$  and  $\mathbb{T}_T$  may also be considered to be different owing to the different types of stress states, i.e. the functional mapping  $f(\cdot)_S$  and  $f(\cdot)_T$  may be different. In addition, we might consider the availability of some labeled data Thus, techniques such as a modified TrAdaboost algorithm as introduced by Dai et al. [170], or improving SVM accuracy by training on auxiliary data similar to Wu et al. [171].

#### 2.4.3.3.1 TrAdaboost: Adaboost for Transfer Learning

*TrAdaboost* algorithm assumes that  $\mathcal{X}_S$  and  $\mathcal{X}_T$  may consist of exactly similar features and labels, however, their distribution over the data are different, i.e.  $P(X)_S \neq P(X)_T$ . In addition, Dai et al. address the possibility that due to this difference in data distribution, some instances from the source domain data may not enhance the classifier performance for prediction on the target domain, which might even harm the classifier [169]. To overcome this, *TrAdaboost* iteratively re-weights the source-domain data to reduce the effect of the “bad” source instances while rewarding the “good” source-data towards improving target-domain classification. Then, for each iteration, the *TrAdaboost* trains the base classifier on this weighted dataset, while the error is only calculated on the target data. The strategy used by *TrAdaboost* is same as the basic *Adaboost* to update the incorrectly classified examples

in the target domain, while it applies a different strategy from *Adaboost* for updating the incorrectly classified source examples in source domain.

In this chapter, till now we have described the symbiotic relation between affect-aware and context-aware computing systems, and delved deeper into the physiological features of the affective state called technostress. We have described methods of computational psychophysiology which can be used to recognize technostress from its physiological signature. In our framework, as described in Chapter 3, we envision using this inference on technostress as method of incorporating implicit affective feedback in a CAIE. In the next section, we describe an example of a potential application domain where we envision the use of our physiology based implicit feedback to improve interaction quality. We have identified the lack of a sense of agency in modern buildings as a potential problem, which our model of technostress based feedback may provide a remedy to.

## 2.5 Implicit Feedback in Human-Building Interaction

Alavi et al. introduced the term *home-biosis* to explain the symbiotic relationship between inhabitants and the building they occupy which is mostly driven by the advancements in artificial intelligence giving modern buildings a new kind of autonomy [172]. Unlike traditional HCI with more well-defined modalities of interaction, a user interacting with a modern building can not terminate an interactive session without abandoning the physical space in case of a negative interaction [1]. Consequently there is a need to thoughtfully design the user-experience for such interactive contexts. From an HCI perspective, an intelligent building scenario can be thought as an extreme amalgamation of *ambient intelligent* (AmI) system and a *tangible interface* [1]. However, it must be noted that the emphasis of an AmI system mainly lies on the interactions with digital artifacts with building as a backdrop and hav-

ing little consideration for user's *comfort*—a space quality that needs to be guaranteed for dwelling in the building context.

### 2.5.1 Perceived Comfort in Built Environments

A ubiquitous technology in the context of modern buildings are Building Automation Systems (BAS), which often prioritize energy conservation and are motivated by optimization of technological installations [1]. Energy conservation based BAS systems use quantitative metrics of performance, with little consideration for uncertainty and diversity of user behavior, inadvertently resulting in negative user experiences such as automated window blinds with seemingly erratic behavior; a feeling of being controlled by the building or sick-building syndrome [1]. With the projected trajectory of technological advancements, such automation systems will progressively be employed in our homes. These situations could be averted if the automation systems were enhanced by an understanding of the inhabitant's *comfort* instead of optimizing only on the basis of energy conservation [1].

For the thermal variables, perceived user comfort is measured at a building scale by Fanger model [173] which measures user's discontent with a score called Predicted Mean Vote (PMV) as a function of building variables such as air and radiant temperature, humidity etc. as well as user variables such as activity, metabolic rate and clothing [1, 174]. However, PMV has been shown to be inadequate for modeling thermal comfort for small groups of people or buildings without centrally controlled systems (very common in a *home* context) [174]. Recently two groups have independently used wearable devices for modelling thermal comfort [175, 174]. Huang et al. have proposed to address this problem recently by utilizing off-the-shelf wearable wristbands measuring physiological signals to model individual perceived thermal comfort using skin- and near-body temperature (to model ambient temperature),

skin conductance (to model skin sweat) and heart-rate, step count and estimated calorie consumption (to model activity level) [174]. Ranjan et al. have demonstrated using infrared thermal imaging to derive a model of user comfort which outperforms BAS on the energy conservation metric [176].

### 2.5.2 Frustration and Lack of Agency

In addition to user's perceived comfort, Nembrini et al. argue that built environments may also induce negative emotions such as *annoyance* or *frustration* due to automated services whose behaviors do not match a user's mental model, such as behavior of the automatically controlled blinds [1]. Karjalainen et al. have shown that the inability to act upon the environment in order to reconfigure it to one's comfort induces increased frustration [177, 1]. This is similar to Alavi et al.'s argument of provisioning for user's *sense of agency* in order to set or change the preferences of services in human-building interactions (HBI). Such negative emotions may in-turn also influence a user's perception of the environment, akin to assigning personality traits to intelligent homes. For example, Mennicken et al. in [178] imparted two groups of personality traits into a smart home while providing intelligent services —(a) conscientiousness and agreeableness in CKC home taking a passive approach, and (b) extraversion and openness in EC home taking a more proactive approach in providing services. One of their key findings was that the users wanted to feel in control even when interacting with computational agents providing automated services, as the *proactive* home induced a feeling of lack of agency in them [178].

We hypothesize that such negative emotions of *frustration* are not tied to any particular service, and may occur in response to services that produce arbitrary or unexpected behavior for any automated services it may provide (some examples of such services are described in



Section 5.3.1). This is akin to the concept of *technostress*, as alluded-to in the introductory paragraphs. From the systems perspective, being able to recognize such frustrated states may provide a novel way of implicitly communicating user's feedback about the appropriateness of the service to the intelligent system. Thus, a technostress based service-appropriateness feedback signal may be useful in the HBI context, in order to adapt the intelligent services for attaining, and improving, *comfort*. It is important to note here that in this article, we do not intend to study users' *comfort* directly, rather we intend to highlight it as a way of defining an evaluating parameter for the success of our technostress based framework when used in the HBI context.

## 2.6 Conclusion

In this chapter, we have shown the state-of-the-art in interface design for intelligent environments, thereby arguing in favor of how incorporating an affective feedback loop may be the key to more human-centric interfaces for context-aware intelligent environments. Technostress in an ICT interface has been researched intensively for atleast a couple of decades, however, we have pointed out some prevailing research gaps in literature. In addition, we claim that our use of technostress to ascertain a user's approval of an intelligent service, is novel. A brief survey on a holistic model of *context*, assimilating affect and intention in its purview, was presented to bring previous discussions into perspective.

Later, we have described the physiological processes elicited by stress-reactions on the electrodermal and cardiovascular subsystems of autonomic nervous system. We have discussed statistical features relevant to computational analysis of EDA and HRV signals, electrical devices and specification for measurement of these signal-streams, pre-processing steps in applied their computational analysis. Further, we have also presented a short survey

of machine-learning algorithms which we found to be well-suited for stress recognition in general, and we hypothesize may be valuable towards a technostress recognition system. Finally, we have discussed the transfer learning paradigm which gives a framework for learning a user's physiological profile from different proven stressors for ground-truth data collection.

# Chapter 3

## The CAfFEINE Framework

---

*“...rather than making computer-interfaces for people, we want to make people-interfaces for computers...”*

— Michael Coen [179]

---

In the previous chapter, we have discussed the need for a paradigm shift in interaction techniques for CAIE compared to traditional desktop computing, and the challenges it poses

for the designer. We have presented a case for the introduction of affective computing (AC) technologies in intelligent environments (IE). As noted by Klien et al. [57, 8], a mere *perception* of affective understanding in a computer-interaction reduces the user’s perception of frustration. Interesting as these results are, Section 2.1.3 also highlights that such affective feedback loops are an understudied area of intelligent environments. In Section 2.5, we have presented an argument in support of introducing an implicit affective feedback in a human-building interaction domain. In this chapter we will present a detailed overview of the envisioned interaction schema in our CAFFEINE framework and go through its numerous components and methods of evaluation.

## 3.1 Components of the CAFFEINE Framework

We have named our interaction framework for incorporating implicit affective feedback in an intelligent environment as CAFFEINE, which stands for Context-aware Affective Feedback in Engineering Intelligent Naturalistic Environments. While describing CAFFEINE, we will present our scheme of implicit affective feedback in a CAIE using human physiological signals, methods of collecting ground-truth data, methods of evaluating the feasibility of the interaction scheme presented in CAFFEINE, and also propose a method of estimating a quality metric for the technostress inference on the aforesaid affective-feedback.

### 3.1.1 User Feedback in a CAIE

*Intelligent Environments* have been used as an umbrella term by various researchers to define “spaces in which computation is seamlessly used to enhance ordinary activity” (by Coen et al. [179]), and a “physical space in which the actions of numerous networked controllers,

each controlling a specific aspect of the environment, is orchestrated by self-programming pre-emptive processes in such a way as to create an interactive holistic functionality that enhances inhabitant's experiences" (by Augusto et al. [6]). Augusto et al. also notes that an IE should be *intelligent* and *sensible* to identify when and how it should provide a service to the user to help them achieve their current goals while also preserve their privacy, safety and autonomous behavior. To achieve the level of intelligence and autonomy being hinted here, coupled with our goal of designing "human centered interfaces" for a CAIE, we argue that the IE needs to gather user's *preferences* to tailor it's services as well as their *feedback* regarding appropriateness of these services after they are provided. As Coen et al. contend in [179] that Intelligent Environments will be severely limited in scope without incorporating various branches of artificial intelligence for preference learning and interaction design; we argue that lack of any means to gather user's feedback would inhibit communication with and adoption of these computing spaces.

Incorporating user-feedback gathering techniques usually entails creating a channel of communication, either explicit or implicit [9], that helps determine the relevance of the system's intelligent interventions towards fulfilling the user's goals. However, in the context of intelligent environments, provisioning for explicit channels of communication becomes particularly challenging due to the absence of a direct-manipulation based interface [180]. As rightly pointed out by Coen et al. in [179], "The user-interface primitives of these systems are not menus, mice and windows but gesture, speech, affect, and context; their applications are not spreadsheets and word processing but intelligent rooms and personal assistants." In such interfaces with de-centralized focus of attention and interaction, designing mechanisms for gathering feedback unintrusively, becomes challenging.

### 3.1.2 Why Technostress? Hint: *Explicit vs Implicit* Feedback

We have discussed in the introduction chapter that one of the requirements, for a pervasive computational system to disappear from the user's cognitive front, is to acquire the ability to assimilate into the temporal workflow of the user at all possible times without posing a hindrance to them. To achieve this, the computational system may need to continuously infer a user's goals, intentions and instantaneous context from an instrumented environment surrounding the user. A key aspect of an inference-based context-aware system is the probabilistic nature of *recognizing* a user's context and thereafter delivering *suitable* services. As a result, such an intelligent computational system may render itself to situations where it is delivering services which may not align properly with a user's instantaneous needs for achieving his broader goals. This lack of proper understanding of a user's instantaneous expectations from the system may arise from various reasons ranging from dynamic user preferences to improper modeling of a user's context, goals or intentions (which in turn may arise due to limited information).

*Explicit* feedback is a potent way of capturing a user's intentions and needs by proactively asking the user about the appropriateness of the service being currently provided. As discussed in Section 2.1.3, such explicit feedback pathways using questionnaires or facial expressions to convey (dis)like for a particular service, have been tried by Benta et al. [67]. However, as discussed previously, they result in reduced quality of interaction due to the need for orienting towards the camera to trigger the sensing of facial expression. Thus, although reliable, explicit feedback comes at the cost of user's cognitive resources. In a way, this defeats the purpose of having an inference-based context-aware pervasive computational system in the first place, if it has to stop and ask the user each time after delivering the services about the appropriateness of the service.

On the other hand, if the pervasive computational system could infer a user's *(dis)approval* about the service's (in)appropriateness from their implicit behavioral, physiological or physical *cues*, the need for stopping and asking the user could be completely eliminated. An *Implicit* feedback, thus, provides a way to infer user's approval of system interventions by sensing non-verbal cues which often are generated at a subconscious level [71] thereby bypassing the need for their cognitive resources. We found that there is a need for and a gap in the knowledge about incorporating implicit feedback channels in intelligent environments. Such seamless communication with the pervasive computational system, akin to human-to-human implicit communication, will in essence improve the overall usability and adoption of such systems.

Brave et al. argue that interfaces which lack the capacity to understand and reciprocate emotions, can dramatically impede performance [49]. Section 2.1 amply demonstrates the existence of a gap in knowledge on methods to incorporate such implicit feedback channels in an IE. We plan to address this gap by using techniques developed in Affective Computing (AC) paradigm, which we think are potent methods for designing an implicit feedback channel. This follows from the argument presented widely over the last decade that though computers are inanimate objects, humans naturally tend to treat computers like social agents [181, 49]. Bourguet rightly notes that "in particular, user's expectations of a system's capabilities and user's mental models of how a multi-modal interface works are often inadequate, and these trigger interaction problems" [182]. This "interaction problem", when appraised by the user, imparts a temporary "technostress", which is the phenomenon we want to exploit to infer a user's approval of the services provided by the IE. As described in Section 2.3, the phenomenon of technostress is an appropriate non-verbal pre-conscious event-tied reactionary process that stimulates various biological and nervous system producing measurable on-body changes. Hence, sensing these changes following any given event in an IE, provides

for a reasonable method to infer a user’s subjective (dis)like of a service’s (in)appropriateness. Next we will discuss the interaction paradigm in an affect-aware CAIE design.

### 3.1.3 Implicit Feedback using Technostress

A schematic for a CAIE is depicted in Figure 3.1 which operates by sensing a user’s context employing various sensors embedded in the IE to detect constituent variables such as environmental (e.g., user identity, time, location and nearby objects), physical (e.g., current activity being performed by the user, intentions and goals set by the user), social (e.g., other users nearby, their relation with current user), and affective (e.g., users psy-

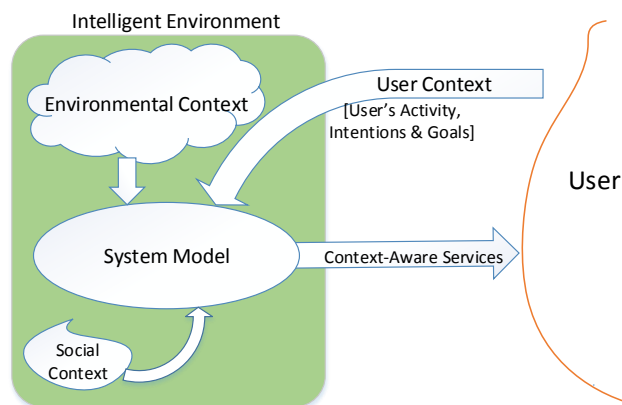


Figure 3.1: Intelligent services provided by a CAIE on sensing a user’s contextual cues

chological state arising from appraisal of progress towards goals). A detailed discussion on the constituents of a user’s *context* can be found in Section 2.2.2. Social context is a very important aspect, especially for intelligent environments that are supposed to be cohabitated by multiple users. We argue that future CAIEs may benefit from incorporating a sense of social-context to avoid socially awkward situations such as the recent AI-chatbot-Twitter-fiasco by Microsoft Tay [183, 184].

On successful identification of a user’s context with reasonable confidence, intelligent services are provided by the CAIE, to assist them in their current activity. This assistance can be in the form of either a *proactive* or a *reactive* service. An example of a proactive service would be: *an intelligent reminder to get milk on the way back from your office as the system suggests*



you take a different route to home due to heavy traffic on your usual route. An example of a reactive service would be: *the inhabitant asking over a voice prompt for instructions on how to cook paneer-butter-masala*. A schematic for this process of technostress generation is shown in Figure 3.1.

There is often a distinct mismatch between a user’s mental model (or expectation) of how a system should work vis-a-vis what the system designers thought appropriate at the time of design freeze [182]. This mismatch in the expectation, accompanied by unforeseen computing bugs and system failure raises various interaction issues feeding into the main reasons for a phenomenon called *technostress*. As already discussed in Section 2.3, *technostress* is a user’s physiological response to stressors from computer-based-interactions such as unexpected network delay, system crash or unexpected behavior such as unresponsive hardware etc. A schematic for this process of technostress generation is shown in Figure 3.2.

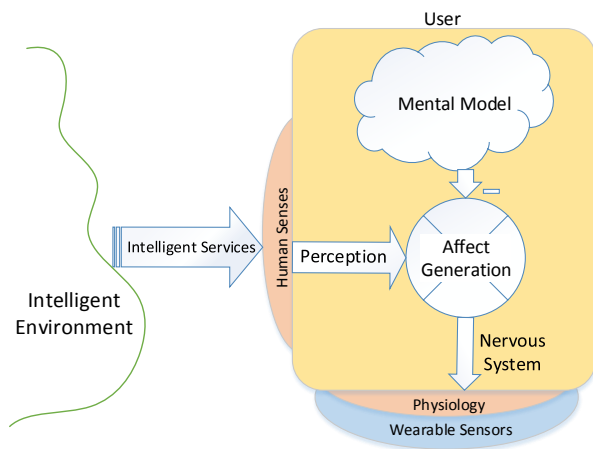


Figure 3.2: Affect generation in a user on appraisal of services received from a CAIE, helping or hindering progress vis-a-vis their goals

As already discussed, provisioning to sense explicit user feedback in pervasive computing systems is not optimal due to various reasons such as lack of a single focal point of interaction. It also defeats the purpose of an intelligent system if it has to stop and ask about the appropriateness of a service after presenting each context-aware service to the user. Inference based implicit feedback sensing is a viable solution, and *technostress* is a potent candidate phenomenon which can be

used as a means to implement the feedback loop. From system design perspective, the user’s response in the form of *technostress* is an information-rich signal, which in the absence of an

affect-sensing system, is lost. The CAFFEINE framework envisions employing human physiological data-streams to infer technostressed states, thereby creating an implicit affective feedback loop for the CAIE. Next section highlights the rationale behind using physiological sensing for our framework.

### 3.1.3.1 Rationale for Physiology based Implicit Feedback

In a pervasive interaction scenario, with interaction spanning a dynamic set of devices, the single consistent factor is the inhabitant and their physiological responses to various services they receive from the CAIE. As pointed out in Section 2.1.3, camera based facial-expression based user-feedback sensing explored by Benta et al. in [67] is not an optimal solution in a naturalistic environment, due to the need for the user to orient their faces towards the camera, in addition to other limitations of a vision based system such as variation in lighting conditions (e.g. dim lighting during night etc.). Voice recognition based user interfaces are appropriate for pervasive multi-device interaction. Speech carries rich affective information in prosody, intonation, rhythmic variation in utterances etc. Recognizing patterns in these speech features may be a viable modality for sensing a user's affective state [185, 186]. However, for sensing a user's (dis)approval of the (in)appropriateness of a service, just after they have received a service from the CAIE, the user has to consciously interact with the CAIE using their voice, introducing an aspect of *explicit* communication into the interaction. Thus, similar to facial-expression based feedback system explored by Benta et al., we argue such a system with voice-interaction based feedback would put the onus back on the user's cognitive resources.

Moreover, human affective states are comprised of a complex mixture of composite signals from various biological subsystems. As we have seen in Section 2.1.3, affective states have historically been inferred using various modalities such as audio-visual, behavioral, gestural

as well as physiological sensing [185]. Physiological sensing provides a reliable modality of non-invasively capturing responses directly from autonomic nervous system (ANS) reflecting implicit responses that may occur without a subject’s conscious awareness or are beyond cognitive intent [71]. Thus, ANS sensing provides a pathway into measuring nascent emotional states unaltered by conscious efforts of emotion regulation. Kreibig [131] presented an extensive review of relationship between human physiological responses and emotional states, in which it was pointed out that ANS responses appear more pronounced in negative emotions (such as stress, anger, frustration etc.) compared to positive emotions (such as happy, excited etc.). Thus, direct measurement of biological subsystems innervated by ANS nerves, is a potent modality for *technostress* sensing. In addition to this, on-body sensing alleviates the need for a centralized focal point of interaction, compared to a vision or audio based interfaces. Thus, we argue that on-body wearable sensing of biosignals provide the most pervasive sensing modality, well suited for our proposed scheme of technostress based CAFFEINE. In the next section, we plan to introduce the systems model of implicit feedback using physiology based technostress detection in a naturalistic CAIE.

### 3.1.4 Interaction Schematic Overview

With the ubiquity of technology and pervasiveness of IEs, users are surrounded by potential technostressors practically all the time. Following our discussion above, we propose to employ human physiology based technostress detection—a potent modality for implicit communication among agents—to design a service appropriateness feedback loop for a CAIE. Real-time detection of such technostressed states using wearable bio-sensors presents novel opportunities for inferring relevance of services provided by the CAIE, thereby providing knowledge useful in refining its services.

