IMPACT OF THE INTERACTION BETWEEN THERMAL ENVIRONMENT AND COGNITIVE LOAD ON THERMAL COMFORT, TASK COGNITIVE PERFORMANCE DOMAIN AND STRESS - EXPERIMENTAL AND DATA-DRIVEN APPROACH

ZHANG XUANGE

PhD

The Hong Kong Polytechnic University

2026

The Hong Kong Polytechnic University

Department of Building and Real Estate

Impact of the Interaction between Thermal Environment and Cognitive Load on Thermal Comfort, Task Cognitive Performance Domain and Stress - Experimental and Data-Driven Approach

Zhang Xuange

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

November 2025

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_______________________________ (Signed)

ZHANG XUANGE_______________ (Name of student)

Abstract

Abstract of thesis entitled: Experimental and Data-Driven Analysis of Thermal Comfort, Task Cognitive Performance Domain, and Stress under Thermal and Task Load Interactions

Submitted by : Zhang Xuange

For the degree of :Doctor of Philosophy

At The Hong Kong Polytechnic University in Oct 2025

This thesis investigates how thermal environments and task difficulty jointly shape human thermal sensation, thermal satisfaction, task performance, workload perception, and stress level, and develops a multimodal predictive framework for modelling such interactions. The study responds to a long-standing gap in thermal-comfort research—its limited integration of cognitive and affective processes—by combining controlled experimentation with statistical and machine-learning approaches to reveal the mechanisms of human adaptation under concurrent thermal and mental load.

Because the thermal experience of office occupants is influenced not only by temperature but also by cognitive demand, the research employed a controlled laboratory design with three PMV-defined thermal states (slightly cool, neutral, slightly warm) and three task-difficulty levels (easy, moderate, hard). Thirty-two subjects conducted computer-based cognitive tasks while their subjective ratings, physiological indicators (skin temperature, electrodermal activity, heart rate), and

thermal-facial images were recorded. Statistical analyses using Friedman tests and post-hoc comparisons assessed the significance and interaction of environmental and cognitive effects, whereas a progressive modelling pipeline—ranging from single-target algorithms to multimodal CNN–MLP architectures—translated these empirical findings into predictive form.

The results revealed clear main and interaction effects between thermal environment and task difficulty. Higher task demand increased workload and stress while moderating thermal sensitivity; neutral temperature conditions produced the best overall performance and well-being. Comfort, workload, and stress were strongly correlated, confirming that thermal and cognitive stressors operate through a shared adaptive system. Predictive modelling further demonstrated that task-related variables and physiological responses substantially enhance the explainability of comfort and performance. While early multi-output models captured inter-response relationships with limited precision, the inclusion of physiological parameters greatly improved stability, and the final CNN–MLP multimodal fusion model, incorporating thermal-image features, achieved the highest accuracy and generalisability. This model effectively represented the intertwined variations among comfort, performance, and psychological strain.

The findings make three principal contributions. Theoretically, they extend conventional thermal-comfort theory by establishing thermal–cognitive interaction as a dynamic regulatory process: comfort is not a static physical equilibrium but a context-dependent balance among environmental, cognitive, and affective factors. Methodologically, they demonstrate the feasibility of combining non-parametric statistics with multimodal deep learning to link descriptive and predictive perspectives on human responses. Practically, the study offers a foundation for human-centred intelligent environmental control, enabling adaptive thermal management that aligns energy efficiency with occupant well-being.

In conclusion, this research reframes thermal comfort as a multidimensional, embodied, and cognitively moderated phenomenon. It provides both conceptual and computational evidence that human performance and mental state are inseparable from thermoregulatory adaptation. By bridging environmental science, cognitive psychology, and data-driven analytics, the thesis contributes an integrated understanding of how people experience, evaluate, and adapt to their thermal surroundings—paving the way for adaptive, empathetic, and sustainable built environments.

Acknowledgements

The results of this thesis were supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China.