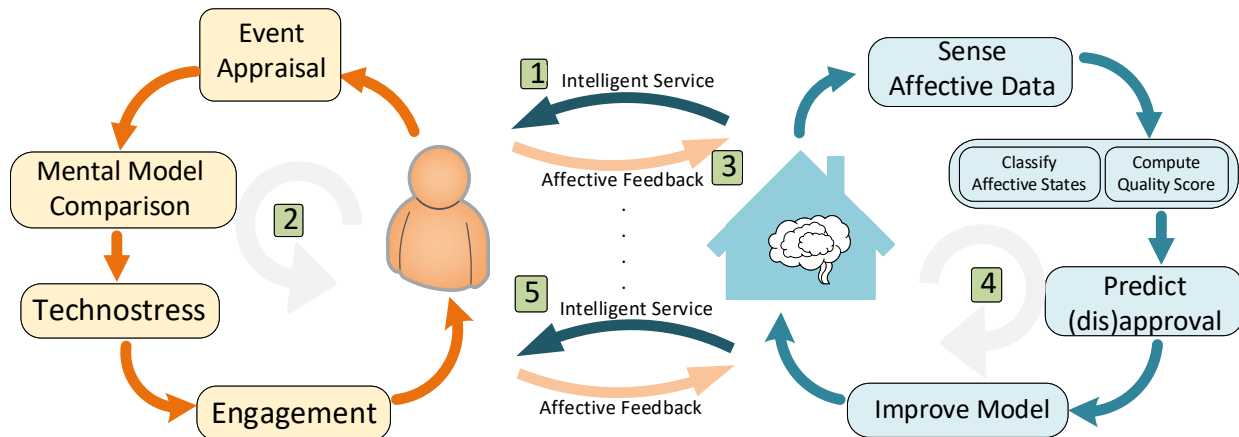


Figure 3.3: Interaction-process Model of Service Delivery and Affect Generation in CAF-EINE Framework. Modified from Saha et al. [41]

In our interaction schematic, the CAIE recognizes a user’s (dis)approval of the intelligent services by detecting their affective states, i.e. technostress in our case, resulting from the appraisal of the appropriateness of the services. On inferring disapproval from the user with ample confidence (possibly derived from the quality of the physiological signals), the CAIE responds by modifying and refining these services in the next iteration of the services. A schematic of such an interaction with an intelligent environment and the user is presented in the interaction-process model in Figure 3.3. The processes in the schematic are marked from steps one through five, indicating a chronological order:

- Step 1: A smart-service is presented from the CAIE such as a smart home or intelligent manufacturing unit,
- Step 2: Thereafter the user appraises the usefulness of the service vis-a-vis their current goals thereby producing technostress and eventually the affective feedback using behavioral traits and physiological signals,
- Step 3: The smart-system infers a user’s implicit feedback by sensing wearable physiological data and predicting user’s disapproval of the service using a user-specific *physiological*

*profile* for technostress,

**Step 4:** On inferring and assessing the implicit feedback, the system modifies the service-delivery model (to be used by subsequent services), in order to match the user's mental model of the service.

**Step 5:** The system provides the next service in the next turn by accessing the modified service-delivery model.

In this systems model of interaction in CAFFEINE, there are some assumptions about the intelligent system, such as: (i) the intelligent system has built a *physiological profile* of the users over time, which will act as a ground-truth repository of technostress physiology. This ground-truth repository will eventually assist the intelligent system to recognize and infer technostress in the wild (in Step 3). (ii) the intelligent system possesses the capability to independently assess the validity of the inference drawn on technostress using physiological signals (in Step 4). In the next few sections, we will delve deeper into each one of these assumptions and also present our methodology to address them.

## 3.2 Psychophysiological Profile Learning

Collecting ground truth data from natural settings such as an intelligent environment, poses practical limitations in realizing such systems. In addition, although very promising, it should be mentioned here that physiological sensing presents its own set of unique multiple other challenges such as (i) inevitable presence of natural confounders like food or caffeine intake and physical activity, (ii) daily and time-of-day variation of physiological subsystems, and (iii) wide between-person variations in responses to stress originating from the different coping strategies adopted by people with different personalities etc. which makes it difficult

to design a generalised classifier model for a wide range of users [76].

Plarre et al. [76] argue that knowledge of proven stressors such as mental arithmetic, public speaking etc. can be used as a more practical method to collect and annotate ground truth data. Various recent works have used validated laboratory stressors to collect ground-truth data for ideographic studies [76, 74, 187, 188, 73, 189, 77]. A quick survey of some recent works demonstrate the successful use of stressors such as mental workload (e.g. timed arithmetic [76, 74]), physical stressor (e.g. cold pressor [76, 74]), cognitive stressor (e.g. computer screen freeze, timed typing [73], paced Stroop test [187, 188], virtual-reality Stroop test [189], n-back memory test [188], Flanker test [188]), social stressor (e.g. public speaking [76, 74]) or combination of these such as MIST stressor (e.g. mental arithmetic followed by social evaluative threat [77]), as individualized physiological ground-truth learning framework.

We hypothesize that incorporating such validated stressors in the interaction scheme at optimum periodic intervals would help in first learning and then periodically re-calibrating the system's *prior knowledge* of a user's physiological profile. This, in a way, alleviates the problem arising from temporal differences in biological subsystems for each user. The periodic learning is provisioned in the schematic shown in Figure 3.6 as *Physio-Response-Learning-Loop*. For better usability of the system, we hypothesize such training (in our case *re-calibrating*) phases of the system should be incorporated into non-intrusive recreational activities which the user voluntarily participates in. A few such examples may be casual games, semi-serious games and music.

In a recent position-paper, we have highlighted the possibility of using validated musical stimuli for the physiological profile learning component of CAFFEINE framework [190]. This position-paper builds upon the work from Huang et al. [191] which shows the existence of groups of people with similar physiological responses to musical features, taken from a music-physiology database generated from the work of Bortz et al. [192].

In this dissertation, we demonstrate the use of Paced Stroop Test (PST) for collecting ground-truth data in CAFFEINE framework which is described in the following section.

### 3.2.1 Paced-Stroop Test (PST)

Stroop Color-Word Interference Test is an extensively studied paradigm in neuropsychology and a validated stress induction stimuli which has known relation to anterior cingulate cortex (ACC) [189]. ACC is known to modulate attention by regulating cognitive and emotional processing, and it has been shown to be activated during Stroop test by neuroimaging studies

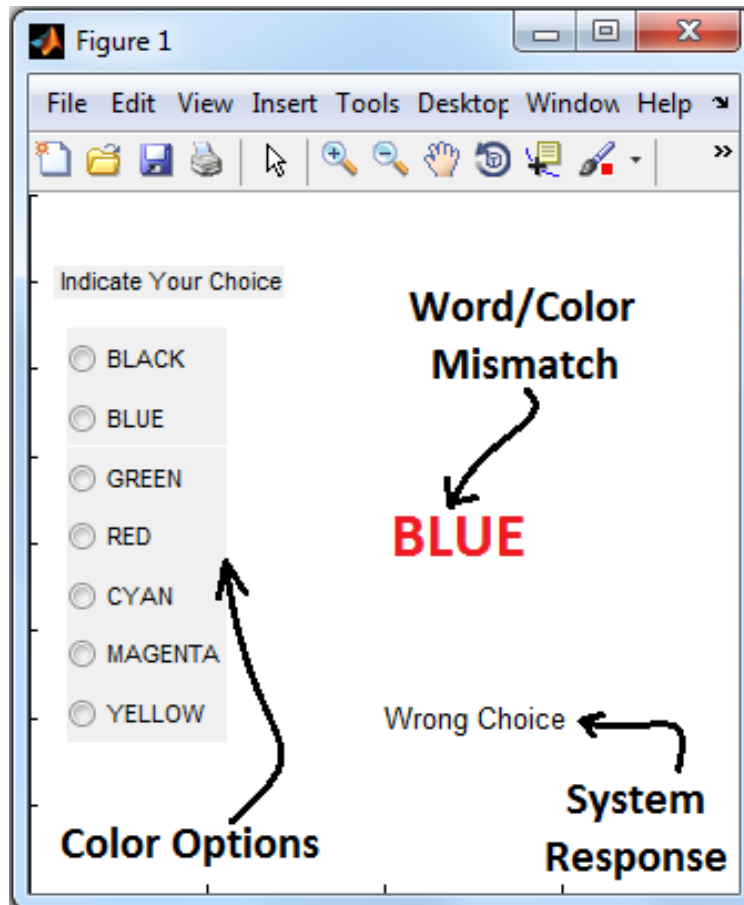


Figure 3.4: Screenshot of Incongruent Phase of Paced Stroop Test

[189]. Increased difficulty during Stroop test has also been shown to evoke autonomic nervous system changes [193] such as cardiovascular arousal [194, 195].

In its classical form, Stroop color-word interference test demands that the user chooses the font color of a word which is depicting the name of either the same color as the font's color or a different color. In the congruent phase of the test, the font color of the word and the name of the color depicted by the word match, whereas in the incongruent version, they do not. This test has been shown to produce classical behavioural effects such as *reaction-time-lengthening* [196] due to the need for suppressing involuntary visual processing of the word, and generating a new word for the color of the font. Stroop color-word interference test has been used as a standard cognitive stressor for laboratory use, which is capable of inducing heightened ANS activity on users [193, 195, 187, 194]. A modified version of this test is called Paced Stroop test, where each iteration of the Stroop test is programmed to be active for a stipulated time, say 3 seconds [187]. This task-pacing during the Stroop test has been shown to enhance the stress-inducing capability of Stroop test as compared to self-paced Stroop test, due to the need to expend increased amount of mental/cognitive effort in producing the correct response [193]. We will see later in Chapter 4, how we use PST in our CAFFEINE protocol.

In the next section, we present our novel derivation of a physiological signal based quality-metric. From the overall systems perspective, the real-time detection of the technostressed state is a valuable service-appropriateness feedback signal. However, the detection of technostress is based on inferences drawn on physiological signals using computational psychophysiology techniques. There is a need to quantitatively assess the quality of this technostress inference.



### 3.3 Physiological Signal based Quality-Metric

As discussed briefly in Section 3.1.4, the inference on technostress-based service- appropriateness from our CAFFEINE interaction framework may become actionable based on a quantifiable measure of the quality of inference. Such a quality-metric can be derived from the nature of the physiological signal captured during each session, when compared to a response to a known stimulus, say a sonic impulse (e.g. a balloon pop sound) acquired during the physiological profile-learning phase (as described in Section 3.2). In this dissertation, we have used the electrodermal activity (EDA) datastream to derive such a quality-metric. In the next section, we describe our method in detail.

#### 3.3.1 EDA Signal Analysis and Impulse Response Function

EDA signal is composed of a slow varying background *tonic* component and a peakier *phasic* component overriding the tonic. For this decomposition of EDA data, Ledalab software is a widely used software package which uses a linear-time invariant model based approach, where the measured EDA signal is assumed to be a convolution of a canonical impulse response function (IRF) with the underlying sudo-motor nerve activation (SMNA) signal (or driver) as shown in Benedek et al. [197]. This work by Benedek et al. follows from the earlier work of Alexander et. al. [198] who proposed a bi-exponential *Bateman* function based IRF with optimal values of parameters  $T_{fit}=(\tau_1, \tau_2)=(0.75s, 2s)$ , fitted from large datasets. Bateman function is a well-known pharmacokinetic function used to model the time-course of drug invasion into and elimination out of a compartment body model [199]. The Bateman function as noted in [199], with the variables  $(a, b)$  transformed with time-constants  $(\tau_1, \tau_2)$  is given by:

$$B(t) = k \frac{\tau_2}{\tau_2 - \tau_1} (\exp^{\frac{-t}{\tau_2}} - \exp^{\frac{-t}{\tau_1}}) \quad (3.1)$$

Alexander et. al. argue that so long as the  $(\tau_1, \tau_2)$  values are chosen such that individual *phasic* peaks are separated and the driver time-series falls back to a baseline in between the peaks, their exact values are inconsequential. However, Benedek et al. while proposing a continuous measure of sympathetic activation in [197], posit that the large inter-individual variations in rise- and fall-times of *phasic* response shapes can harm the accuracy of deconvolution-based EDA decomposition using fixed IRF shapes. Ledalab framework designed by Benedek et. al., addresses this issue by taking a gradient descent optimization approach to find a better fit to the measured EDA time-series, by optimizing for the IRF time-constant parameters  $T_{ledalab}$  for each individual dataset, while minimizing a cost function defined by indistinctiveness of phasic peaks and their negativity. An outcome of this approach of decomposition is that it allows for IRF shapes that may not be physiologically valid [197]. Benedek et. al. posited “artifacts in the recorded EDA” data as a possible explanation for such behavior of the IRF shapes resulting from Ledalab’s optimization approach. To quote Benedek et. al. about these IRFs—“they thus give information on the quality of the extraction algorithm and of the original SC (skin-conductance) data” [197]. Thus, taking cues from this, we can use the resulting IRF output by Ledalab for each dataset to define a measure of confidence for each user.

### 3.3.2 Towards Defining an EDA Quality Measure

In order to capture an independently obtained physiological response to an impulse stimulus, we are collecting an individual user’s physiological response to a sonic impulse stimulus,

such as a balloon-pop sound (a common method of eliciting distinct EDA responses [198]). This is described in the experimental protocol section Section 4.3.1.1. A few samples of unfiltered, zero-baselined and normalized EDA impulse responses from our dataset can be seen in Figure 4.5. In Section 3.3.1, we discussed that the *Bateman* IRF function with parameters  $T_{ledalab}$ , resultant from Ledalab optimization is an indicator of the quality of EDA decomposition and of the original EDA data. Thus, we can use an individual's response to sonic impulse to derive an EDA *quality*-metric for our dataset.

Our proposed method is noted in the following steps:

1. Find a parametric fit of the zero-baselined normalized impulse responses and extract the Bateman function time-constants  $T_{sonic}=(\tau'_1, \tau'_2)$ .
2. Define an *error* term  $E_{irf}$  derived from the mean of the percentage of absolute deviations of  $T_{sonic}$  from  $T_{ledalab}$
3. Define a *quality* term  $Q_{eda} = (1 - E_{irf})*100$ .

For Step-1, we take 5 seconds of zero-baselined normalized impulse response data starting from the response onset as the independent variable, Equation (3.1) as the target model function with  $T_{ledalab}$  as the initial values for the parameter estimation using least-squares curve fitting method. We obtain  $T_{sonic}=(\tau'_1, \tau'_2)$  as a result of this curve-fitting. For Step-2 and Step-3, we compute the  $E_{irf}$  and  $Q_{eda}$  as described above.

To interpret our quality metric, a higher  $Q_{eda}$  score implies that the experimental IRF closely matches the Ledalab optimization based IRF, signifying good quality EDA signal with negligible artifacts corrupting it, thus enabling good inferences on sympathetic activation. On the other hand, low  $Q_{eda}$  score may imply inadequacy of the decomposition method, or higher artifacts in the datastream feeding into erroneous inferences on sympathetic activation.

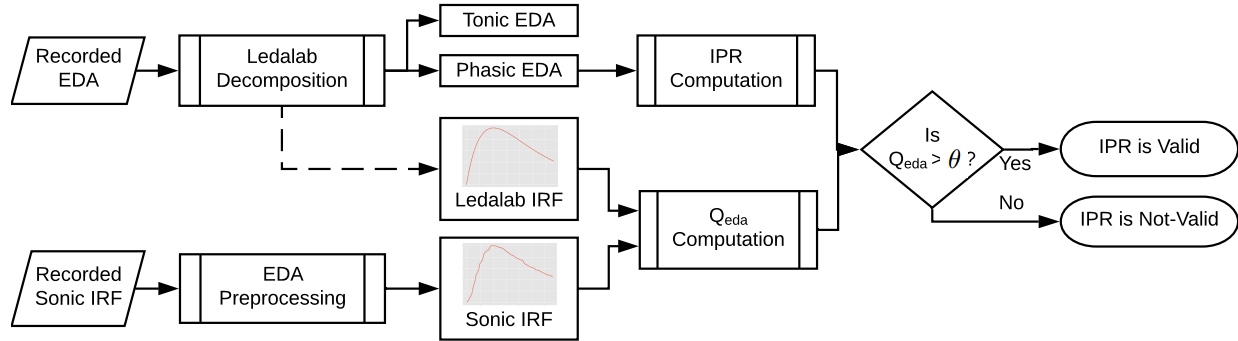


Figure 3.5: Schematic of  $Q_{eda}$  Computation Process and assessment of Validity of Integrated Phasic Response (IPR) feature based on  $Q_{eda}$  Score.

To operationalize the  $Q_{eda}$  score, we may define a threshold  $\theta$ , say 50%, acting as a surrogate for a confidence measure on the inferences drawn using EDA features. The inference system in a CAIE will operate on the following condition:

$$\text{EDA Features} = \begin{cases} \text{Valid,} & \text{if } Q_{eda} > \theta \\ \text{Invalid,} & \text{if } Q_{eda} < \theta \end{cases} \quad (3.2)$$

The schematic of this overall process of  $Q_{eda}$  score computation and interpretation is shown in Figure 3.5.

### 3.4 CAFFEINE Framework: Bringing It All Together

Till now, we have presented various components of our CAFFEINE framework, the interaction overview with a CAIE, discussed the assumptions and our approach towards addressing them. To summarize our discussion in a coherent interaction framework, we present our final interaction schematic, as depicted in Figure 3.6. The chronological order of the processes as

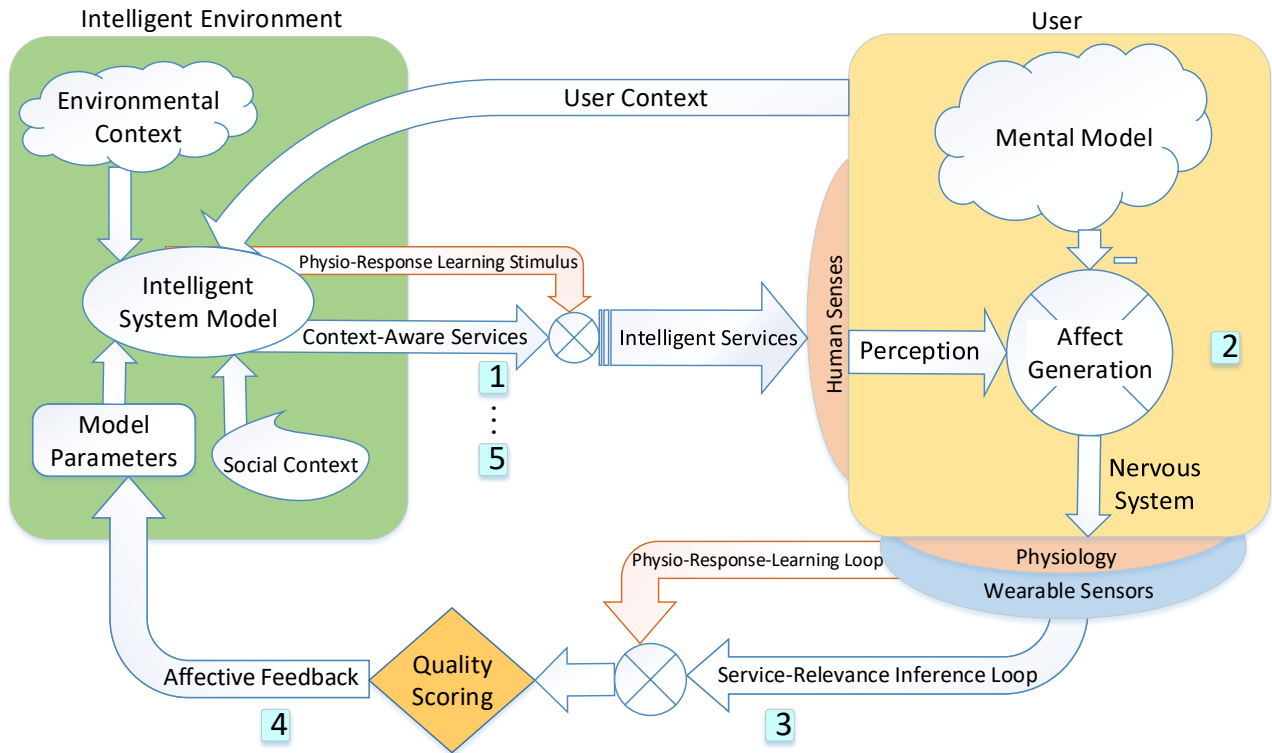


Figure 3.6: Affective Feedback in an Intelligent Environment

described in Section 3.1.4 are marked with the same step numbers in both Figure 3.3 and Figure 3.6.

As discussed in Section 3.1.3, this scheme operates by inferring a user’s context from various environmental sensors and provides a smart service, depicted using the primary blue arrows (Step 1) in Figure 3.6. The user appraises the appropriateness of this service against their mental model of the service and generates affective reaction (Step 2), which is sensed using the physiological channel. The intelligent system, which has over-time built an individualized physiological profile acting as the ground-truth about the user, can now infer technostress arising due to the smart service (Step 3). The *physiological profile learning loop* is shown in Figure 3.6 using the orange secondary arrows. Using the  $Q_{eda}$  based quality metric scoring method discussed in Section 3.3, the intelligent system can independently assess the validity

of the technostress inference, and accordingly update it's own service-delivery model for subsequent use (Step 4). In the next iteration of the service delivery, this updated model will be used.

### 3.5 Conclusion

In this chapter, we presented the envisioned interaction scheme in our CAFFEINE framework. To summarize, the CAIE system is a service provider that has *prior knowledge* of a user's affective and physiological profile, provides smart services by recognizing the user's present context consisting of their physical, environmental, affective, and social situations. This schema for the IE is illustrated in Figure 3.6, which shows the implicit feedback loop implemented for an IE using physiology-based *technostress* recognition framework for determining service-relevance. On successful identification of a user's context, the IE presents either a proactive or a reactive service to assist them towards their goals. The perception of (in)appropriateness of this service, when appraised by the user, causes technostress (or lack thereof), thereby signifying (dis)approval for the service. Successful detection of such a technostressed state is a valuable feedback signal, which may be used to reconfigure the services. To assess the validity of this inference, thereby enabling the CAIE to quantify the need to reconfigure it's service, we presented a physiological signal based quality-metric. In later chapters, we have presented validation of our interaction framework, method of ground-truth collection as well as the quality-metric.

# Chapter 4

## Experimental Setup and Results

---

*“How often people speak of art and science as though they were two entirely different things, with no interconnection. An artist is emotional, they think, and uses only his intuition; he sees all at once and has no need of reason. A scientist is cold, they think, and uses only his reason; he argues carefully step by step, and needs no imagination. That is all wrong. The true artist is quite rational as well as imaginative and knows what he is doing; if he does not, his art suffers. The true scientist is quite imaginative as well as rational, and sometimes leaps to solutions where reason can follow only slowly; if he does not, his science suffers.”*

— Isaac Asimov in *The Roving Mind* [200]

---

In this document, till now we have presented our CAFFEINE framework proposed in this dissertation work. In the framework, we have proposed a novel interaction scheme for inferring a user’s disapproval for a context-aware service presented by a CAIE, by detecting their psychological state of technostress induced by the service. We have discussed our rationale for using physiological signals as a potent information stream that alleviates the need for a continuous medium of interaction, envisioned to solve a unique Human-Computer Interaction (HCI) issue posed by the IE domain. We have also presented a novel method of deriving a quality-metric from the physiological signals, for estimating the plausibility of our inference on user’s disapproval of a service. Towards that end, we have conducted a set of experiments to first validate and then delve deeper into the physiological underpinnings of our framework, revealing important discriminatory features as well as a novel method for deriving a physiological-signal based inference quality metric.

In this current chapter, we present the experimental validation of the CAFFEINE framework from three of our experiments and their associated results. In experiment one as described in Section 4.1, we have prototyped a warehouse scenario wherein the picker-personnel are given a context-aware service to help them complete their task in a more efficient and error-free manner. We have developed a machine-learning system to recognize technostressed states in response to *wrong* responses from the system using data collected from a user’s physiological signals. We demonstrated that recognition of technostressed states was indeed possible, and the system performance improved when we included data from PST in our training set. These results, originally presented in [40], provide an experimental validation of our interaction framework.

Following the experimental validation of a critical link in our framework—namely, the per-



sonalized ground-truth learning component—we set out to validate the physiological underpinnings of the psychological state of technostress. In particular, our quest was to identify physiologically validated features of psychological stress (such as indicators of increased sympathetic activation), also arising during situations of technostress in a CAIE. Results from this study would validate our hypothesis of using technostress as a service-appropriateness signal in the wild, as well as help us in identifying individual discriminatory features. In order to design a more realistic and immersive experimental setup, we have prototyped a general intelligent supermarket scenario in a position tracked virtual reality space at our research institute, the Institute for Creativity, Arts and Technology, which is described in more details in Section 4.2. From the results of this experiment, we identified a few physiologically validated features from the EDA datastream that contain discriminatory information which may potentially improve the technostress classifier performance described in Section 4.1.2 [41].

On validating both the components of our CAFFEINE framework, namely, the ground-truth learning component and the use of technostress as a service-appropriateness signal in the wild, we needed a physiologically grounded method of quantifying the quality of a technostress inference. This quality score is envisioned to dynamically assess the correctness of the technostress recognition, thereby helping the CAIE to decide if it needs to reconfigure its service or discard the inference altogether. In order to design an experimental method to assess this, we designed an experimental setup by modifying the previous experiment, as described in Section 4.3. This modification is used to ascertain the quality of inferences from our framework using the underlying phenomenon of a canonical impulse-response function (IRF), which is integral to EDA signal analysis, specifically a widely used tool called Ledalab. Results validate our idea of a physiological-signal based quality metric computation, considering an empirical threshold [42]. We now present the three experiments in the next three

sections.

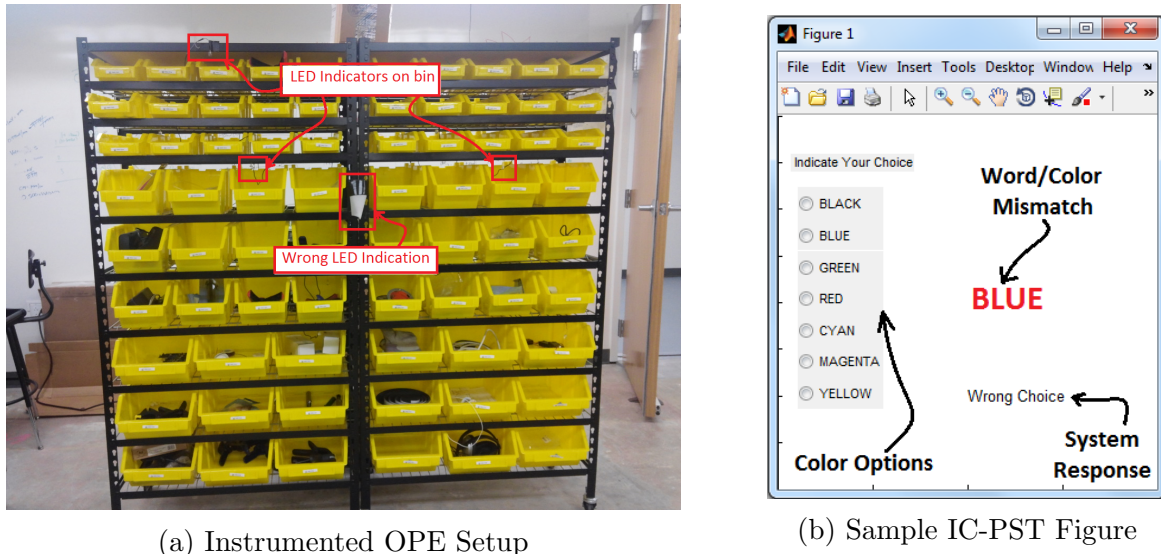
## 4.1 Experiment One: Warehouse Picker Experiment

An order picking task is the process of collecting supply items corresponding to a particular order from warehouse racks, and sorting them as per order requisitions for delivery. Order picking is one of the major tasks in warehouses across the globe, accounting for upto 60% of their operating costs [201]. To reduce this operational cost as well as human errors, order-picking personnel are provided with various context aware task-assistances from the IE. Commonly used task-assistance systems in the industry are paper pick-lists (pick-by-list or LST), illuminated bin indicators (pick-by-light or LHT) and heads-up-display (HUD) assisted picking [201]. Such systems are usually fitted with laser trackers in order to determine the current pick by detecting personnel’s reach in each bin. However, these laser-trackers are frequently prone to mistrigger errors which is a major source of irritation for these personnel [201], inducing “achievement stress” which is one of the potent reasons of *technostress* [35]. In practice, these mistrigger errors have to be corrected by explicitly indicating to the IE about the error using buttons.

Our prototype (shown in Figure 4.1a) is based on the pick-by-light scheme, wherein the bins are fitted with bin-indicating LEDs (bLED) showing the personnel which bin to pick from, as well as wrong-pick LEDs (wLED) indicating a wrong pick.

### 4.1.1 Experimental Protocol

We have collected data from 7 participants (5 males, 2 females) in the age bracket 20-30 yrs under a research protocol approved by *Virginia Tech* IRB#14-689. Participants represented



(a) Instrumented OPE Setup

(b) Sample IC-PST Figure

Figure 4.1: Snapshot for Order Picking Experiment Setup in our Laboratory at Institute for Creativity, Arts and Technology.

a wide range of nationality, ethnicity and physique, though no conscious effort was made to select participants based along any discriminatory attributes. Each participant had to go through the two phases of the experiment as mentioned below.

#### 4.1.1.1 Order Picking Experiment (OPE) Phase

During our experiments, we purposefully indicate a wrong pick at predetermined times, even though the participants know they are picking from the correct bin, thereby simulating the situation of mistrigger error in this IE which causes *technostress*. Our system is operated in a Wizard-of-Oz fashion, wherein the services and the mistrigger/pick-place error indications are both triggered by the experimenter. In our experiment, each user is provided with a paper-pick list containing 14 order bin numbers (i.e., 14 items per task), out of which 5 orders have no task-assistance (LST), 5 orders have correct task-assistance ( $\{LHT + bLED\}$  states) and 4 orders have incorrect task-assistance ( $\{LHT + wLED\}$  states). The participants were requested to finish their tasks in the minimum possible time.

#### 4.1.1.2 Paced-Stroop Test (PST) Phase

As described in Section 3.2, there is a need to address the individual variations in physiological signals by designing a provision to collect validated ground-truth dataset for each user. Our CAFFEINE framework has designed a provision for collecting such ground-truth using a validated laboratory experiment called Paced-Stroop Test, as described in Section 3.2.1. There are many possibilities of using various other validated stress induction instruments such as n-back memory test, Flanker test or MIST, as described briefly in Section 3.2. For the purpose of this experiment, we have used task pacing time of 3 seconds between each Stroop figure, running for a total of 180 seconds. In the PST experiment, one block of 60 seconds (i.e., 20 pairs) of C-PST is preceded and followed by 60 seconds each of IC-PST (i.e., 2 x 20 pairs). A snapshot of the IC-PST phase of the experiment is shown in Figure 4.1b.

#### 4.1.1.3 Picker Experiment Hardware Setup

Our setup involves acquiring EDA and ECG data using BioEmo and BioBeat sensors from Biocontrol Systems<sup>1</sup> respectively. BioEmo is an exosomatic skin conductance sensor, designed to be worn on the medial or distal phalanges of the fingers in direct contact with skin. BioBeat is the ECG sensor which comprises of gold plated electrodes worn in the form of a chest band. Both these sensors are connected with iCubeX wiMicroDig digitizers sampling at 200Hz, which are configured to stream data wirelessly over bluetooth to a nearby laptop acting as a terminal for running both OPE and PST while also logging physiological data.

<sup>1</sup>[www.biocontrol.com](http://www.biocontrol.com), accessed 01/11/18

## 4.1.2 Picker Experiment Data Analysis

### 4.1.2.1 Data Normalization and Preprocessing

As discussed earlier in Section 2.4.1, significant individual differences are observed in the baseline value for skin-conductance levels. The raw time-series data is normalised by computing the studentized residuals, making the algorithm self-calibrating to personal baseline differences, which improves classification [76, 74].

For EDA preprocessing we have used a modified version of Jaimovich's EDA preprocessing MATLAB subroutines [152]. The algorithms take the raw EDA time-series, resamples it to 50Hz, removes electrical noise using an FIR filter of 0.5Hz cut off frequency and gives time annotated SCR and SCL values as output. For computing the ER.SCR related features, we have used the procedure charted in Kim et al.'s work [202, 130]. For ECG preprocessing, we have used the Jaimovich's ECG preprocessing MATLAB subroutines [152], which gives a time annotated instantaneous HR time series as output. The raw ECG data, obtained from the ECG sensor setup as described in Section 4.1.1.3, is detrended and filtered using an FIR high-pass filter with Kaiser Window having a cut-off frequency of 3Hz, followed by heart-rate extraction at each beat using a moving window with a thresholding parameter of 2 standard-deviations (SD) and beat change-ratio of 20%.

### 4.1.2.2 Feature Extraction and Feature Reduction

We extract a set of fourteen features from GSR and ECG time-series data, that have been reported in literature as distinguishing for stress related studies [187, 76, 74]. Following Figner's report [143] stating a common window of interest for EDA feature extraction to be limited upto 6 seconds after the stimulus onset, features in our analysis are extracted from

a window of 6 seconds called *StimWin*. Only time domain EDA features were included in our analysis which include mean amplitude, rise-time and fall-time of the phasic ER-SCR. Both time and frequency domain features are extracted from the HRV time-series, derived from the ECG data as described in Section 2.4.2.2. Time domain features include mean and SD of HR computed at each beat, mean and SD of R-R peak intervals, root-mean-square of successive difference of R-R peak intervals, percentage of all R-R peak intervals in the *StimWin* that are greater than 20ms and 50ms. Frequency domain features used are total spectral power in LF band, HF band, and the ratio of these total powers in LF and HF bands.

In order to properly project observations onto a space with independent basis vectors, we use principal component analysis (PCA) which while orthogonalizing features also preserves the variance of the dataset along these orthogonal bases. Thus, it can also be used to discard those dimensions which do not explain significant amount of variance of the dataset, essentially reducing dimensions and alleviating the *Curse-of-Dimensionality*. A reasonable cut-off for variance is 99%, i.e. accepting at least N dimensions such that they explain a total variance of more than 99% [165].

#### 4.1.2.3 Support Vector Machines (SVM)

Support Vector Machine is a very widely used linear discriminative classification algorithm which, being data distribution independent, is known to successfully classify a wide variety of problems with good accuracy. Prior works on physiology based stress recognition have shown SVM outperforms various other classifiers [75, 77, 76, 74]. SVM is predominantly a binary classifier where the primary objective is to come up with a maximum margin classifying hyperplane between the classes such that there are minimum number of support vectors inside the margin, where margin is defined as the minimum distance between the

classifying hyperplane and a point in the dataset. So for a training dataset of labeled points  $\mathbf{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^n$  with  $y_i \in \{+1, -1\}$ , soft-margin SVM has the following dual formulation :

$$\begin{aligned} \text{Objective :} \quad & \max_{\alpha} L_{dual} = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{K}(x_i, x_j) \\ \text{Constraints :} \quad & 0 \leq \alpha_i \leq C \quad \forall i \in \mathbf{D} \quad \text{and} \quad \sum_{i=1}^n \alpha_i y_i = 0 \end{aligned}$$

where  $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)$  is the kernel function used to map data vectors to a more expressive feature space which aids in classification of non-linear datasets. In Section 4.1.3.1, we present our kernel function comparison study to choose a good kernel for the CAFFEINE framework dataset.

### 4.1.3 Picker Experiment Results

The goal of this current work is to correctly identify physiological states corresponding to the onset of mental stress (i.e., *technostress*) induced by the incorrect responses from the IE, i.e.,  $\{LHT + wLED\}$  states in OPE experiment. This goal is achieved by learning a statistical model from a subset of this labelled dataset as well as the ground-truth obtained from a laboratory stressor i.e., the PST dataset. The model is verified per user by predicting the class i.e., stressed (S) vs. not-stressed (NS), of a previously unseen input sample from the OPE dataset using the leave-one-sample-out-cross-validation (LOSO-CV) method. Following data pre-processing, fourteen features were extracted from segments of window length *StimWin* from the onset of each stimulus. These features were presented to the pattern recognition pipeline to learn a statistical model.

We must note here that for this system must not tolerate *any* false-positives (FP) even at the

cost of a low recognition accuracy of *technostressed* states. This is intuitive that an FP (i.e., when the system falsely senses user to be in *technostressed* state when they are actually not) in this system will directly deteriorate its performance, as it will try to reorient its services even though the service was actually helpful to the user.

As described in Section 3.2.1, physiological data collected during PST is used as ground truth data corresponding to cognitive states of S or NS, C-PST being related to NS state and IC-PST to S state of the user. Based on our experiment design, we have the following sets of results to present here: *a) CASE-I: Train* only on OPE data and *cross-validate (CV)* on OPE data, *b) CASE-II: Train* only on PST data and *Predict* on OPE data, and *c) CASE-III: Train* on Combined PST+OPE data and *CV* on OPE data. Evidently we are particularly interested in the results from the prediction/CV performed on the OPE data.

#### 4.1.3.1 SVM Kernel Selection

In order to experimentally validate a good kernel for our technostress based CAFFEINE framework, we tested 3 kernel functions, namely Gaussian, Polynomial and Sigmoid kernels on a small pilot dataset. The confusion matrix comparison is shown below in Table 4.1. Specifically, in comparison studies 1, 2 and 3, Sigmoid kernels outperform Polynomial kernels, both on lower number of False Positives and overall accuracy. In comparison study 4, Sigmoid kernel outperforms Gaussian kernel on classification accuracy metric. In comparison study 5 and 6, they perform equally. Given these findings, Sigmoid kernel clearly is the best performing kernel for our CAFFEINE framework. Zhai et al. [203] showed that Sigmoid kernel outperformed various other kernels for stress recognition. This is in line with our findings, which prompted us to use Sigmoid kernel for compiling our results.



Table 4.1: Comparison Table for Polynomial, Gaussian and Sigmoid Kernel Functions. In each confusion matrix, S/NS pair represents stressed/non-stressed states. A typical confusion matrix result, for say, Comparison Study 2, Polynomial Kernel will read as TP=3, TN=8, FP=2, FN=1.

Comparison Study 1						Comparison Study 2						Comparison Study 3					
Sigmoid			Polynomial			Sigmoid			Polynomial			Sigmoid			Polynomial		
	NS	S		NS	S		NS	S		NS	S		NS	S		NS	S
NS	10	<b>0</b>	NS	4	<b>6</b>	NS	10	<b>0</b>	NS	8	<b>2</b>	NS	10	<b>0</b>	NS	9	<b>1</b>
S	0	4	S	2	2	S	2	2	S	1	3	S	3	1	S	1	3
Comparison Study 4						Comparison Study 5						Comparison Study 6					
Sigmoid			Gaussian			Sigmoid			Polynomial			Sigmoid			Polynomial		
	NS	S		NS	S		NS	S		NS	S		NS	S		NS	S
NS	8	<b>2</b>	NS	8	<b>2</b>	NS	8	<b>2</b>	NS	8	<b>2</b>	NS	9	<b>1</b>	NS	9	<b>1</b>
S	1	3	S	2	2	S	1	3	S	1	3	S	3	1	S	3	1

#### 4.1.3.2 Performance Evaluation Criteria

Classification accuracy metric is not an adequate metric for evaluating classifier performance of class-imbalance learning problems such as ours [204, 205]. Stressed states in practical situations can reasonably be assumed to be rare states, compared to normal non-stressed states; hence, a stress-classifier in practice has to, almost always, deal with imbalanced classes. A confusion matrix and its derived measures such as precision ( $p$ ), recall ( $r$ ),  $G$ -score,  $F_\beta$ -score are used for quantifying classifier performance in imbalanced cases and are defined as:  $G = \sqrt{pr}$ ,  $F_\beta = \frac{(\beta^2+1)pr}{\beta^2p+r}$  where  $\beta$  is used to tune the effect of  $p$  and  $r$ . Sasaki [206] suggests that for  $\beta < 1$ ,  $F_\beta$  becomes increasingly precision oriented. As discussed in Section 4.1.3, our problem statement calls for heavy penalty for *any* FP while also rewarding a good classification for *technostressed* states; so we must formulate an  $F_\beta$ -score that rewards very low FP (i.e., more dominant on  $p$ ); hence we select  $\beta = 0.1$ .

Table 4.2: User-wise Confusion Matrix,  $G$ -score and  $F_{\beta}$ -score Calculations (described in Section 4.1.3.2) for Case I (discussed in Section 4.1.3). A typical confusion matrix result for say User F, Case-I will be read as TP=3, TN=7, FP=3, FN=1.

Confusion Matrix for Case-I					
User		NS	S	$G$ -score	$F_{\beta}$ -score
A	NS	9	1	0.35	0.49
	S	3	1		
B	NS	6	4	0.41	0.33
	S	2	2		
C	NS	7	3	0	0
	S	4	0		
D	NS	8	2	0.29	0.33
	S	3	1		
E	NS	9	1	0	0
	S	4	0		
F	NS	7	3	0.61	0.50
	S	1	3		
G*	NS	6	4	0.26	0.20
	S	2	1		

Table 4.3: User-wise Confusion Matrix,  $G$ -score and  $F_{\beta}$ -score Calculations (described in Section 4.1.3.2) for Case II (discussed in Section 4.1.3). A typical confusion matrix result for say User F, Case-II will be read as TP=1, TN=9, FP=1, FN=3.

Confusion Matrix for Case-II					
User		NS	S	$G$ -score	$F_{\beta}$ -score
A	NS	10	0	0.50	0.97
	S	3	1		
B	NS	5	5	0.53	0.38
	S	1	3		
C	NS	6	4	0.22	0.20
	S	3	1		
D	NS	10	0	1.00	1.00
	S	0	4		
E	NS	3	7	0.47	0.30
	S	1	3		
F	NS	9	1	0.35	0.49
	S	3	1		
G*	NS	6	4	0.65	0.43
	S	0	3		

Table 4.4: User-wise Confusion Matrix,  $G$ -score and  $F_{\beta}$ -score Calculations (described in Section 4.1.3.2) for Case III (discussed in Section 4.1.3). A typical confusion matrix result for say User F, Case-III will be read as TP=3, TN=9, FP=1, FN=1.

Confusion Matrix for Case-III					
User		NS	S	$G$ -score	$F_{\beta}$ -score
A	NS	10	0	0.71	0.99
	S	2	2		
B	NS	8	2	0.67	0.61
	S	1	3		
C	NS	9	1	0.35	0.50
	S	3	1		
D	NS	10	0	0.71	0.99
	S	2	2		
E	NS	8	2	0.67	0.61
	S	1	3		
F	NS	9	1	0.75	0.75
	S	1	3		
G*	NS	9	1	0.67	0.67
	S	1	2		

#### 4.1.3.3 Model Performance Evaluation

The results from training an SVM classifier with Sigmoid Kernel for each user are presented in Table 4.6. Evidently the model underperforms on both  $G$ -score and  $F_{\beta}$ -score metrics in CASE-I where the classifier trains only on the OPE data, and could not recognise a single correct *technostressed*-state for User C and User E. In CASE-II, we trained on PST and cross-validated on OPE data. The results clearly show that the classifier performance has improved, benefiting from the improved ground truth data provided by the PST dataset. In CASE-III, we used the combined PST and OPE datasets to train our classifier. The performance results have improved both on  $G$ -score and  $F_{\beta}$ -score metrics, compared to both the previous cases. Particularly, the FP has reduced and classification for *technostressed*-states has increased for all users. Thus, it is safe to conclude that the model has improved from the improved ground truth data provided by the combined PST and OPE datasets.