First of all, I would like to give my sincere appreciation to three greatest women in my thirty-year-life who have had the most profound influence on me, my Chief Supervisor, Prof. Lee Minhyun, my mother, Ms. Wang Hui, and my grandmother, Ms. Li Jingzhi. It was you who taught me how to be an upright person, how to live joyfully, and how to keep learning.

I still remember what my Chief Supervisor told me during our first meeting in April 2022: "Since you have chosen to pursue a PhD, I hope you can learn to be a real scholar." I have always kept this statement in mind. During the three years, Prof. Lee Minhyun guided me not only with her profound academic insight but also with her patience, integrity, and kindness. From the very beginning, she encouraged me to think independently, to defend my ideas with logic and evidence, and to stay humble in the pursuit of truth. Her high standards pushed me to refine every detail of my research—from experimental design and data analysis to the interpretation of results. She inspired me to treat every setback as an opportunity for deeper understanding and never allowed me to lose faith in the value of persistence. I am also deeply grateful to Prof. Lee for her generosity and care beyond academic supervision. She used her own research funding to fully support me, turning me from a self-financed student into a funded member of her research group. This gesture went far beyond financial assistance—it represented her confidence in my potential and her belief that genuine scholarship deserves to be nurtured regardless of one's starting point. Her trust not only relieved the financial pressure that once weighed on me but also strengthened my sense of belonging and responsibility as a young researcher. I will always be indebted to her for this kindness, which has had a lasting influence on both my academic path and personal growth.

I would like to sincerely appreciate Prof. Seo Joonoh for serving as my co-supervisor and for his kind guidance and support during my study. His comments and advice, though concise, were valuable in refining my work and broadening my academic perspective. I am grateful for his time and encouragement.

I would like to appreciate my beloved family, to my father, Mr. Zhang Yiqun, my maternal grandfather, Mr. Wang Liangtian, my aunt, Ms. Wang Ying, and my cousin, Mr. Zhou Zhexi, for their unconditional love and support throughout my life and studies. It is also lovingly dedicated to my late grandparents, whose kindness, wisdom, and strength continue to inspire me every day. Their memory reminds me why perseverance and integrity matter, and their spirits accompany me in every step of this journey.

I would also like to express my heartfelt thanks to all the members of our research group, Miss Cui Xue, Dr. Mohammad Nyme Uddin, Mr. Luo Junjiang, Dr. Huang Lijie, Dr. Zhang Ruixiaoxiao, Mr. Vincent Gbouna Zakka, Mr. Gao Wenjun, Miss Mei Shiqi, and Miss Xu Lili for their kind help, encouragement, and companionship during my doctoral study. Their support in experiments, data collection, and daily research discussions made my work easier and more enjoyable. I truly appreciate the moments we shared, the ideas we exchanged, and the friendship that made these three years both productive and memorable.

I would also like to express my heartfelt gratitude to my dear friends and classmates who have accompanied me from childhood to adulthood, Miss Lu Yawei, Mr. Zhu Yuzhao, Mr. Lin Boyu, Mr. Liu Hongming, Mr. Sun Yu and Mr. Cheng Liming, etc. Although many of them are far away, their friendship, humour, and constant encouragement have always been a source of strength for me. The memories we share remind me of where I started and keep me grounded while pursuing my goals. Their faith in me helped me through difficult times and made this journey much warmer and more meaningful.

Finally, to all the people who have appeared in my life, those I have mentioned here and those I may not have named, thank you for making my world brighter, broader, and more meaningful.