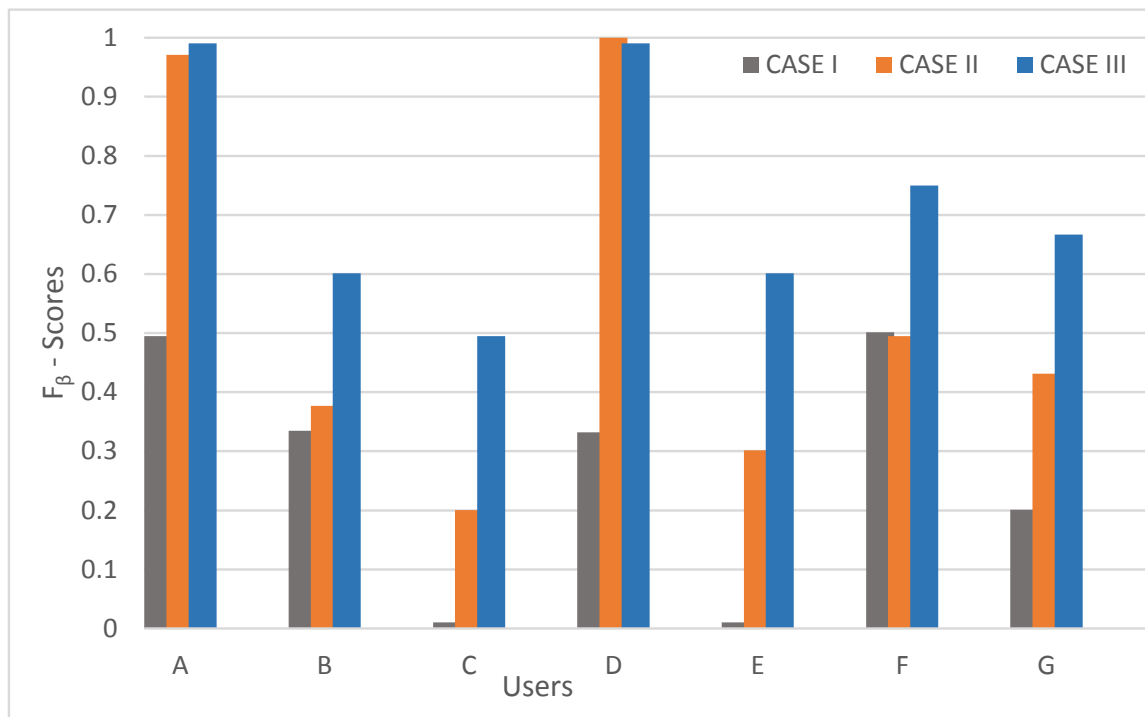


Figure 4.2: User-wise  $F_{\beta}$ -score Comparison For Case I, II and III

#### 4.1.4 Picker Experiment Discussion

Using this user-study, we sought to answer two basic questions for real-life IE, namely whether it is possible: *a)* to create an *implicit* channel of communication between a user and CAIE by recognizing technostressed states, *b)* to use laboratory stressors as ground truth for real-life stressors during ambulatory sensing of stress. The results produced from this user study as shown in Section 4.1.3, depict that by training the SVM models individually for each user, we were able to find similarity in the patterns of physiological data acquired during sessions where two kinds of stressors were presented to a user, namely *technostress* in the OPE experiment and *cognitive stress* in the PST experiment. This is evident from the improvement in results in *Case-III* for all users (except User-D) as compared to *Case-II* and *Case-I* of the experiment. This is depicted in Figure 4.2. These results provide preliminary

evidence of computationally learning statistical parameters corresponding to stress related physiological responses elicited from a proven laboratory stressor and using these parameters to classify stress responses in real-life settings (*technostress* in this case). Section 2.3 describes how *technostress* is elicited when a system malfunctions, thereby hindering a user's progress. Thus, the parametric model capable of classifying *technostress*-states, can be used to provide a user feedback, employing the *implicit*-channel of communication [9], thereby completing an affective feedback loop in an intelligent environment.

As described in Section 4.1.3, our goal was to train a classifier that produces the least number of false positives (FP) while also accurately classifying stressed states. Results shown in CASEE-III of Table 4.6 are very encouraging, demonstrating a consistent increase in  $F_\beta$ -score with the introduction of PST dataset in training phase. It should be noted here that by tuning the hyperparameters for the SVM model, we were able to reduce FP count to zero for all users; however, it resulted in near-zero correct classifications for stressed states, which is why they are not included in the results. The OPE experiment was conceptualised to mimic a real-life setting for an IE which senses user context and provides relevant services. This also introduces a lot of noise sources, primarily motion artifacts into the sensor data. Although, we have used adhesive tapes to affix the EDA sensors, thus, reducing sensor fitting issues. However, the sensors used for this experiment are not designed for use in ambulatory settings. There were instances of data corruption, which were dropped during the pre-processing stages.

#### 4.1.5 Picker Experiment Conclusion

In this section, we set out to explore the possibility of determining the relevance of services provided by an intelligent environment by creating an affective feedback loop. Our results

show that we could indeed identify technostressed states in response to wrong responses from the system. Following a few recent works, we also hypothesized that proven laboratory stressors such as PST can be used as effective ground truth collection instruments even in ambulatory settings. Results from this study are encouraging and show that this idea of using Paced-Stroop Test for a personalized ground-truth learning framework helps improve the recognition of technostressed states. Thus, these results present an experimental validation of the ground-truth learning component of our CAFFEINE framework.

Having validated the ground-truth learning framework component, our next quest is to prove the validity of using technostress as a feedback signal, in an immersive intelligent environment providing context-aware smart services. In the next section, we present our modeling of the immersive intelligent environment in the form of an Intelligent Supermarket in a position tracked virtual reality space providing navigation assistance. We used this setup to detect physiological signature of technostress from physiological signals.

## 4.2 Experiment Two: Virtual SuperMarket with Navigation Assist Experiment

In a descriptive essay by Wahlster et al. [207], supermarkets are envisioned as a good candidate to model as a CAIE, wherein various intelligent interactions are discussed such as RFID tagged objects and web-connected shopping lists. In our setup, we modeled such a service which we call *navigation-assist* in a supermarket, that finds the best direct path to the items on a dynamically changing shopping list from the current location of the user. There has been recent commercial interest in smart supermarkets such as an in-store navigation-assist introduced by Lowe's Supermarkets [208], and a smart shopping experience requiring

no checkout, independently designed by Amazon Inc. and Stanford Cognition Labs [209].

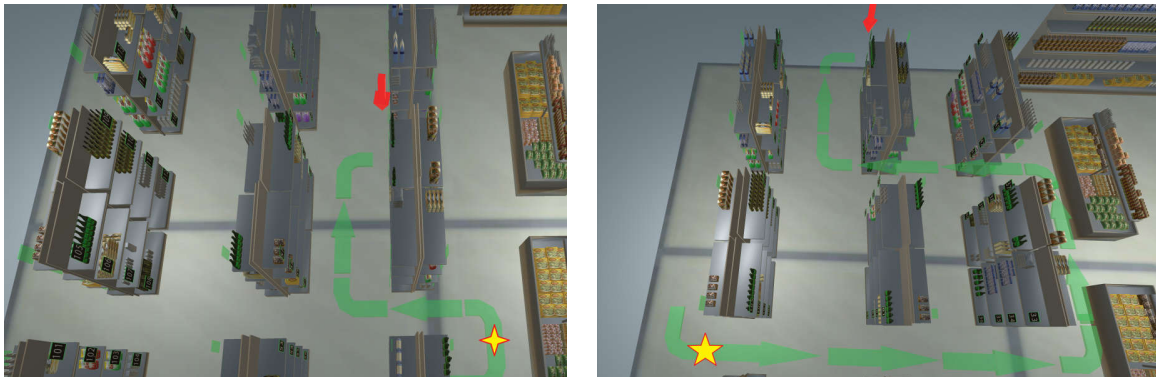
### 4.2.1 Intelligent Supermarket in Virtual Reality Setup

Our prototype intelligent supermarket system simulates a navigation-assist system which is designed to show the best direct-path from user's current location to the destination obtained from a grocery list. However, due to the dynamic nature of the list or real-life sensing issues such as uncertainty in indoor location tracking, the system may not always come up with an optimum path. Since the navigation-assist service is intended to help the customer achieve their goals faster, the wrong services (i.e. winding path) may cause achievement stress which is a potent cause of technostress.

In order to design an immersive experimental setup, we have modelled our supermarket experiment in a fully position-tracked virtual reality (VR) setup in The Cube at Virginia Tech. Although this is a controlled environment, aspects of human behavior have been widely compared in virtual and real worlds, and have been found to be following similar patterns [210]. For example, it has been shown that users follow similar social norms [211], demonstrate similar perception of proxemics [212], and follow similar economic behavior [213] in virtual worlds as they do in real world. These and various other studies have shown the effectiveness of using virtual worlds as viable option for designing our intelligent environment setup. In the next section, we present the details of our experimental setup.

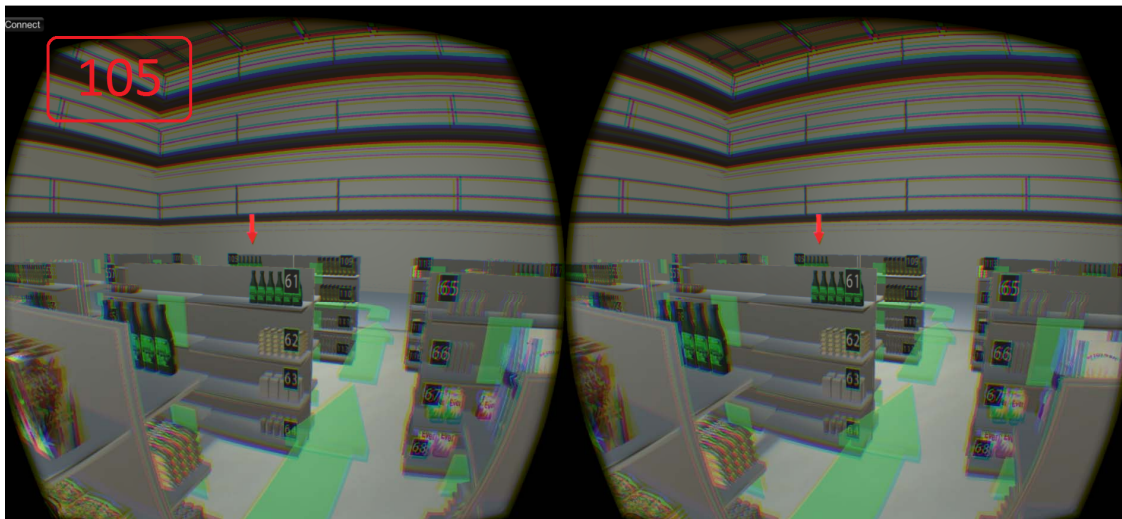
#### 4.2.1.1 Experimental Protocol in VR Setup

Participants wear an Oculus headset and walk in a position tracked space, which simulates walking in a supermarket. The hardware setup has been described in Section 4.2.2. In our model, items are placed on shelves marked with serial numbers. Participants were informed



(a) Top View of A Path

(b) Top View of A Path



(c) Snapshot of a Virtual-Reality View

Figure 4.3: Our Experimental Setup: “Smart Supermarket with Navigation-Assist Service”. Reproduced from Saha et al. [41]

that their shopping list was pre-populated with 10 items, and item numbers corresponding to the next item will be shown as an overlay on the supermarket scene in their VR-glasses (shown in Figure 4.3c). This unseen item-list gives a perception of a dynamically changing shopping list. Participants were informed that the system will highlight a direct path from current position to destination using horizontal green path-arrows, while a vertical red-arrow will indicate the final destination visible from their current location (shown in Figure 4.3c). This red-arrow is essential for the user to create a mental model of the smart-service (i.e. a



direct path), violating which may impart technostress, as already discussed in Chapter 3.

In reality, the experiment was conducted in a Wizard-of-Oz fashion, where-in the experimenter would listen to the participant speaking the next item number on their VR headset and activate the next path arrow by pressing a *hotkey* in the experimenter's view (see Section 4.2.2). Some of the paths were an obvious direct path, while some paths were deliberately made winding, to create an impression of a system failure thus imparting technostress. Participants were asked to always follow the path indicated by green arrows, even if it is not a direct path. Out of the 10 items on the list, we provided correct service (CS: direct path) and wrong service (WS: winding path) for 5 items each.

## 4.2.2 Hardware Setup

Our setup consists of a Qualisys Motion Capture system with 24 Oqus5+ cameras for tracking reflective marker based rigid-bodies and an Oculus Rift DK2 as our VR headset. The Oculus Rift DK2 is connected to a laptop (say, Oculus Computer (OC)) which is running the supermarket model in Unity, and is being carried by the participant in a backpack. For performing the experiment in Wizard-of-Oz fashion, the experimenter has a mirrored view of the participant's VR view (running on OC), onto a local computer, say Experimenter Computer (EC) communicating over the local wireless network. Physiological data is collected using Empatica E4 wristband which streams time-stamped biosignal datastreams over Bluetooth Low Energy (BLE) to OC. The *hotkey* presses (VR event onsets) are time-stamped with the OC machine-time along with Empatica E4 data using custom code<sup>2</sup>.

<sup>2</sup>(available at <https://github.com/debapratimsaha/EmpaticaUnityBLEClient>)

## 4.2.3 Virtual Supermarket Experiment Data Analysis

### 4.2.3.1 Electrodermal Activity (EDA) Analysis

As already described in details in Section 2.4.1, EDA is a reliable indicator of activation of the sympathetic division of ANS (SNS), which shows heightened activity during the experience of technostress [131, 96]. It is arguably the only physiological system that is activated solely by SNS, uncontaminated by the PSNS, making it a well established marker for SNS activity [40]. EDA signal is composed of a slow varying *tonic* and a rapidly changing *phasic* components. For this experiment dataset, we decompose measured EDA into phasic and tonic, using a deconvolution based method (see Ledalab [197]), wherein the measured EDA is deconvolved with an impulse response function (IRF) waveform to obtain the underlying compact sudomotor nerve-activity (SMNA) pulses. The IRF is modeled as a biexponential *Bateman* function  $f(t) = \exp\frac{-t}{\tau_1} - \exp\frac{-t}{\tau_2}$  that explains the physiological processes of EDA generation [197], refer to Figure 4.4 for details.

### 4.2.3.2 Integrated Phasic Response (IPR) Analysis

Ledalab<sup>3</sup> can decompose superposed EDA peaks into independent SMNA pulses, thus enabling the separation of phasic peaks. An advantage of Ledalab is the resulting phasic EDA has a zero baseline, enabling us to compute the time-integral of phasic EDA over a response window, which is a measure of sympathetic activation from the stimulus [197]. After decomposition, we slice individual SMNA peaks and reconvolve them with the IRF to obtain individual non-overlapping zero-baseline phasic EDA peaks. We take time-integral of these peaks, to obtain an EDA scoring measure defined as integrated phasic response (IPR) [197].

<sup>3</sup>Code available: <https://github.com/brennon/Pypsy>

Table 4.5: User-wise integrated phasic response (IPR) (in  $\mu Ss$  units) and Peak Count Analysis. Higher scores indicating stronger sympathetic activation, in each Correct Service (CS)/Wrong Service (WS) pair for each user are bold-faced.

(a) IPR in Service Groups			(b) Number of Phasic peaks			(c) IPR per Peak in Service Group		
User	CS	WS	User	CS	WS	User	CS	WS
A	523.15	<b>555.43</b>	A	<b>66</b>	63	A	39.16	<b>44.3</b>
B	235.7	<b>363.21</b>	B	21	<b>36</b>	B	<b>55.91</b>	50.06
C	80.6	<b>151.85</b>	C	10	<b>30</b>	C	23.88	<b>25.52</b>
D	51.41	<b>78.87</b>	D	14	<b>21</b>	D	16.38	<b>20.24</b>
E	7.77	<b>52.7</b>	E	17	<b>33</b>	E	1.69	<b>8.8</b>
F	14.54	<b>88.33</b>	F	16	<b>37</b>	F	3.34	<b>11.688</b>
G	67.65	<b>86.29</b>	G	22	<b>41</b>	G	<b>15.32</b>	11.39

#### 4.2.4 EDA Analysis Results

Our goal for this experiment was to identify instances of higher sympathetic activation (which can be used as an indicator of technostress, given a known context), due to wrong-services from a CAIE based on validated physiological indicators. We conducted a user study and have collected data from 7 participants (6 males, 1 female) under a research protocol approved by Virginia Tech (IRB-15-1193). Participants represented a wide range of nationality and ethnicity. The results from our batch analysis of EDA features, accumulated per event type show heightened sympathetic activation during WS events when compared to CS events based on validated physiological indicators. The results of this analysis will enable the CAIE to decide, *when* to ask clarifying questions in an adaptive-window based multi-turn interaction as discussed in Section 3.1 section.

The number of significant phasic peaks and time-integral of phasic peaks are widely used EDA features, wherein a higher number represents stronger sympathetic activation [214]. To perform the IPR analysis, the individual phasic peaks are thresholded to above 5% of the userwise maximum peak-amplitude to mark the significant peaks. Time-integral of these individual phasic peaks, where time is measured in seconds and phasic EDA in  $\mu S$ , are

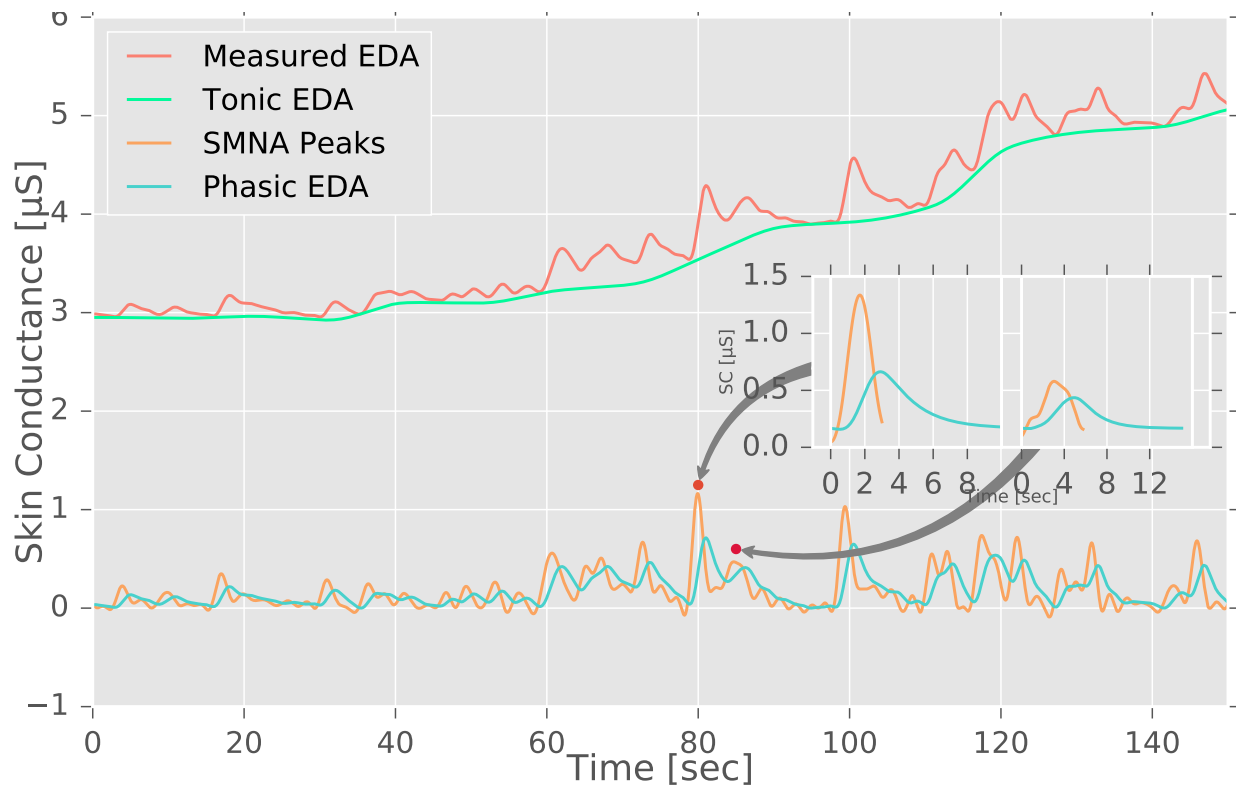


Figure 4.4: EDA Decomposition using Ledalab. Observe the *tonic* EDA follows the measured EDA signal, while sliced individual SMNA peaks are convolved with IRF to obtain zero-baselined *phasic* EDA (see inset). Notice the overlapped *phasic* peaks are separated as individual peaks. Image reproduced from Saha et al. [41].

computed and accumulated for each type of services i.e. correct (CS) and wrong (WS) within their respective windows to obtain the IPR values (in  $\mu Ss$  units). The results are compiled in Table 4.5 where the bold-faced numbers are higher among the CS/WS pairs for each user. We can see that for all users, IPR during the WS events is higher than that during the CS events. In addition, IPR per Peak is computed by dividing the total IPR by the number of peaks following a service, then accumulating for each service type. We see that for five users, the IPR per Peak is higher during the WS events. The number of significant phasic EDA peaks is also compiled, and barring User A, we obtain higher number of significant phasic peaks during WS compared to CS events. Time-spans for each events

depend on the length of the paths, however, WS events induce *higher* number of phasic peaks each with *greater* IPR (as seen in Table 4.6b-4.5c) indicating stronger SNS activation. It must be noted, that with a more liberal thresholding (say, 15%) for peak significance, the results for User A in Table 4.6b and for User B and G in Table 4.5c are consistent with the overall results.

Although, there are some users (esp. B and G) for whom the physiological indicators did not reflect these patterns, we have learned that such differences may arise from factors such as personality [112]. We do not have personality related data in our current dataset, however, adding such qualitative data collection methods for future studies should help in data analysis.

#### 4.2.5 Virtual Supermarket Experiment Discussion

With this current work, we sought to identify user-independent physiological indicators of stress experienced by users in CAIE, when they receive an inappropriate service. Table 4.5 shows that the number of phasic EDA peaks and average IPR in these peaks is higher during WS events, i.e. more numbers of larger phasic EDA peaks are produced during WS events. From this, we can infer that users show higher sympathetic activation during the WS events compared to the CS events. Thus, from our experimental dataset in a VR environment, we observed patterns in EDA signal across users during such WS (i.e. inappropriate or wrong services), that have been shown to be correlated with negative emotional states [131] such as frustration. The hypothesis behind our interaction framework as discussed in Section 3.1 section, rests on the successful identification of such affective states from physiological data. Our results from the batch analysis show greater number of phasic EDA peaks each having higher average IPR, both of which are independent evidences of stronger sympathetic

activation in users while experiencing technostress in a CAIE. Although individual event-wise analysis is not conclusively consistent across all users, however, with further analysis of more EDA features and HRV signals, we hope to improve upon the granularity of these discriminatory inferences to, possibly, a single window following each service. Nevertheless, the patterns from this group analysis will enable a CAIE to improve a multi-turn interaction (see Section 3.1 section) using features of technostress.

In addition to continued analysis of the physiological signals, we are also refining our experimental protocol, in order to gain more insights into known influencers of human affective responses such as their personality [112], thereby helping us improve our inferences. For instance, a recent work has demonstrated that the daily usage pattern of a mobile phone is predictive of a person's personality types [79]. While collecting such mobile usage data is out of the scope of our work, we intend to add qualitative data collection methods such as personality questionnaire.

#### 4.2.6 Virtual Supermarket Experiment Conclusion

In this section, we have proposed a novel system architecture to employ affective computing techniques to identify a user's states showing sympathetic activation arising from wrong services. Successful identification of such states (a surrogate for technostressed states) following a service, implying it's inappropriateness, can be used as a feedback signal in order to refine the services in subsequent turns. To evaluate this hypothesis, we have designed a controlled experimental platform in a VR setup providing intelligent services, and occasionally provided wrong services, while collecting real-time physiological data. The results from EDA signal analysis from our study conducted in the experimental platform show heightened sympathetic activation during wrong services, indicating onset of negative-emotional states such as

technostress. These results are encouraging as we continue to refine our setup and analysis. The results from this experiment validate another hypothesis in our CAFFEINE framework where we hypothesized the feasibility of using technostress as a feedback signal in a CAIE. This experiment shows that technostress indeed produces a physiological signature similar to psychological stress with heightened sympathetic activation. Thus, recognition of physiological features of technostress using ambulatory wearable sensors is indeed possible. Having validated both the components of the CAFFEINE framework, our final quest in this dissertation work is to find a method to computationally estimate the quality of the technostress inference using physiological signals. Towards that end, in the next section, we present a novel algorithmic and experimental approach towards defining a physiological signal based quality metric.

### **4.3 Experiment Three: Defining an Inference Quality Metric using Physiological Signals**

In this section, we will discuss our experimental setup to derive a physiological signal based inference-quality metric. The experimental protocol and interaction in the VR supermarket remain largely unchanged from the setup described in Section 4.2 (also presented in Saha et al. [15]), please refer there for a detailed explanation. In this section, we present a highlight of the modifications made to the existing protocol for the Intelligent Supermarket in Virtual Reality, in order to answer the 4th research question presented in Section 1.2.1, namely to derive a computational method for assessing the quality of technostress inference using physiological signals.

## 4.3.1 Modified Experimental Protocol

### 4.3.1.1 Baseline Sonic-Impulse Phase

As described in Section 3.3.2, our central idea for assessing the quality of technostress inference using physiological signals is to use a comparison between an individual's physiological response to a known stimulus and that during the experience of technostress. To assess each subject's response to a known stimulus, we used a sudden excitation (or impulse), such as a sonic impulse stimulus e.g. a balloon-pop sound, physiological response to which is used in our confidence measure computation (please refer Section 3.3.2). A few samples of unfiltered, zero-baselined and normalized EDA impulse responses from our dataset can be seen in Figure 4.5. To create a uniform experimental condition, users were asked to put on an isolating headphone to listen to a calming stimulus (such as a uniform white-noise sound) for one minute. Following this, a short-duration (approx. 100ms) impulse stimulus was played preceded and succeeded by silence for a minimum of five seconds. Physiological data was captured on a local computer being streamed from Empatica E4 device (see Section 4.3.2) while users listened to these sounds.

### 4.3.1.2 VR SuperMarket Phase

Participants were informed that their shopping list was pre-populated with 16 items, and item numbers corresponding to the subsequent item will be shown as an overlay on the supermarket scene in their VR headsets. Participants were informed that the system will highlight a direct path from their current position to the destination using horizontal green path-arrows, while a vertical red-arrow will indicate the final destination visible from their current location. In reality, the experiment was conducted in a Wizard-of-Oz fashion, where-



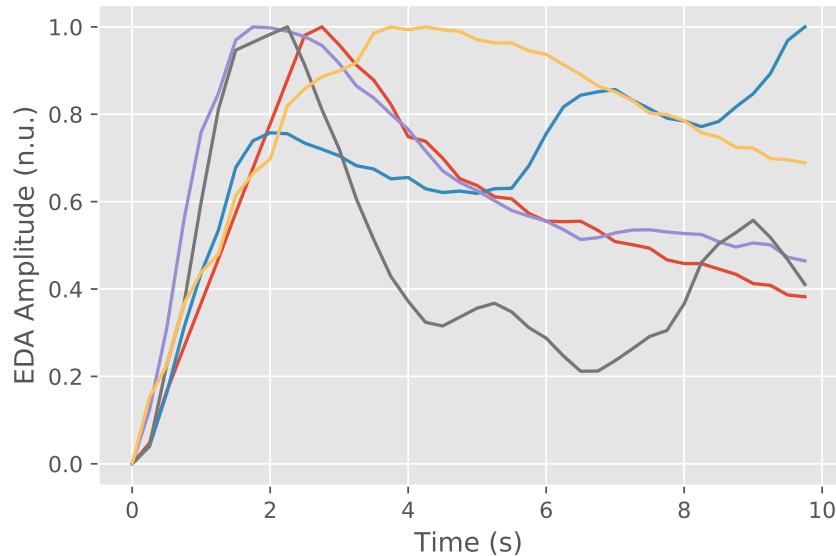


Figure 4.5: A few samples of normalized EDA responses to sonic impulse stimulus in our dataset (refer to Section 4.3.1.1 for details). Figure reproduced from Saha et al. [42]

in the experimenter would listen to the participant reading the next-item-number on their VR headset and activate the next-path-arrow by pressing a *hotkey* in the experimenter’s view (see section Modified Hardware Setup). Some of the paths were an obvious direct path, while some paths were deliberately made winding to create an impression of a system failure thus imparting technostress. Participants were asked to always follow the path indicated by green arrows after locating the destination red-arrows from their current location.

Out of the 16 items on the list, we provided correct service (CS: direct path) and wrong service (WS: winding path) for 8 items each, interspersed in groups of **G1**(6 CS) → **G2**(5 WS) → **G3**(2 CS) → **G4**(3 WS) in that order. Our hypothesis behind this interspersing was to experimentally discard the ordering effect of service groups i.e. our hypothesis was to observe repeating patterns of physiological indicators of NS → S → NS → S states corresponding to CS/WS services.

### 4.3.2 Modified Hardware Setup

The position tracking setup in The Cube at Virginia Tech consists of a Qualisys Motion Capture system with 24 Oqus5+ cameras for tracking reflective marker based rigid-bodies attached to an Oculus headset. The Qualisys system provides the 3D translation coordinates, whereas the rotation coordinates are read from the Oculus headset. The Oculus is connected to a VR Backpack computer (named Oculus Computer (OC)) which is running the supermarket model in Unity. For performing the experiment in Wizard-of-Oz fashion, the experimenter has a mirrored view of the participant's VR view (running on OC), onto a local computer (named Experimenter Computer (EC)) and communicating over the local wireless network. We used snap-fit finger electrodes from Lafayette Instruments<sup>4</sup> attached to a modified version of Empatica E4 device to collect physiological data from the user. The time-stamped biosignal datastreams are streamed over Bluetooth to OC. The *hotkey* presses (VR event onsets) are time-stamped with the OC machine-time along with Empatica E4 data packets using custom code<sup>5</sup>.

### 4.3.3 Quality Metric Methods and Analysis

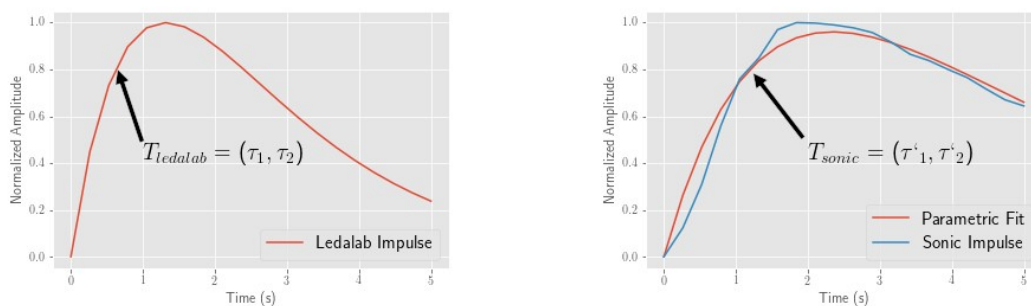
#### 4.3.3.1 EDA Signal based Quality Metric

As described in Section 3.3, we employ the signal quality of captured EDA datastream to derive a quality metric for our inference on affective feedback. Our quality metric calculation hinges on the idea that using a canonical EDA impulse response function may result in sudomotor nerve activity (SMNA) pulsetrain having negative values, mainly due to individual differences in EDA responsivity [197]. To overcome this, we proposed an idea of using

<sup>4</sup><https://www.lafayetteinstrument.com/>

<sup>5</sup>(available at <https://github.com/debapratimsaha/EmpaticaUnityBLEClient>)

the response of each user to a known sonic impulse stimulus to calibrate our system. A detailed overview of our quality metric computation is presented in Section 3.3, reader is highly recommended to read the section before proceeding ahead. In the next section, we will briefly discuss our experiments with deriving an algorithm for computing EDA based quality metric.



(a) A sample of EDA IRF obtained from Ledalab Decomposition of Intelligent Super-Market Data for a user.

(b) A sample of EDA IRF obtained from the Sonic Impulse phase and the parametric curve-fit shown for the same user.

Figure 4.6: EDA Signal based quality metric computation by comparing IRF obtained from Ledalab decomposition and sonic impulse.

#### 4.3.3.1.1 Point-wise Error Computation Algorithm

An initial approach for comparing the IRF response obtained from sonic impulse and Ledalab framework optimization process, was to use point-to-point comparison of the filtered and normalized waveforms. For a Sonic IRF timeseries denoted by  $SIRF_k \forall (k = 0 \dots n)$  and Ledalab IRF timeseries denoted by  $LIRF_k \forall (k = 0 \dots n)$ , an error term ( $E_{irf}$ ) was defined as the absolute percentage deviation between each pair of  $|SIRF_k - LIRF_k| \forall (k = 0 \dots n)$ , averaged over all the  $n$  values. However, we found that the error term obtained using this method is highly dependent on measurement noise as well as prone to shape outliers.

Following initial investigation, we decided to use parametric comparison instead of point-wise comparison. The algorithm for parametric comparison has been described in Section 3.3. For the sake of completeness, the algorithm will be briefly described in the next section.

#### 4.3.3.1.2 Parametric Error Computation Algorithm

EDA impulse response has been hypothesized to follow the Biexponential Bateman function (see Equation (3.1)) which is a pharmacokinetic model representing the time-course of drug diffusion in a compartment body model. This equation consists of a set of two parameters ( $T = \tau_1, \tau_2$ ). Ledalab optimization framework provides the IRF in this parametric form, denoted by  $T_{ledalab}$ . For the sonic IRF, we use parametric curve-fitting with Equation (3.1) as the target function to obtain  $T_{sonic}$ . Following this, an error term ( $E_{irf}$ ) was defined as the absolute percentage deviation between each pair of  $\tau$  parameters. For a detailed discussion, please see Section 3.3. A sample of Ledalab IRF compared with Sonic IRF is shown in Figure 4.6.

### 4.3.4 Quality Metric Results and Discussion

#### 4.3.4.1 EDA Signal based Parametric Quality Metric Results

Our goal for this work was to experimentally validate a *quality* metric derived from physiological signals to interpret the implicit AF inference. Towards this end, we refined our experimental setup to collect user’s physiological response to a sudden excitation (refer Section 4.3.1 for the modifications). We present the validation of our  $Q_{eda}$  metric using validated discriminatory EDA features [41]. We have used Ledalab<sup>6</sup> to decompose the measured EDA signal into *tonic* and *phasic* components as well as their respective drivers. The resulting

<sup>6</sup>Code available: <https://github.com/brennon/Pypsy>

Table 4.6: User-wise  $Q_{eda}$  measure, Cumulative IPR and Number of Phasic Peaks grouped by Event type CS (correct service) and WS (wrong service). User E sensor data dropped out in the middle of G4.

(a) $Q_{EDA}$ Measure		(b) NSPP in CS/WS Group			(c) IPR in CS/WS Group		
User	$Q_{eda}$	User	CS	WS	User	CS	WS
A	-	A	34	<b>52</b>	A	438.1	<b>673.3</b>
B	34.6	B	27	<b>31</b>	B	<b>685.1</b>	553.7
C	48.9	C	21	<b>37</b>	C	<b>284.9</b>	261.3
D	57.0	D	29	<b>39</b>	D	130.9	<b>146.9</b>
E*	58.4	E*	39	<b>50</b>	E*	663.2	<b>757.5</b>
F	54.5	F	38	<b>71</b>	F	100.8	<b>347.3</b>

phasic driver (being a zero-baselined signal) can be used to compute a continuous measure of phasic activity called Integrated Phasic Response (IPR)—refer Saha et al. [41] for a discussion on calculating this feature. We also compute an EDA signal based  $Q_{eda}$  metric following the method described in Section 3.3.2. These features and the  $Q_{eda}$  metric constitute our basic computational framework for evaluating the hypothesis in our interaction model defined in Chapter 3. We have collected data from 6 participants under a research protocol approved by Virginia Tech (IRB-15-1193). Participants were recruited using advertisement emails, and special care was taken to discard participants who had already participated in Phase-1, to avoid precedence effects.

The number of significant phasic peaks (NSPP) and the time-integral of phasic peaks (reported by IPR value) following a stimulus have been reported to be reliable indicators of sympathetic activation [186]. For computing IPR, we have followed the method described in [41]. Significant phasic peaks are computed above a threshold of 5% of userwise maximum phasic driver amplitude (a figure considered in [197] as well). Time-integrals of these significant phasic peaks, measured in ( $\mu Ss$ ) units, are computed and grouped by service type (CS: CORRECT service) and (WS: WRONG service). The results compiled from EDA analysis as well as Quality analysis are reported in Table 4.6. Please note that for User-A, the impulse response was not included in the experiment, so we do not have a  $Q_{eda}$  measure for the user.

In Table 4.6c and Table 4.6b, the bold faced numbers are higher among the **CORRECT**/**WRONG** pairs of events for each user. We see that for all users, NSPP are higher during **WRONG** events compared to **CORRECT** events. For four users, IPR is higher during **WRONG** events compared to **CORRECT** events. We will discuss the case of Users B and C in Section 4.3.4.2.

As discussed in Section 4.3.1, we have interspersed **CORRECT** and **WRONG** events in groups, with a hypothesis to observe higher sympathetic activation during all **WRONG** event groups compared to all **CORRECT** event groups in the sequence  $G1 \rightarrow G2 \rightarrow G3 \rightarrow G4$ . In Table 4.7, we report the groupwise per-event IPR and per-event NSPP analysis results. We observe that for Users A, E and F, per-event IPR during all **CORRECT**/**WRONG** event group transitions matched our hypothesis. But, for Users B, C and D, per-event IPR do not match our hypothesis during transitions  $G1 \rightarrow G2$ . Similarly, on per-event NSPP analysis feature, all users except User B match our hypothesis for all transitions of service groups. The case of Users B and C will be discussed in Section 4.3.4.2.

#### 4.3.4.2 Interpreting the $Q_{eda}$ Quality Score

Our hypothesis for the experiment was to observe high sympathetic activation during groups of **WRONG** events compared to groups of **CORRECT** events. Following Equation (3.2) in Section 3.3.2, and assuming  $\theta = 50\%$ , we observe in Table 4.6a that for Users B and C,  $Q_{eda} < 50\%$ . This implies low confidence in EDA features for Users B and C, which may explain the observed mismatch with our hypothesis as seen in Table 4.6c, where IPR during **CORRECT** events is higher compared to **WRONG** events. In Table 4.7 also, we observe that the transition  $G1 \rightarrow G2$  violates our hypothesis for Users B and C on per-event IPR feature and for User B on per-event NSPP feature. On the other hand, for Users D, E and F,  $Q_{eda} > 50\%$ , and their IPR for **WRONG** events is higher than **CORRECT** events in Table 4.6c, matching our hypothesis. In Table 4.7 also, we see Users E and F match our hypothesis on both the features.

Table 4.7: Per-Event IPR and Per-Event NSPP in sequence of Event groups: G1(5CS), G2 (5WS), G3 (2CS) and G4 (2WS) (CS:correct service, WS:wrong service). User E sensor data dropped out in the middle of G4.

User	G1		G2		G3		G4	
	NSPP	IPR	NSPP	IPR	NSPP	IPR	NSPP	IPR
A	4.4	56.0	6.4	103.5	4.0	46.7	6.0	78.0
B	3.8	91.2	3.8	60.8	2.5	15.2	5.0	39.7
C	2.8	46.8	4.8	31.7	2.0	5.9	4.0	36.7
D	3.8	17.5	5.3	15.2	3.5	11.3	5.5	16.8
E	5.0	84.2	6.8	117.7	4.5	85.7	7.5	84.5
F	4.4	14.3	8.3	45.8	6.5	14.4	10.0	42.4

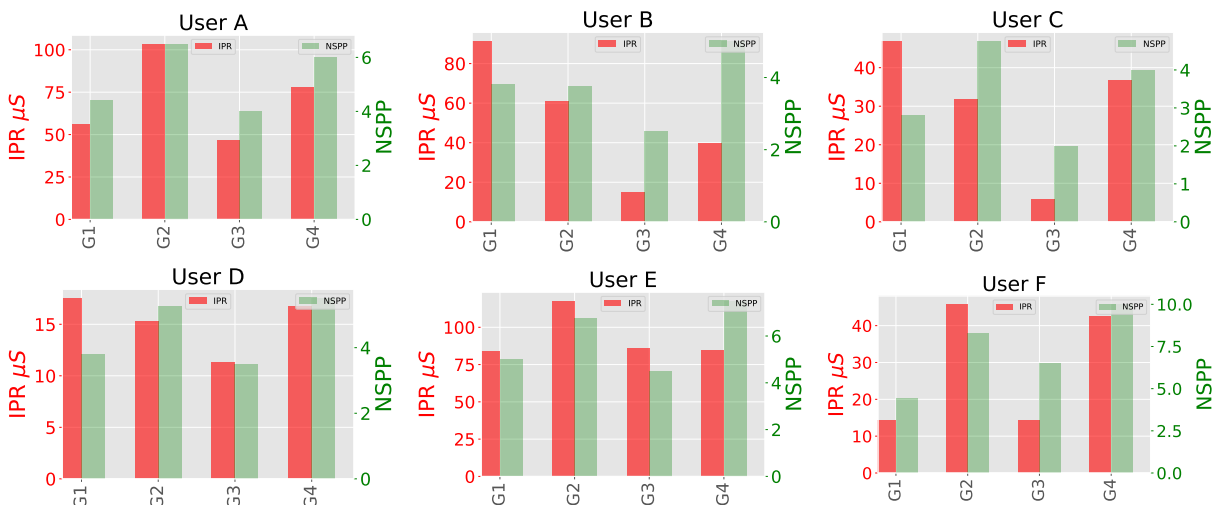


Figure 4.7: Groupwise Per-Event IPR and Per-Event NSPP Analysis Plots

The only outlier to our quality/confidence model is for User D in per-event IPR feature for G1→G2 transition. Thus, such a  $Q_{eda}$  thresholding may enable the CAIE to calibrate the implicit AF loop by ignoring EDA features for the window where  $Q_{eda} < 50\%$ . We must note here that, the example of a threshold presented here is empirical, more experimental validation is needed to arrive at a more accurate  $Q_{eda}$  threshold.