19 October 2025

Table of contents

List of Figures

List of Tables

Nomenclature

Heating, ventilation and air conditioning	HVAC
International Energy Agency	IEA
Indoor environment quality	IEQ
Predicted mean vote	PMV
Analysis of variance	ANOVA
Multivariate regression analysis	MRA
Skin temperature	ST
Electrodermal activity	EDA
Heart rate	HR
NASA Task Load Index	NASA-TLX
State-Trait Anxiety Inventory	STAI
Body mass index	BMI
Thermal sensation	TSENS
Thermal Sensation Vote	TSENSV
Mean radiant temperature	MRT
Thermal preference	TPREF
Thermal Preference Vote	TPREFV
Thermal acceptability	TACT
Thermal Acceptability Vote	TACTV
Thermal satisfaction	TSAT
Thermal Satisfaction Vote	TSATV
Post-occupancy evaluation	POE
Thermal Comfort Vote	TCV
Thermal Expectation	TEXP
Perceived Control	PCON
Overall Comfort	OCOM
Body mass index	BMI
Basal metabolic rate	BMR
Task cognitive–performance domain	TCPD
Core body temperature	T_c
Skin temperature	T_{sk}
Electrodermal activity	EDA

Galvanic skin response	GSR
Heart rate	HR
Blood pressure	BP
Inter-beat interval	IBI
Blood volume pulse	BVP
Heart rate variability	HRV
Low-frequency/high-frequency	LF/HF
Electrocardiography	ECG
Photoplethysmography	PPG
Oxygen uptake	VO_2
Respiratory exchange ratio	RER
Infrared thermography	IRT
Electroencephalograph	EEG
Functional near-infrared spectroscopy	fNIRS
Objective performance index	OPI
Skin conductance responses	SCR
Skin conductance level	SCL
Region of interest	ROI
Logistic Regression	LR
K-Nearest Neighbours	KNN
Random Forest	RF
Multilayer perceptron	MLP
Convolutional neural network	CNN
Area Under the Receiver Operating Characteristic Curve	AUROC
Area Under the Precision-Recall Curve	AUPRC
Interquartile range	IQR

Chapter 1 Introduction

1.1 Background

A comfortable living environment has always been a significant goal pursued by human society. There exists a close link between human health conditions and environmental factors. Human development in modern society dictates that most people spend much time indoors. Because most daily activities, such as working, sleeping, resting, dining, relaxing, and entertaining take place indoors. Even special age groups, such as the elderly, infants, and people with certain illnesses, may live only in indoor environments. According to statistics, modern humans spend 90% of their time living in indoor spaces [1, 2]. Whilst living in an indoor environment insulates us from many of the fluctuations and influences of the outdoor environment, it is worth noting that the man-made environment introduces new risks and affects the health of its occupants in a different way. Therefore, understanding how to design a comfortable and healthy indoor environment is an important guarantee of long-term sustainable human development. On the other hand, unlike outdoor environments, which can be uncontrollable due to location, climate, weather, neighbourhood amenities, etc., indoor environments can be adjusted in various ways to make them more comfortable for occupants. In conclusion, the comfort of the indoor environment is a very important factor in the design and management of buildings in today's society, and it also contributes to the health status of the occupants. As a result, both designers and users of buildings are increasingly concerned with building better indoor environments.

In addition to ensuring satisfactory living conditions, the study of indoor thermal comfort plays a pivotal role in the optimisation of energy consumption within buildings. Heating, ventilation and air conditioning (HVAC) systems are among the primary contributors to a building's overall energy demand, accounting for approximately 40–50% of total building energy use in typical office or commercial buildings [3, 4]. According to the International Energy Agency (IEA), improvements in building design and operation, including optimising thermal comfort, could reduce HVAC-related energy consumption by 20–40% without compromising occupants' well-being [5]. For instance, a study published by the World Business Council for Sustainable Development found that implementing advanced thermal comfort control strategies, such as dynamic setpoint adjustments and personal comfort systems, can lead to energy savings of up to 30% in modern office environments [6]. By deepening our understanding of thermal comfort, it is possible to develop strategies and technologies that maintain acceptable comfort levels while reducing unnecessary energy expenditure. This not only benefits building owners and occupants through lower operational costs, but also contributes to broader environmental goals by decreasing greenhouse gas emissions and supporting sustainable development initiatives.