### 4.3.5 Quality Metric Discussion

In this experiment, we sought to define a quality measure for the user independent discriminatory physiological features that can be used as indicators of *frustration* arising due to experiencing technostress from a CAIE. For defining a quality score for the recorded EDA signal and the decomposition method, we observed that the shape of the IRF derived as a result of Ledalab optimization process, can be compared with an individual user's EDA response to a sonic impulse (see Section 3.3.2 for details). Subsection 4.3.4.2 demonstrates, that a threshold of  $\theta = 50\%$  on  $Q_{eda}$  can be used to arrive at meaningful interpretations of results in Table 4.6 and Table 4.7. The results demonstrate that all users with  $Q_{eda} > 50\%$ , match our hypothesis in groups of CORRECT/WRONG events and their transitions. For users with  $Q_{eda} < 50\%$ , the data does not match our hypothesis on the IPR and NSPP features both cumulatively and in group transitions. Although it is worth mentioning, that transitions G2→G3→G4 always matched our hypothesis. In this phase of our experiment, we have used finger electrodes attached to the Empatica wristband, which captures higher EDA amplitudes compared to the wrist EDA enabling better phasic peak scoring. However, finger electrodes with longer wires may pick more motion artifacts, which could be a possible explanation for low score on  $Q_{eda}$  metric.

Thus, such a thresholded  $Q_{eda}$  score may be effectively used as a surrogate for confidence in the discriminating quality of the features, enabling the CAIE to calibrate it's response in case of an (in)appropriate service. Additionally, in interpreting  $Q_{eda}$  score with an empirical threshold  $\theta$ , we may generalize the rule such that it may be used as surrogate for a direction of change in EDA feature values, instead of a hard-set binary threshold. The  $Q_{eda}$  threshold presented here is purely empirical at this stage, and needs more experimental validation to arrive at a meaningful value and a suitable way to interpret it's applicability (i.e. yes/no



binary threshold versus direction of change threshold). Nevertheless, such a score may be a valuable step towards defining a final confidence score. For example, this  $Q_{eda}$  score based approach might be used as a weight parameter in the  $F_\beta$  score based classifier design presented in [40]. Inferences used in an affective feedback for our interaction framework will need to be from multimodal signals, each of which may have such physiologically grounded quality scores. These quality scores may finally be combined together along with the probability outputs from a classifier to form a final confidence score on the service-(in)appropriateness inference, guiding the strategy for reconfiguring subsequent services, shown using process marked “5” in Figure 3.3 at the user-system interface.

### 4.3.6 Quality Metric Experiment Conclusion

In this section, we have proposed a novel method of deriving a quality metric ( $Q_{eda}$ ) for the EDA datastream using an individual’s response to a sonic impulse. This quality metric is envisioned to be used as a surrogate for a confidence measure in the inference drawn on technostressed states in our interaction framework. We have collected a new dataset from a refined version of a previous experimental setup wherein users are immersed in a virtual reality setup and occasionally given a wrong service. The EDA analysis reveals that users show heightened sympathetic activation while they are given wrong services from the CAIE. In addition, our proposed  $Q_{eda}$  metric improves our capability to explain the observed results from EDA analysis on validated physiological markers such as IPR and NSPP. These results are encouraging, as we continue to refine our proposed quality metric to derive validated thresholds.

# Chapter 5

## Summary, Conclusions and Reflections

---

*“A clear understanding of negative emotions dismisses them.”*

— Vernon Howard

---

The work presented in this dissertation was motivated by the need to design a mechanism in a context-aware intelligent environment (CAIE) that can implicitly infer a user’s feedback

on the appropriateness of the services received from the CAIE. In this work, we set out to study methods and develop protocols in computational psychophysiology in order to design the aforementioned implicit feedback loop in a CAIE using physiological signals. In doing so, we have borrowed from and advanced the discussion on the symbiotic relationship between *Affective Computing* and *Ubiquitous Computing* domains, during experiment design and user-evaluation of our research methods. We proposed a computational framework, called CAFFEINE (presented in Chapter 3), and conducted various user-studies to demonstrate the feasibility of multiple components of the framework (presented in Chapter 4). We developed new experimental protocols, algorithms, research methods and experimentation tools as a part of the process to evaluate our hypotheses presented as components of the CAFFEINE framework.

In the Introduction chapter, we stated our overarching research question as:

**What are the methods in which affective computing techniques can be used to design an automatic service-appropriateness feedback in a real-life human-centric CAIE?**

This question was further broken into a few sub-questions as listed below:

- (i) What is the interaction schema to effectively incorporate an implicit-feedback loop in a context-aware intelligent environment?
- (ii) What are the basic design parameters, signal-features and methods for evaluation of the interaction scheme of a physiological-signal based affective feedback loop in a CAIE?
- (iii) Which parameters of the physiological signals are critical in performance improvement of a technostress based service-relevance feedback loop?
- (iv) How to estimate the quality of the inference drawn from the physiological signals?

## 5.1 Discussion on Research Questions

In this dissertation, we proposed a novel interaction framework called CAFFEINE, which may enable a CAIE to establish a human-centric and naturalistic communication channel with its users. We posit that inferring implicit feedback from users about the appropriateness of services provided by these CAIEs will reduce the cognitive load on the occupants rather than requiring them to explicitly provide feedback, thereby allowing the smart environment to disappear from a user’s consciousness, and hopefully achieving the ideal goal of “invisible, everywhere computing”, originally described by Mark Weiser in his vision for ubiquitous computing. Consequently it is envisioned to restore the sense of agency in the occupant as pointed out in Section 2.5, thereby improving the system’s usability. In this work, we validated multiple critical components of our CAFFEINE framework using a battery of experiments. Specifically, we chronologically addressed the research questions presented in the introduction.

### 5.1.1 Validating CAFFEINE Interaction Scheme

To address research questions 1 and 2, we proposed an interaction scheme for CAIEs wherein the user-feedback for service-appropriateness is computationally inferred by detecting technostressed states. In our framework, the ground-truth for these states for individual users is envisioned to be collected from validated stress-induction instruments such as Paced-Stroop Test (PST). We demonstrated that recognition of technostressed states in response to wrong services from a CAIE was indeed possible using physiological signals collected from wearable sensors. In doing so, in experiment one as described in Section 4.1, we have prototyped a warehouse scenario wherein an order-picker personnel is given a context-aware service to

help them complete their task in a more efficient and error-free manner. These services sometimes failed in providing correct services, thereby causing technostress. We have developed a machine-learning system using Support Vector Machine (SVM) classifier to recognize technostressed states in response to *wrong* responses from the system using data collected from user's physiological signals, specifically electrodermal activity and electrocardiography signals. We demonstrated that recognition of technostressed states was indeed possible, and the system performance improved when we included data from PST in our training set. These results, originally presented in [40], provide an experimental validation of our interaction framework. In addition, the results validated the ground-truth learning component of CAFFEINE framework by showing improvements in the system performance when using PST data as a training set for the classifier.

### 5.1.2 Identifying Technostress Physiology

Following the experimental validation of the personalized ground-truth learning component of our framework, we set out to validate the physiological underpinnings of technostressed states in order to answer the research question 3. In particular, our quest was to identify physiologically validated markers of psychological stress that also arise during situations of technostress in a CAIE, such as indicators of increased sympathetic activation. In order to design an immersive experimental setup, we have prototyped a general intelligent supermarket scenario in a position tracked virtual reality space at our research institute, the Institute for Creativity, Arts and Technology, which is described in more details in Section 4.2. In this experiment, a user is immersed in a virtual reality setup of an intelligent supermarket which provides navigation assistance highlighting a direct path from their current location to a destination. Sometimes these services fail to provide a direct path, causing technostress. We have identified a few physiologically validated features from the EDA datastream such

as integrated phasic response and number of phasic peaks which are indicative of heightened sympathetic activation. From our user study, these features have been demonstrated to contain discriminatory information for differentiating correct services from wrong services, which may potentially improve the technostress classifier performance described in the previous section. The results of this study, originally presented [41], validates our hypothesis of using technostress as a service-appropriateness signal in intelligent environments, as well as help us in identifying validated physiological features.

### 5.1.3 Deriving A Quality Metric

A computational framework needs a method to assess the quality of its inferences. For CAFFEINE framework, our quest was to design a physiologically validated method of quantifying the quality of a technostress inference in order to address research question 4. We proposed a novel method of computing EDA signal based quality metric using the underlying phenomenon of an impulse-response function used for EDA decomposition, and through a user-study, we empirically validated the idea. This quality score is envisioned to dynamically assess the correctness of the technostress recognition, thereby helping the CAIE to decide if it needs to reconfigure its service or discard the inference altogether. In our approach, we modified the previous experimental setup, as described in Section 4.3, by adding a sonic impulse phase and recording the impulse response shape from each user. This shape is considered as a user-specific impulse response template which is compared with the canonical impulse-response function (IRF) obtained from the EDA decomposition tool called Ledalab. Results from this experiment, originally presented in [42], validate our idea of a physiological-signal based quality metric computation considering an empirical threshold.

## 5.2 Observations and Reflections

In this section, we will present some of our observations and reflections from the evaluation methods, user-studies and applications presented in this dissertation, and discuss some logical next steps for extending this work.

### 5.2.1 New Sensing Modalities and Instruments

**New Sensing Modalities and Instruments:** This dissertation has proposed the use of electrodermal activity and cardiovascular activity in order to assess the technostress states of users in a CAIE. However, collecting motion-artefact free data from locations suitable for better wearability in ecological environments is challenging. Encouraged by the evidence produced by the experiments using electrodermal activity and heart-rate variability measured from wearable wristband form-factor devices, a logical next step is to broaden our sensing modalities suited for a CAIE, to include data streams containing information on both emotional arousal and valence. In addition to ECG and EDA, some recent stress recognition studies have shown effective use of respiration rate [76, 74], pupillary dilation [187] and trapezius muscle electromyography (EMG) [215] as good predictors of mental stress.

Apart from wearable devices, environmental sensing of affective and behavioral signals, such as using infrared thermographic (IRT) cameras [216] or sensory chairs (for office use) [217] can be helpful in deriving the affective/behavioral states of the user. IRT cameras capture facial temperature variations which is a potent signal stream for sensing changes in human ANS, as well as enable us to infer various secondary physiological parameters such as breathing-rate, cardiac-pulse and cutaneous blood perfusion as noted in Cardone et al. [216]. Apart from these, IRT images have been recently shown to be useful in inferring changes in skin-

sweat, correlating with changes in electrodermal activity, thereby enabling off-body sensing of physiological indicators of affect [216]. These open up new vistas of modeling and sensing stress-episodes.

Additionally, recent prior works such as Plarre et al. [76] and Shi et al. [74] have used various other kinds of social, mental and physical stressors such as public speaking, mental arithmetic, cold pressor challenge *etc.* as stress induction instruments for improved physiological profile learning component (described in Section 3.2). An interesting future direction may be to explore the use of varied classes of laboratory stressors to model the ground truth of a real-life stressor by experimentally weighting the contributions from each kind of stressor that might constitute a real-life stressor.

**Wearable Sensor Placement:** Sensor placement has an impact on the quality of acquired data, as well as the usability of the wearable devices. For example, in experiment 3, we captured EDA data using Empatica at the fingertips whereas in experiment 2 we captured it at the wrist (placement as intended in original device design). Signal amplitudes are diminished in the wrist sensor placement compared to fingertip resulting in lower signal-to-noise ratio impacting feature extraction and inference. However, using fingertip sensor in the wild is not scalable. Similar observations can be made about heart-rate signal captured from photoplethysmography at the wrist (used in experiments 2 and 3) compared to electrocardiography at the chest (used in experiment 1). Thus, novel sensing locations for EDA and HR datastreams, such as a finger ring form-factor device may prove helpful. In addition, future designs may provision for embedding EDA electrodes with co-located pressure sensors which may be used to provide control for motion artifacts induced by changes in electrode pressure, which have considerable effect on signal quality [141].



## 5.2.2 New Computational methods

**Computational Models for Technostress Recognition:** In future work, this dissertation can be extended by incorporating a continuous temporal model based Dynamic Bayes' Networks and which use an accumulation-decay model, on similar lines as the work by Plarre et al. [76]. The central idea behind an aggregation model of perceived stress, presented in the work by Plarre et al., is the fact that the perceived stress for each individual may take a long time to decay after a stressor is withdrawn, in addition to accumulation of physiological effects from repeated successive stress events which may show pronounced effects and take even longer to fade away. Plarre et al. implemented such a model in [76] using a Dynamic Bayes Network with personalized accumulation and decay constants ( $\alpha, \beta$ ) to account for inter-individual difference of stress-responsivity (these parameters are fit using experience sampling questionnaires).

**Transfer Learning Approach for Physiological Profile Learning:** The physiological profile-learning architecture proposed in this dissertation, as discussed in Section 3.2, argues the need for collecting ground-truth data from proven stressors such as paced Stroop test or validated musical stimuli. Using this ground-truth data, we have already designed a computational stress-recognition pipeline for which the target task is to recognize technostressed states in response to services received from the CAIE as described in Section 4.1. From this architecture we can surmise that the train and the test tasks are similar, i.e. learn a person specific stress-model from physiological data and predict a technostressed state. However, we hypothesize that the distribution of the data in the two phases may not be similar owing to the difference in the nature of stress elicitation. Nevertheless, this kind of problem is well-suited for this specialized area of *transfer learning*, wherein the learned representation of stressed states from PST can be intelligently transformed to recognize technostress

states, thereby improving the classifier performance of Section 4.1. In Section 2.4.3.3, we have briefly delved into the details of transfer learning technique and how to use it for this framework. An interesting recent work uses inductive transfer learning techniques to address the well-known problem of individual differences [218], and report improved recognition accuracy on a similar task—Virtual-Reality Stroop Test which is a modified version of the Stroop color-word interference test that we are using.

## 5.3 Future Applications for CAFFEINE Framework

During the course of this dissertation, we have received several suggestions to extend our work in novel application domains. One particularly interesting scenario is in creating novel human-building interaction applications explained here.

### 5.3.1 Novel Systems for Human-Building Interaction

A new sensor-rich building at the nexus of arts and technology, called Moss Arts' Center at Virginia Tech has been used to design a novel human-centered intelligent infrastructure called Mirror-Worlds (MW), in the mould of a broad idea of Ubiquitous Computing sketched out by Mark Weiser [219]. The MW system was designed to study building-sized interaction such as crowd-simulation, path-finding and navigation, spatial-analytics and visualization, and augmented environments. The underlying cyberinfrastructure, named Fusality, is modeled as a real-time server-client architecture being fed data from *producers* and projecting real-time analytics on *consumers*, in addition to enabling real-time interaction with a virtual world from a physical world using *producers+consumers* [3]. The *producers* in the MW system mainly consist of fisheye optical cameras affixed to the ceiling which uses computer-vision

techniques to analyze coarse human movements around the smart-building, and projects these anonymized movements on MW clients which are a gamut of display devices around the building enabling real-time visualization. This sensing modality captures coarse group analytics data enabling services and applications such as path planning and evacuation, classroom student behavior analysis etc. However, it lacks the specificity and individuality of physiological sensors which we argue are needed to broaden the ambit of the MW system's services and may enable the system to create individualized implicit feedback loops envisioned in Section 2.5.

In our quest to broaden the scope of our system from experimental laboratory setting to a real-world building-sized intelligent infrastructure, we have introduced a novel *producer* to this MW system in the form of a physiological sensor by streaming real-time data from Empatica wristband to the Fusality servers. The Empatica wristband streams electrodermal activity, blood-volume pulse, skin temperature and hand movement data. In addition, we are also in the process of integrating another *producer* to the MW infrastructure in the form of an infrared thermographic (IRT) camera as described in Section 5.2.1. Both these sensors enables the MW system to capture an individual's responses to a certain intelligent service it provided, and thereby infer the service's appropriateness based on the identification of a technostressed state as already described in this dissertation. Such an infrastructure consisting of real-time synchronized datastream of physiological and behavioral data will help in defining novel interactions with the building, such as inferring thermal comfort using physiological signals as already discussed in Section 2.5. A very recent work by Nkurikiyeyezu et al. [220] shows discriminatory features in HRV signals indicative of thermal comfort. Thus, our CAFFEINE framework may be used to design a more organic natural interaction with modern adaptive buildings.

## 5.4 Concluding Remarks

Our primary vision for this dissertation was to study computational methods that can be used to incorporate implicit feedback into context aware intelligent environments. In doing so, our work advances a discussion on the symbiotic relationship between context-aware computing and affective computing domains. This dissertation presents an exploration into various aspects of designing a computational framework, which entails validating its component processes by conducting multiple user studies and proposing novel algorithmic methods and interaction protocols envisioned to be used in ecological environments. We have chosen real-life applications and use-cases for our evaluation process, nevertheless, the scope of the study remained in controlled settings. While there remains numerous open questions that need to be addressed, this dissertation revealed several interesting results that open further lines of investigation into novel research problems in this domain. Multiple ongoing student research projects have fanned out as a result of our investigations and identified gaps in knowledge. To the best of our knowledge, this work takes a few small, yet concrete steps towards filling a knowledge gap in the field of creating implicit feedback mechanisms using physiological signals as applicable to intelligent environments. Ultimately, we hope to spur general interest in this area by laying out the groundwork and demonstrating a feasibility of this idea.

# Bibliography

- [1] J. Nembrini and D. Lalanne, “Human-Building Interaction: When the Machine Becomes a Building,” in *Human-Computer Interaction - INTERACT 2017*, ser. Lecture Notes in Computer Science. Springer, Cham, Sep. 2017, pp. 348–369.
- [2] M. Weiser, “The Computer for the 21st Century,” *Scientific American*, vol. 265, no. 3, pp. 94–104, Sep. 1991.
- [3] N. F. Polys, B. Knapp, M. Bock, C. Lidwin, D. Webster, N. Waggoner, and I. Bukvic, “Fusality: An Open Framework for Cross-platform Mirror World Installations,” in *Proceedings of the 20th International Conference on 3D Web Technology*, ser. Web3D ’15. New York, NY, USA: ACM, 2015, pp. 171–179.
- [4] M. Weiser, “Some Computer Science Issues in Ubiquitous Computing,” *Commun. ACM*, vol. 36, no. 7, pp. 75–84, Jul. 1993.
- [5] S. A. N. Shafer, B. Brumitt, and J. J. Cadiz, “Interaction Issues in Context-aware Intelligent Environments,” *Hum.-Comput. Interact.*, vol. 16, no. 2, pp. 363–378, Dec. 2001.
- [6] J. C. Augusto, V. Callaghan, D. Cook, A. Kameas, and I. Satoh, “Intelligent Environments: A manifesto,” *Human-centric Computing and Information Sciences*, vol. 3, no. 1, p. 12, 2013.

- [7] A. Mehrabian, *Silent Messages: Implicit Communication of Emotion and Attitude*. Belmont, CA: Wadsworth, 1981.
- [8] N. Thompson and T. McGill, “Affective Human-Computer Interaction,” in *Encyclopedia of Information Science and Technology, Third Edition*, 2015, pp. 3712–3720.
- [9] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. Taylor, “Emotion recognition in human-computer interaction,” *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32–80, Jan. 2001.
- [10] R. B. Knapp, J. Kim, and E. André, “Physiological Signals and Their Use in Augmenting Emotion Recognition for Human–Machine Interaction,” pp. 133–159, 2011.
- [11] R. W. White, *Interactions with Search Systems*. Cambridge University Press, Mar. 2016.
- [12] G. Pasi, “Implicit Feedback through User-system Interactions for Defining User Models in Personalized Search,” *Procedia Computer Science*, vol. 39, pp. 8–11, Jan. 2014.
- [13] H. A. Feild, J. Allan, and R. Jones, “Predicting Searcher Frustration,” in *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR ’10. New York, NY, USA: ACM, 2010, pp. 34–41.
- [14] A. Hassan and R. W. White, “Personalized Models of Search Satisfaction,” in *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management*, ser. CIKM ’13. New York, NY, USA: ACM, 2013, pp. 2009–2018.
- [15] R. W. White and R. Ma, “Improving Search Engines via Large-Scale Physiological Sensing,” in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR ’17. New York, NY, USA: ACM, 2017, pp. 881–884.

- [16] Y. Moshfeghi and J. M. Jose, “An Effective Implicit Relevance Feedback Technique Using Affective, Physiological and Behavioural Features,” in *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR '13. New York, NY, USA: ACM, 2013, pp. 133–142.
- [17] J. Sanders and G. Taubman, “Ranking query results using biometric parameters,” USA Patent, Feb., 2013.
- [18] A. K. Dey, “Understanding and Using Context,” *Personal and Ubiquitous Computing*, vol. 5, pp. 4–7, 2001.
- [19] M. Weiser, “Ubiquitous Computing,” <http://www.ubiq.com/hypertext/weiser/UbiHome.html>, 1996.
- [20] H. Prendinger, S. Mayer, J. Mori, and M. Ishizuka, “Persona Effect Revisited,” in *Intelligent Virtual Agents*, ser. Lecture Notes in Computer Science, T. Rist, R. S. Aylett, D. Ballin, and J. Rickel, Eds. Springer Berlin Heidelberg, Sep. 2003, no. 2792, pp. 283–291.
- [21] E. Hudlicka, S. Payr, R. Ventura, C. Becker-Asano, K. Fischer, I. Leite, and C. Von, “Social interaction with robots and agents: Where do we stand, where do we go?” in *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009*, Sep. 2009, pp. 1–6.
- [22] E. Hudlicka, “Virtual Training and Coaching of Health Behavior: Example from Mindfulness Meditation Training,” *Patient education and counseling*, vol. 92, no. 2, pp. 160–166, Aug. 2013.

- [23] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, “Affect recognition for interactive companions: Challenges and design in real world scenarios,” *Journal on Multimodal User Interfaces*, vol. 3, no. 1-2, pp. 89–98, Mar. 2010.
- [24] G. Castellano, Q. Mary, R. Aylett, P. W. Mcowan, Q. Mary, and A. Paiva, “Affect Recognition for Interactive Companions,” 2009.
- [25] S. D’Mello, R. Picard, and A. Graesser, “Toward an Affect-Sensitive AutoTutor,” *IEEE Intelligent Systems*, vol. 22, no. 4, pp. 53–61, Jul. 2007.
- [26] E. Hudlicka, “Affective Game Engines: Motivation and Requirements,” in *Proceedings of the 4th International Conference on Foundations of Digital Games*, ser. FDG ’09. New York, NY, USA: ACM, 2009, pp. 299–306.
- [27] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, “It’s all in the game: Towards an affect sensitive and context aware game companion,” in *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009*, Sep. 2009, pp. 1–8.
- [28] C. Holmgard, G. Yannakakis, K.-I. Karstoft, and H. Andersen, “Stress Detection for PTSD via the StartleMart Game,” in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*, Sep. 2013, pp. 523–528.
- [29] A. Paiva, M. Costa, R. Chaves, M. Piedade, D. Mourão, D. Sobral, K. Höök, G. Andersson, and A. Bullock, “SenToy : An Affective Sympathetic Interface,” *Int. J. Hum.-Comput. Stud.*, vol. 59, no. 1-2, pp. 227–235, Jul. 2003.
- [30] E. Hudlicka, C. L. Lisetti, D. Hodge, A. Paiva, A. A. Rizzo, and E. Wagner, “Panel on Artificial Agents for Psychotherapy.” 2008.



- [31] J. Scheirer, R. Fernandez, and R. W. Picard, “Expression Glasses: A Wearable Device for Facial Expression Recognition,” in *CHI '99 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '99. New York, NY, USA: ACM, 1999, pp. 262–263.
- [32] R. el Kaliouby, R. Picard, and S. Baron-Cohen, “Affective Computing and Autism,” *Annals of the New York Academy of Sciences*, vol. 1093, no. 1, pp. 228–248, Dec. 2006.
- [33] F. Doctor, R. Iqbal, and V. Zamudio, “Introduction to the thematic issue on Affect Aware Ubiquitous Computing,” *Journal of Ambient Intelligence and Smart Environments*, vol. 7, no. 1, pp. 3–4, Jan. 2015.
- [34] C. Brod, *Technostress : The Human Cost of the Computer Revolution*. Addison-Wesley, 1984.
- [35] R. Riedl, H. Kindermann, A. Auinger, and A. Javor, “Computer Breakdown as a Stress Factor during Task Completion under Time Pressure: Identifying Gender Differences Based on Skin Conductance,” *Advances in Human-Computer Interaction*, vol. 2013, Oct. 2013.
- [36] ———, “Technostress from a Neurobiological Perspective - System Breakdown Increases the Stress Hormone Cortisol in Computer Users,” *Business & Information Systems Engineering*, vol. 4, no. 2, pp. 61–69, Mar. 2012, tODO.
- [37] A. Jaimes, N. Sebe, and D. Gatica-Perez, “Human-centered Computing: A Multimedia Perspective,” in *Proceedings of the 14th ACM International Conference on Multimedia*, ser. MM '06. New York, NY, USA: ACM, 2006, pp. 855–864.
- [38] K. Vredenberg, S. Isensee, and C. Righi, *User-Centered Design: An Integrated Approach*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

- [39] R. Riedl, “On the Biology of Technostress: Literature Review and Research Agenda,” *SIGMIS Database*, vol. 44, no. 1, pp. 18–55, Nov. 2012.
- [40] D. P. Saha, T. L. Martin, and R. Knapp, “Towards incorporating affective feedback into context-aware intelligent environments,” in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, Sep. 2015, pp. 49–55.
- [41] D. P. Saha, R. B. Knapp, and T. L. Martin, “Affective Feedback in a Virtual Reality Based Intelligent Supermarket,” in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, ser. UbiComp ’17. New York, NY, USA: ACM, 2017, pp. 646–653.
- [42] D. P. Saha, T. L. Martin, and R. B. Knapp, “Towards Defining a Quality-Metric for Affective Feedback in an Intelligent Environment,” in *IEEE International Conference on Pervasive Computing and Communications Workshops, 2018. PerCom Workshops’18*, Athens, Greece, Mar. 2018.
- [43] Z. Hammal and M. Teodosia Suarez, “Towards Context Based Affective Computing,” in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*, Sep. 2013, pp. 802–802.
- [44] T. Erickson, “Some problems with the notion of context-aware computing,” *Communications of the ACM*, vol. 45, no. 2, Feb. 2002.
- [45] D. C. Holzman, “What’s in a Color? The Unique Human Health Effects of Blue Light,” *Environmental Health Perspectives*, vol. 118, no. 1, pp. A22–A27, Jan. 2010.

- [46] H. Ishii and B. Ullmer, “Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms,” in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '97. New York, NY, USA: ACM, 1997, pp. 234–241.
- [47] J. Rekimoto, “Pick-and-drop : A Direct Manipulation Technique for Multiple Computer Environments,” in *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '97. New York, NY, USA: ACM, 1997, pp. 31–39.
- [48] M. Mozer, “The neural network house: An environment that adapts to its inhabitants,” *Proceedings of the American Association for Artificial Intelligence*, 1998.
- [49] S. Brave and C. Nass, “Emotion in human-computer interaction,” in *The Human-Computer Interaction Handbook*, J. A. Jacko and A. Sears, Eds. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 2003, pp. 81–96.
- [50] B. Bortz, I. Bradburn, A. Scrapa, M. Bell, D. P. Saha, D. Swain, and R. B. Knapp, “Emotion Regulation in Children with Autism through Music-based Biofeedback,” in *4th International Conference on Music and Emotion (ICME4)*, Geneva, Switzerland, 2015.
- [51] R. W. Picard, “Affective Computing: From Laughter to IEEE,” *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 11–17, Jan. 2010.
- [52] T. Partala and V. Surakka, “The effects of affective interventions in human–computer interaction,” *Interacting with Computers*, vol. 16, no. 2, pp. 295–309, Apr. 2004.
- [53] A. Öhman, A. Flykt, and F. Esteves, “Emotion drives attention: Detecting the snake in the grass,” *Journal of Experimental Psychology: General*, vol. 130, no. 3, pp. 466–478, 2001.

- [54] P. Vuilleumier and S. Schwartz, “Beware and be aware: Capture of spatial attention by fear-related stimuli in neglect,” *Neuroreport*, vol. 12, no. 6, pp. 1119–1122, May 2001.
- [55] A. K. Anderson and E. A. Phelps, “Lesions of the human amygdala impair enhanced perception of emotionally salient events,” *Nature*, vol. 411, no. 6835, pp. 305–309, May 2001.
- [56] D. Reisberg and P. Hertel, *Memory and Emotion*. Oxford University Press, Nov. 2003.
- [57] J. Klein, Y. Moon, and R. W. Picard, “This computer responds to user frustration: Theory, design, and results,” *Interacting with Computers*, vol. 14, no. 2, pp. 119–140, Jan. 2002.
- [58] J. A. Sánchez, O. Starostenko, P. Medina, O. Cervantes, and W. Wan, “Affordable Development of Animated Avatars for Conveying Emotion in Intelligent Environments,” in *Workshop Proceedings of the 10th International Conference on Intelligent Environments, Shanghai, China, June 30 - July 1, 2014*, 2014, pp. 80–91.
- [59] C.-Y. Chen, J. Forlizzi, and P. Jennings, “ComSlipper : An Expressive Design to Support Awareness and Availability,” in *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '06. New York, NY, USA: ACM, 2006, pp. 369–374.
- [60] W. Sanchez, A. Martinez, W. Campos, H. Estrada, and V. Pelechano, “Inferring Loneliness Levels in Older Adults from Smartphones,” *J. Ambient Intell. Smart Environ.*, vol. 7, no. 1, pp. 85–98, Jan. 2015.
- [61] W. Campos, A. Martinez, H. Estrada, and N. Castro, “Measuring Social Isolation in Older Adults through Ambient Intelligence and Social Networking Sites,” in *Work-*

- shop Proceedings of the 10th International Conference on Intelligent Environments, Shanghai, China, June 30 - July 1, 2014*, 2014, pp. 48–59.
- [62] J. Navarro, V. Zamudio, F. Doctor, C. L. Ramírez, R. Baltazar, C. Martínez, J. Torres, L. Bermúdez, and B. Gutiérrez, “Game Based Monitoring and Cognitive Therapy for Elderly.” in *Intelligent Environments (Workshops)*, ser. Ambient Intelligence and Smart Environments, J. A. Botía and D. Charitos, Eds., vol. 17. IOS Press, 2013, pp. 116–127.
- [63] N. Doulamis, “An adaptable emotionally rich pervasive computing system,” in *Signal Processing Conference, 2006 14th European*, Sep. 2006, pp. 1–5, tODO.
- [64] S. Moncrieff, S. Venkatesh, G. West, and S. Greenhill, “Incorporating Contextual Audio for an Actively Anxious Smart Home,” in *Proceedings of the 2005 International Conference on Intelligent Sensors, Sensor Networks and Information Processing Conference, 2005*, Dec. 2005, pp. 373–378.
- [65] N. V. Jungum and E. Laurent, “Emotions in Pervasive Computing Environments,” *arXiv:0912.1810 [cs]*, Dec. 2009, iMP TODO.
- [66] E. Hudlicka and M. D. McNeese, “Assessment of User Affective and Belief States for Interface Adaptation: Application to an Air Force Pilot Task,” *User Modeling and User-Adapted Interaction*, vol. 12, no. 1, pp. 1–47, Feb. 2002.
- [67] K.-I. Benta, C. Marcel, and T. Valeriu, “Towards an Affective Aware Home,” in *Ambient Assistive Health and Wellness Management in the Heart of the City*, ser. Lecture Notes in Computer Science, M. Mokhtari, I. Khalil, J. Bauchet, D. Zhang, and C. Nugent, Eds. Springer Berlin Heidelberg, Jan. 2009, no. 5597, pp. 74–81.

- [68] J. Broekens, “Emotion and Reinforcement: Affective Facial Expressions Facilitate Robot Learning,” in *Artificial Intelligence for Human Computing*, ser. Lecture Notes in Computer Science, T. S. Huang, A. Nijholt, M. Pantic, and A. Pentland, Eds. Springer Berlin Heidelberg, 2007, no. 4451, pp. 113–132.
- [69] P. M. Niedenthal, S. Krauth-Gruber, and F. Ric, *Psychology of Emotion: Interpersonal, Experiential, and Cognitive Approaches*. New York, NY, US: Psychology Press, 2006, vol. xi.
- [70] B. J. Peters, N. C. Overall, and J. P. Jamieson, “Physiological and cognitive consequences of suppressing and expressing emotion in dyadic interactions,” *International Journal of Psychophysiology*, vol. 94, no. 1, pp. 100–107, Oct. 2014.
- [71] S. H. Fairclough, “Fundamentals of physiological computing,” *Interacting with Computers*, vol. 21, no. 1-2, pp. 133–145, Jan. 2009.
- [72] L. Maat and M. Pantic, “Gaze-X: Adaptive Affective Multimodal Interface for Single-user Office Scenarios,” in *Proceedings of the 8th International Conference on Multimodal Interfaces*, ser. ICMI ’06. New York, NY, USA: ACM, 2006, pp. 171–178.
- [73] J. Hernandez, P. Paredes, A. Roseway, and M. Czerwinski, “Under Pressure: Sensing Stress of Computer Users,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 51–60.
- [74] Y. Shi, M. H. Nguyen, P. Blitz, B. French, S. Fisk, O. D. L. Torre, A. Smailagic, D. P. Siewiorek, M. A. Absi, E. Ertin, T. Kamarck, and S. Kumar, “Personalized Stress Detection from Physiological Measurements,” in *2010 International Symposium on Quality of Life Technology*, 2010, tODO.

- [75] J. Hernandez, R. R. Morris, and R. W. Picard, “Call Center Stress Recognition with Person-Specific Models,” in *Affective Computing and Intelligent Interaction*, ser. Lecture Notes in Computer Science, S. D’Mello, A. Graesser, B. Schuller, and J.-C. Martin, Eds. Springer Berlin Heidelberg, 2011, pp. 125–134.
- [76] K. Plarre, A. Rajj, S. Hossain, A. Ali, M. Nakajima, M. al’Absi, E. Ertin, T. Kamarck, S. Kumar, M. Scott, D. Siewiorek, A. Smailagic, and L. Wittmers, “Continuous inference of psychological stress from sensory measurements collected in the natural environment,” in *2011 10th International Conference on Information Processing in Sensor Networks (IPSN)*, Apr. 2011, pp. 97–108.
- [77] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Troster, and U. Ehlert, “Discriminating Stress From Cognitive Load Using a Wearable EDA Device,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 410–417, Mar. 2010.
- [78] C. Ramos, “A Survey on the use of Emotions, Mood, and Personality in Ambient Intelligence and Smart Environments,” in *Handbook of Research on Ambient Intelligence and Smart Environments Trends and Perspectives*, 2011, pp. 88–107.
- [79] G. Chittaranjan, J. Blom, and D. Gatica-perez, “Who’s who with Big-Five: Analyzing and Classifying personality traits with smartphones,” in *In Proc. Int’l Symposium on Wearable Computers (ISWC)*, 2011.
- [80] G. Chittaranjan, J. Blom, and D. Gatica-Perez, “Mining large-scale smartphone data for personality studies,” *Personal and Ubiquitous Computing*, vol. 17, no. 3, pp. 433–450, Mar. 2013.
- [81] B. Schilit, N. Adams, and R. Want, “Context-Aware Computing Applications,” in *Proceedings of the 1994 First Workshop on Mobile Computing Systems and Applications*, ser. WMCSA ’94. Washington, DC, USA: IEEE Computer Society, 1994, pp. 85–90.

- [82] M. Baldauf, S. Dustdar, and F. Rosenberg, “A Survey on Context-Aware Systems,” *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 2, no. 4, pp. 263–277, Jun. 2007.
- [83] A. K. Dey, “Context-Aware Computing: The CyberDesk Project. AAAI 1998 Spring Symposium on Intelligent Environments,” Tech. Rep., 1998.
- [84] P. Prekop and M. Burnett, “Activities, context and ubiquitous computing,” *Computer Communications*, vol. 26, no. 11, pp. 1168–1176, Jul. 2003.
- [85] T. Hofer, M. Pichler, G. Leonhartsberger, J. Altmann, and W. Retschitzegger, “Context-awareness on mobile devices – the hydrogen approach,” in *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, 2002, pp. 292–302.
- [86] C. Burghardt and T. Kirste, “Inferring Intentions in Generic Context-aware Systems,” in *Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia*, ser. MUM '07. New York, NY, USA: ACM, 2007, pp. 50–54.
- [87] F. Sadri, “Logic-Based Approaches to Intention Recognition,” in *Handbook of Research on Ambient Intelligence and Smart Environments Trends and Perspectives*, 2011, pp. 346–375.
- [88] P. A. Jaques and R. M. Vicari, “A BDI approach to infer student’s emotions in an intelligent learning environment,” *Computers & Education*, vol. 49, no. 2, pp. 360–384, Sep. 2007.
- [89] A. Vinciarelli, M. Pantic, and H. Bourlard, “Social signal processing: Survey of an emerging domain,” *Image and Vision Computing*, vol. 27, no. 12, pp. 1743–1759, Nov. 2009.



- [90] A. Ortony, G. Clore, and A. Collins, “The Cognitive Structure of Emotion,” in *Cognitive Structure of Emotions*. Cambridge University Press, 1988.
- [91] A. Ortony, “On making believable emotional agents believable,” in *Emotions in Humans and Artifacts*. The MIT Press, 2002, pp. 189–211.
- [92] S. C. Marsella and J. Gratch, “EMA: A process model of appraisal dynamics,” *Cognitive Systems Research*, vol. 10, no. 1, pp. 70–90, Mar. 2009.
- [93] S. Marsella, J. Gratch, N. Wang, and B. Stankovic, “Assessing the validity of a computational model of emotional coping,” in *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009*, Sep. 2009, pp. 1–8.
- [94] S. Marsella and J. Gratch, “A Step Toward Irrationality: Using Emotion to Change Belief,” in *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 1*, ser. AAMAS '02. New York, NY, USA: ACM, 2002, pp. 334–341.
- [95] A. DeLongis and S. Holtzman, “Coping in Context: The Role of Stress, Social Support, and Personality in Coping,” *Journal of Personality*, vol. 73, no. 6, pp. 1633–1656, Dec. 2005.
- [96] G. S. Everly and J. M. Lating, “The Anatomy and Physiology of the Human Stress Response,” in *A Clinical Guide to the Treatment of the Human Stress Response*. Springer New York, 2013, pp. 17–51.
- [97] R. S. Lazarus and S. Folkman, *Stress, Appraisal, and Coping*. Springer publishing company, 1984.
- [98] M. Eysenck, *Fundamentals of Psychology*. Psychology Press, 2014.

- [99] A. Monat and R. S. Lazarus, “Introduction : Stress and coping—some current issues and controversies,” in *Stress and Coping: An Anthology*, 1991, pp. 1–15.
- [100] R. A. Hudiburg, “Psychology of Computer Use: Xvii The Computer Technology Hassles Scale: Revision, Reliability, and Some Correlates,” *Psychological Reports*, vol. 65, no. 3f, pp. 1387–1394, Dec. 1989.
- [101] —, “Psychology of Computer Use: Vii. Measuring Technostress: Computer-Related Stress,” *Psychological Reports*, vol. 64, no. 3, pp. 767–772, Jun. 1989.
- [102] B. B. Arnetz and M. Berg, “Techno-Stress : Psycho-physiological Consequences of Poor Man-Machine Interface,” *ADVANCES IN HUMAN FACTORS ERGONOMICS*, vol. 19, pp. 891–891, 1993.
- [103] H. H. Emurian, “Physiological responses during data retrieval: Comparison of constant and variable system response times,” *Computers in Human Behavior*, vol. 7, no. 4, pp. 291–310, 1991.
- [104] W. Kuhmann, W. Boucsein, F. Schaffer, and J. Alexander, “Experimental investigation of psychophysiological stress-reactions induced by different system response times in human-computer interaction\*,” *Ergonomics*, vol. 30, no. 6, pp. 933–943, Jun. 1987.
- [105] D. Lunn and S. Harper, “Using Galvanic Skin Response Measures to Identify Areas of Frustration for Older Web 2.0 Users,” in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, ser. W4A '10. New York, NY, USA: ACM, 2010, pp. 34:1–34:10.
- [106] B. L. Steward, “Physiology of Stress,” in *Managing Stress: Principles and Strategies for Health and Wellbeing*, 5th ed., 2009.
- [107] R. Hudiberg, “Assessing and Managing Technostress,” New York, Aug. 1996.

- [108] P. Muter, J. J. Furedy, A. Vincent, and T. Pelcowitz, “User-hostile systems and patterns of psychophysiological activity,” *Computers in Human Behavior*, vol. 9, no. 1, pp. 105–111, 1993.
- [109] —, “User-hostile systems and patterns of psychophysiological activity: Erratum,” *Computers in Human Behavior*, vol. 10, no. 1, p. 173, 1994.
- [110] H. H. Emurian, “Human-computer interactions: Are there adverse health consequences?” *Computers in Human Behavior*, vol. 5, no. 4, pp. 265–275, 1989.
- [111] W. Boucsein, “Forty Years of Research on System Response Times – What Did We Learn from It?” in *Industrial Engineering and Ergonomics*, C. M. Schlick, Ed. Springer Berlin Heidelberg, 2009, pp. 575–593.
- [112] A. Crider, “Personality and Electrodermal Response Lability: An Interpretation,” *Applied Psychophysiology and Biofeedback*, vol. 33, no. 3, pp. 141–148, May 2008, tODO.
- [113] M. E. Dawson, A. M. Schell, and D. L. Filion, “The electrodermal system,” in *Handbook of Psychophysiology, 3rd Ed*, J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, Eds. New York, NY, US: Cambridge University Press, 2007, pp. 159–181.
- [114] L. R. Stroud, P. Salovey, and E. S. Epel, “Sex differences in stress responses: Social rejection versus achievement stress,” *Biological Psychiatry*, vol. 52, no. 4, pp. 318–327, Aug. 2002.
- [115] S. E. Taylor, L. C. Klein, B. P. Lewis, T. L. Gruenewald, A. R., and J. A. Updegraff, “Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight,” *Psychological Review*, vol. 107, no. 3, pp. 411–429, 2000.

- [116] R. M. Eisler and J. R. Skidmore, "Masculine Gender Role Stress Scale Development and Component Factors in the Appraisal of Stressful Situations," *Behavior Modification*, vol. 11, no. 2, pp. 123–136, Jan. 1987.
- [117] B. L. Gillespie and R. M. Eisler, "Development of the Feminine Gender Role Stress Scale A Cognitive-Behavioral Measure of Stress, Appraisal, and Coping for Women," *Behavior Modification*, vol. 16, no. 3, pp. 426–438, Jan. 1992.
- [118] B. M. Kudielka, D. Hellhammer, and C. Kirschbaum, "Sex differences in human stress response," in *Encyclopedia of Stress*. Academic Press San Diego, 2000, vol. 3.
- [119] B. M. Kudielka and C. Kirschbaum, "Sex differences in HPA axis responses to stress: A review," *Biological Psychology*, vol. 69, no. 1, pp. 113–132, Apr. 2005.
- [120] E. Stenstrom, P. Stenstrom, G. Saad, and S. Cheikhrouhou, "Online Hunting and Gathering: An Evolutionary Perspective on Sex Differences in Website Preferences and Navigation," *IEEE Transactions on Professional Communication*, vol. 51, no. 2, pp. 155–168, Jun. 2008.
- [121] R. Riedl, M. Hubert, and P. Kenning, "Are There Neural Gender Differences in Online Trust? An fMRI Study on the Perceived Trustworthiness of Ebay Offers," *MIS Q.*, vol. 34, no. 2, pp. 397–428, Jun. 2010.
- [122] M. Tarafdar, Q. Tu, T. S. Ragu-Nathan, and B. S. Ragu-Nathan, "Crossing to the Dark Side: Examining Creators, Outcomes, and Inhibitors of Technostress," *Commun. ACM*, vol. 54, no. 9, pp. 113–120, Sep. 2011.
- [123] V. B. Elder, E. P. Gardner, and S. R. Ruth, "Gender and age in technostress: Effects on white collar productivity: [computer anxiety among government finance office workers]," *Government finance review.*, vol. 3, no. 6, 1987.

- [124] S. C. Srivastava, S. Chandra, and A. Shirish, “Technostress creators and job outcomes: Theorising the moderating influence of personality traits,” *Information Systems Journal*, vol. 25, no. 4, pp. 355–401, Jul. 2015.
- [125] P. Galluch, “It’s All In Your Personality: Combatting Technostress In The Workplace,” *SAIS 2015 Proceedings*, Jan. 2015.
- [126] D. Sharma and T. K. Gill, “Technostress and Personality Traits – Are they Associated? – Evidence from Indian Bankers,” *International Journal of Computer Science and Technology*, vol. 7, no. 1, pp. 106–111, 2016.
- [127] V. Moreland, “Technostress and Personality Type,” *Online*, vol. 17, no. 4, pp. 59–62, Jul. 1993.
- [128] K. C. Gunthert, L. H. Cohen, and S. Armeli, “The role of neuroticism in daily stress and coping,” *Journal of Personality and Social Psychology*, vol. 77, no. 5, pp. 1087–1100, 1999.
- [129] R. Calvo and S. D’Mello, “Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications,” *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, Jan. 2010.
- [130] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe, “A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments,” *Psychophysiology*, vol. 49, pp. 1017–1034, 2013.
- [131] S. D. Kreibig, “Autonomic nervous system activity in emotion: A review,” *Biological Psychology*, vol. 84, no. 3, pp. 394–421, Jul. 2010.
- [132] A. Colomer Granero, F. Fuentes-Hurtado, V. Naranjo Ornedo, J. Guixeres Provinciale, J. M. Ausín, and M. Alcañiz Raya, “A Comparison of Physiological Signal Analysis

- Techniques and Classifiers for Automatic Emotional Evaluation of Audiovisual Contents,” *Frontiers in Computational Neuroscience*, p. 74, 2016.
- [133] S. K. Yoo, C. K. Lee, Y. J. Park, N. H. Kim, B. C. Lee, and K. S. Jeong, “Neural Network Based Emotion Estimation Using Heart Rate Variability and Skin Resistance,” in *Advances in Natural Computation*, ser. Lecture Notes in Computer Science, L. Wang, K. Chen, and Y. S. Ong, Eds. Springer Berlin Heidelberg, Aug. 2005, no. 3610, pp. 818–824.
- [134] C. W. Darrow, “The Rationale for Treating the Change in Galvanic Skin Response as a Change in Conductance,” *Psychophysiology*, vol. 1, no. 1, pp. 31–38, 1964.
- [135] W. Boucsein, *Electrodermal Activity*. Springer US, Jan. 2012.
- [136] M. Benedek and C. Kaernbach, “Decomposition of skin conductance data by means of nonnegative deconvolution,” *Psychophysiology*, vol. 47, no. 4, pp. 647–658, Jul. 2010.
- [137] D. T. Lykken and P. H. Venables, “Direct Measurement of Skin Conductance: A Proposal for Standardization,” *Psychophysiology*, vol. 8, no. 5, pp. 656–672, Sep. 1971.
- [138] W. Boucsein, D. C. Fowles, S. Grimnes, W. Roth, M. E. Dawson, and D. Filion, “Publication Recommendations for Electrodermal Measurements,” *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, 2012.
- [139] M.-Z. Poh, N. Swenson, and R. Picard, “A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 5, pp. 1243–1252, May 2010.
- [140] M. van Dooren, J. J. G. G.-J. de Vries, and J. H. Janssen, “Emotional sweating across the body: Comparing 16 different skin conductance measurement locations,” *Physiology & Behavior*, vol. 106, no. 2, pp. 298–304, May 2012.

- [141] R. Tartz, A. Vartak, J. King, and D. Fowles, “Effects of grip force on skin conductance measured from a handheld device,” *Psychophysiology*, vol. 52, no. 1, pp. 8–19, Jan. 2015.
- [142] M. Světlák, P. Bob, R. Roman, S. Ježek, A. Damborská, J. Chládek, D. J. Shaw, and M. Kukleta, “Stress-induced alterations of left-right electrodermal activity coupling indexed by pointwise transinformation,” *Physiological Research / Academia Scientiarum Bohemoslovaca*, vol. 62, no. 6, pp. 711–719, 2013.
- [143] B. Figner and R. Murphy, “Using skin conductance in judgment and decision making research,” in *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User’s Guide*. Psychology Press, May 2011.
- [144] J. Kim and E. Andre, “Emotion recognition based on physiological changes in music listening,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.
- [145] G. E. Billman, “Heart Rate Variability – A Historical Perspective,” *Frontiers in Physiology*, vol. 2, Nov. 2011.
- [146] R. C. Melo, M. D. B. Santos, E. Silva, R. J. Quitério, M. A. Moreno, M. S. Reis, I. A. Verzola, L. Oliveira, L. E. B. Martins, L. Gallo-Junior, and A. M. Catai, “Effects of age and physical activity on the autonomic control of heart rate in healthy men,” *Brazilian Journal of Medical and Biological Research*, vol. 38, no. 9, pp. 1331–1338, Sep. 2005.
- [147] E. A. Butler, F. H. Wilhelm, and J. J. Gross, “Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction,” *Psychophysiology*, vol. 43, no. 6, pp. 612–622, Nov. 2006.

- [148] P. Nickel and F. Nachreiner, “Sensitivity and Diagnosticity of the 0.1-Hz Component of Heart Rate Variability as an Indicator of Mental Workload,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 45, no. 4, pp. 575–590, Jan. 2003.
- [149] P. Jönsson, “Respiratory sinus arrhythmia as a function of state anxiety in healthy individuals,” *International Journal of Psychophysiology*, vol. 63, no. 1, pp. 48–54, Jan. 2007.
- [150] G. Parati, G. Mancia, M. Di Rienzo, and P. Castiglioni, “Point : Cardiovascular variability is/is not an index of autonomic control of circulation,” *Journal of Applied Physiology (Bethesda, Md.: 1985)*, vol. 101, no. 2, pp. 676–678; discussion 681–682, Aug. 2006.
- [151] G. S. Everly and J. M. Lating, “Measurement of the Human Stress Response,” in *A Clinical Guide to the Treatment of the Human Stress Response*. Springer New York, 2013, pp. 103–126.
- [152] J. Jaimovich, “Emotion recognition from physiological indicators for musical applications,” Ph.D., Queen’s University Belfast, 2013.
- [153] G. G. Berntson, K. S. Quigley, and D. Lozano, “Cardiovascular psychophysiology,” *Handbook of psychophysiology*, vol. 3, pp. 182–210, 2007.
- [154] Z. Zhang, Z. Pi, and B. Liu, “TROIKA : A General Framework for Heart Rate Monitoring Using Wrist-Type Photoplethysmographic (PPG) Signals During Intensive Physical Exercise,” *IEEE Transactions on Biomedical Engineering*, vol. PP, no. 99, pp. 1–1, 2014.



- [155] R. Yousefi, M. Nourani, S. Ostadabbas, and I. Panahi, "A Motion-Tolerant Adaptive Algorithm for Wearable Photoplethysmographic Biosensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 2, pp. 670–681, Mar. 2014.
- [156] F. Peng, Z. Zhang, X. Gou, H. Liu, and W. Wang, "Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter," *BioMedical Engineering OnLine*, vol. 13, no. 1, p. 50, Apr. 2014.
- [157] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.
- [158] M. Brennan, M. Palaniswami, and P. Kamen, "Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 283, no. 5, pp. H1873–H1886, Nov. 2002.
- [159] G. E. Billman, "The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance," *Clinical and Translational Physiology*, p. 26, 2013.
- [160] R. A. Hoshi, C. M. Pastre, L. C. M. Vanderlei, and M. F. Godoy, "Poincaré plot indexes of heart rate variability: Relationships with other nonlinear variables," *Autonomic Neuroscience: Basic and Clinical*, vol. 177, no. 2, pp. 271–274, Oct. 2013.
- [161] C. K. Karmakar, A. H. Khandoker, A. Voss, and M. Palaniswami, "Sensitivity of temporal heart rate variability in Poincaré plot to changes in parasympathetic nervous system activity," *BioMedical Engineering OnLine*, vol. 10, p. 17, 2011.

- [162] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-aho, and P. A. Karjalainen, “Kubios HRV – Heart rate variability analysis software,” *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 210–220, Jan. 2014.
- [163] H. Samet, *Foundations of Multidimensional and Metric Data Structures (The Morgan Kaufmann Series in Computer Graphics and Geometric Modeling)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.
- [164] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY: Springer New York, 2009.
- [165] A. H. H. Ngu, Q. Z. Sheng, D. Q. Huynh, and R. Lei, “Combining Multi-visual Features for Efficient Indexing in a Large Image Database,” *The VLDB Journal*, vol. 9, no. 4, pp. 279–293, Apr. 2001.
- [166] A. Janecek, W. N. Gansterer, M. Demel, and G. Ecker, “On the Relationship Between Feature Selection and Classification Accuracy.” in *FSDM*, 2008, pp. 90–105.
- [167] S. Theodoridis and K. Koutroumbas, *Pattern Recognition, Fourth Edition*, 4th ed. Academic Press, 2008.
- [168] M. J. Zaki and W. Meira Jr, *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press, 2014.
- [169] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [170] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu, “Boosting for Transfer Learning,” in *Proceedings of the 24th International Conference on Machine Learning*, ser. ICML ’07. New York, NY, USA: ACM, 2007, pp. 193–200.

- [171] P. Wu and T. G. Dietterich, “Improving SVM accuracy by training on auxiliary data sources,” in *In ICML*, 2004, pp. 871–878.
- [172] H. S. Alavi, E. Churchill, D. Kirk, J. Nembrini, and D. Lalanne, “Deconstructing Human-building Interaction,” *interactions*, vol. 23, no. 6, pp. 60–62, Oct. 2016.
- [173] P. O. Fanger, “Thermal comfort. Analysis and applications in environmental engineering.” *Thermal comfort. Analysis and applications in environmental engineering.*, 1970.
- [174] C.-C. J. Huang, R. Yang, and M. W. Newman, “The Potential and Challenges of Inferring Thermal Comfort at Home Using Commodity Sensors,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp ’15. New York, NY, USA: ACM, 2015, pp. 1089–1100.
- [175] K. Knecht, N. Bryan-Kinns, and K. Shoop, “Usability and Design of Personal Wearable and Portable Devices for Thermal Comfort in Shared Work Environments,” in *Proceedings of the 30th International BCS Human Computer Interaction Conference: Fusion!*, ser. HCI ’16. Swindon, UK: BCS Learning & Development Ltd., 2016, pp. 41:1–41:12.
- [176] J. Ranjan and J. Scott, “ThermalSense: Determining Dynamic Thermal Comfort Preferences Using Thermographic Imaging,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp ’16. New York, NY, USA: ACM, 2016, pp. 1212–1222.
- [177] S. Karjalainen, “Should it be automatic or manual—The occupant’s perspective on the design of domestic control systems,” *Energy and Buildings*, vol. 65, no. Supplement C, pp. 119–126, Oct. 2013.

- [178] S. Mennicken, O. Zihler, F. Juldaschewa, V. Molnar, D. Aggeler, and E. M. Huang, “‘It’s Like Living with a Friendly Stranger’: Perceptions of Personality Traits in a Smart Home,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp ’16. ACM, 2016, pp. 120–131.
- [179] M. H. Coen, “Design Principles for Intelligent Environments,” in *Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence*, ser. AAAI ’98/IAAI ’98. Menlo Park, CA, USA: American Association for Artificial Intelligence, 1998, pp. 547–554.
- [180] B. Shneiderman and P. Maes, “Direct Manipulation vs. Interface Agents,” *interactions*, vol. 4, no. 6, pp. 42–61, Nov. 1997.
- [181] C. Creed, “Using Emotion Simulation to Influence User Attitudes and Behaviour,” Ph.D. dissertation, 2005.
- [182] M.-L. Bourguet, “Towards a taxonomy of error-handling strategies in recognition-based multi-modal human–computer interfaces,” *Signal Processing*, vol. 86, no. 12, pp. 3625–3643, Dec. 2006.
- [183] D. Alba, “It’s Your Fault Microsoft’s Teen AI Turned Into Such a Jerk,” Mar. 2016.
- [184] P. Lee, “Learning from Tay’s introduction,” Mar. 2016.
- [185] Z. Zeng, M. Pantic, G. Roisman, and T. Huang, “A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 39–58, Jan. 2009.
- [186] B. Cowley, M. Filetti, K. Lukander, J. Torniainen, A. Henelius, L. Ahonen, O. Baral, I. Kosunen, T. Valtonen, M. Huotilainen, N. Ravaja, and G. Jacucci, “The Psychophysiology Primer: A Guide to Methods and a Broad Review with a Focus on

- Human–Computer Interaction,” *Foundations and Trends® Human–Computer Interaction*, vol. 9, no. 3-4, pp. 151–308, Nov. 2016.
- [187] J. Zhai and A. Barreto, “Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables,” in *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2006. EMBS '06*, 2006, pp. 1355–1358.
- [188] W. J. Tyler, A. M. Boasso, J. D. Charlesworth, M. A. Marlin, K. Aebersold, L. Aven, D. Z. Wetmore, and S. K. Pal, “Suppression of human psychophysiological and biochemical stress responses using high-frequency pulse-modulated transdermal electrical neurosignaling.” *bioRxiv*, p. 015032, Feb. 2015.
- [189] D. Wu, C. G. Courtney, B. J. Lance, S. S. Narayanan, M. E. Dawson, K. S. Oie, and T. D. Parsons, “Optimal Arousal Identification and Classification for Affective Computing Using Physiological Signals: Virtual Reality Stroop Task,” *IEEE Transactions on Affective Computing*, vol. 1, no. 2, pp. 109–118, Jul. 2010.
- [190] D. P. Saha, B. Bortz, W. Huang, T. Martin, and B. Knapp, “Affect-Aware Intelligent Environment Using Musical Cues as an Emotion Learning Framework,” in *2016 12th International Conference on Intelligent Environments (IE)*, Sep. 2016, pp. 178–181.
- [191] W. Huang, B. Bortz, and R. B. Knapp, “Exploring the causal relationships between musical features and physiological indicators of emotion,” in *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference On*. IEEE, 2015, pp. 560–566.
- [192] B. Bortz, J. Jaimovich, and R. B. Knapp, “Emotion in Motion: A Reimagined Framework for Biomusical/Emotional Interaction,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, ser. NIME 2015. Baton Rouge,

- Louisiana, USA: The School of Music and the Center for Computation and Technology (CCT), Louisiana State University, 2015, pp. 44–49.
- [193] P. Renaud and J.-P. Blondin, “The stress of Stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed,” *International Journal of Psychophysiology*, vol. 27, no. 2, pp. 87–97, Sep. 1997, tODO.
- [194] Y. Hoshikawa and Y. Yamamoto, “Effects of Stroop color-word conflict test on the autonomic nervous system responses,” *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 272, no. 3, pp. H1113–H1121, Mar. 1997.
- [195] D. F. Barbosa, F. J. A. Prada, M. F. Glanner, O. d. T. Nóbrega, and C. Córdova, “Cardiovascular response to Stroop test: Comparison between the computerized and verbal tests,” *Arquivos Brasileiros de Cardiologia*, vol. 94, no. 4, pp. 507–511, Apr. 2010.
- [196] J. R. Stroop, “Studies of interference in serial verbal reactions,” *Journal of Experimental Psychology*, vol. 18, no. 6, pp. 643–662, 1935.
- [197] M. Benedek and C. Kaernbach, “A continuous measure of phasic electrodermal activity,” *Journal of Neuroscience Methods*, vol. 190, no. 1, pp. 80–91, Jun. 2010.
- [198] D. M. Alexander, C. Trengove, P. Johnston, T. Cooper, J. P. August, and E. Gordon, “Separating individual skin conductance responses in a short interstimulus-interval paradigm,” *Journal of Neuroscience Methods*, vol. 146, no. 1, pp. 116–123, Jul. 2005.
- [199] H. Knolle, “Parameter Estimation for the Bateman Function via Moments of the Empirical Curve,” *Biometrical Journal*, vol. 28, no. 1, pp. 23–29, Jan. 1986.
- [200] I. Asimov, *The Roving Mind*. Prometheus Books, 1983.

- [201] A. Guo, S. Raghu, X. Xie, S. Ismail, X. Luo, J. Simoneau, S. Gilliland, H. Baumann, C. Southern, and T. Starner, “A Comparison of Order Picking Assisted by Head-up Display (HUD), Cart-mounted Display (CMD), Light, and Paper Pick List,” in *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, ser. ISWC '14. New York, NY, USA: ACM, 2014, pp. 71–78.
- [202] K. H. Kim, S. W. Bang, and S. R. Kim, “Emotion recognition system using short-term monitoring of physiological signals,” *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 419–427, May 2004.
- [203] J. Zhai, A. Barreto, C. Chin, and C. Li, “Realization of stress detection using psychophysiological signals for improvement of human-computer interactions,” in *IEEE SoutheastCon, 2005. Proceedings*, Apr. 2005, pp. 415–420.
- [204] M. Kubat and S. Matwin, “Addressing the Curse of Imbalanced Training Sets: One-Sided Selection,” in *In Proceedings of the Fourteenth International Conference on Machine Learning*. Morgan Kaufmann, 1997, pp. 179–186.
- [205] M. Kubat, R. C. Holte, and S. Matwin, “Machine Learning for the Detection of Oil Spills in Satellite Radar Images,” *Machine Learning*, vol. 30, no. 2-3, pp. 195–215, Feb. 1998.
- [206] Y. Sasaki, “The truth of the F-measure,” *Teach Tutor mater*, pp. 1–5, 2007.
- [207] W. Wahlster, M. Feld, P. Gebhard, D. Heckmann, R. Jung, M. Kruppa, M. Schmitz, L. Spassova, and R. Wasinger, “The Shopping Experience of Tomorrow: Human-Centered and Resource-Adaptive,” in *Resource-Adaptive Cognitive Processes*, ser. Cognitive Technologies, M. W. Crocker and J. Siekmann, Eds. Springer Berlin Heidelberg, 2010, pp. 205–237.

- [208] N. Lowe's, "Lowe's Introduces In-Store Navigation Using Augmented Reality," <https://newsroom.lowes.com/news-releases/lowesintroducesin-storenavigationusingaugmentedreality/>, Mar. 2017.
- [209] Engadget, "The checkout line's death knell — The Future IRL," <https://www.engadget.com/2017/10/31/the-checkout-line-s-death-knell-the-future-irl/>, 2017.
- [210] S. Lazem and D. Gracanin, "Social Traps in Second Life," in *2010 Second International Conference on Games and Virtual Worlds for Serious Applications*, Mar. 2010, pp. 133–140.
- [211] N. Yee, J. N. Bailenson, M. Urbanek, F. Chang, and D. Merget, "The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments," *CyberPsychology & Behavior*, vol. 10, no. 1, pp. 115–121, Feb. 2007.
- [212] D. Friedman, A. Steed, and M. Slater, "Spatial Social Behavior in Second Life," in *Intelligent Virtual Agents*, ser. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, Sep. 2007, pp. 252–263.
- [213] E. Castronova, D. Williams, Cuihua Shen, R. Ratan, Li Xiong, Yun Huang, and B. Keegan, "As real as real? Macroeconomic behavior in a large-scale virtual world," *New Media & Society*, vol. 11, no. 5, pp. 685–707, Aug. 2009.
- [214] B. U. Cowley and J. Torniainen, "A short review and primer on electrodermal activity in human computer interaction applications," *arXiv:1608.06986 [cs]*, Aug. 2016.
- [215] J. Wijsman, B. Grundlehner, J. Penders, and H. Hermens, "Trapezius Muscle EMG As Predictor of Mental Stress," in *Wireless Health 2010*, ser. WH '10. New York, NY, USA: ACM, 2010, pp. 155–163.



- [216] D. Cardone and A. Merla, “New Frontiers for Applications of Thermal Infrared Imaging Devices: Computational Psychophysiology in the Neurosciences,” *Sensors*, vol. 17, no. 5, p. 1042, May 2017.
- [217] B. Arnrich, C. Setz, R. L. Marca, G. Tröster, and U. Ehlert, “What Does Your Chair Know About Your Stress Level?” *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 207–214, Mar. 2010.
- [218] D. Wu and T. D. Parsons, “Inductive Transfer Learning for Handling Individual Differences in Affective Computing,” in *Affective Computing and Intelligent Interaction*, ser. Lecture Notes in Computer Science, S. D’Mello, A. Graesser, B. Schuller, and J.-C. Martin, Eds. Springer Berlin Heidelberg, 2011, no. 6975, pp. 142–151.
- [219] M. Weiser, “Ubiquitous computing,” *Computer*, vol. 26, no. 10, pp. 71–72, Oct. 1993.
- [220] K. N. Nkurikiyeyezu, Y. Suzuki, and G. F. Lopez, “Heart rate variability as a predictive biomarker of thermal comfort,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–13, Aug. 2017.