Furthermore, research into thermal comfort provides the scientific foundation for the development and refinement of building standards and guidelines. National and international organisations increasingly rely on empirical evidence to inform regulations that govern indoor environmental quality. Standards such as ISO 7730 and ASHRAE 55, for example, offer detailed recommendations on acceptable thermal conditions for various indoor settings [7]. By continuously updating these standards in light of emerging research, policymakers and industry professionals can ensure that building environments are designed and operated to meet the evolving expectations of comfort, health and efficiency. This process not only protects occupants' well-being, but also ensures consistency and reliability across the built environment sector.

Thus, the goal of studying indoor thermal comfort is also not only to create a more comfortable and healthier indoor environment for humans. Nicol [8] outlined three objectives for studying the knowledge in the field of thermal comfort:

- To provide occupants with satisfactory living conditions
- To optimise energy consumption
- To inform the development of standards

As a result, a growing number of scholars have begun to pay attention to the field of thermal comfort and regard it as one of the most important factors affecting indoor comfort.

Many aspects such as human behaviour, environmental psychology, thermal environment and thermal comfort, building services, materials science, ergonomics, etc. can influence the comfort of an indoor environment. For occupants of indoor environments, thermal comfort, as one of the environmental factors, is usually more important than auditory and visual comfort and air quality [9]. There are many factors affecting the thermal comfort of an indoor space, the first of which are physical parameters. Physical parameters include ambient air temperature, surface temperature, air humidity, wind speed, radiation intensity, etc., which depend on the indoor space itself. In general, the adjustment of the indoor environment is also the adjustment of these indicators. The occupants' individual factors are also one of the significant determinants of thermal comfort. Individual factors consist of physiological factors including age, sex, races and basal metabolic rate, etc. and psychological factors such as temperature habits, preference, experience, etc. Besides, occupants' physical behaviours, psychological behaviours and clothing insulation can also affect thermal comfort [10]. Therefore, thermal comfort is a perception that represents subjective satisfaction with the current thermal environment, which differs from an objective specific temperature set and depends on the factors mentioned above. As a result, the study of thermal comfort contains several directions including physics, physiology, psychology, mechanical engineering, material science, etc.

In addition to being an environment for human habitation, the indoor environment is also one of the most important environments in which human beings carry out their work and production. Studies have shown that the indoor environment quality (IEQ) has a large impact on job satisfaction and work productivity [11]. In addition, the cost of office personnel is an amount and higher compared to the cost of constructing a work environment. Therefore, improving the work environment is one of the most effective ways to increase work productivity. Previous research has suggested that the indoor thermal environment may affect productivity in ways that go beyond occupant comfort, potentially involving cognitive performance and workload management. Similar to the significant effect of indoor thermal environments on comfort, the effect of indoor thermal environments on recognised productivity in offices is also significant [12, 13]. Predicted mean vote (PMV) is a predictor of indoor thermal perception and satisfaction with the thermal environment. Occupant satisfaction with the thermal environment was highest when PMV equals to 0. In this case, productivity is usually considered to be the highest as well. However, many studies have found that productivity is not highest when comfort in the thermal environment is highest. The reason for this problem may be due to the fact that the level of stress that people perceive from the environment is correlated with the comfort level of the environment. Yerkes and Dodson's research on productivity from a psychological perspective has found that productivity is highest when the level of stress is moderate, while productivity decreases at higher or lower levels of stress [14]. Based on this finding, Inverted-U theory (also known as the Yerkes-Dodson law) has been proposed. This theory describes the relationship between stress and productivity as an inverted U form. Similarly, an inverted U relationship exists between metabolic activity and attention: as metabolism increases, attention increases to a certain point, but further increases in metabolism lead to a decrease in attention [15]. In extrapolating these theories to the relationship between indoor thermal environments and productivity, it is clear that workers may be more productive in moderately stressful environments, such as slightly warmer or slightly cooler environments, where productivity is maximised, than in perfectly comfortable thermal neutrality. Numerous studies have also confirmed the inverted-U relationship between environmental stress and worker productivity [16-19].

As mentioned above, many studies have modelled the relationship between sensation and worker productivity. However, most of these studies suffer from a number of shortcomings. First, most of relevant studies commonly used PMV to assess thermal environments, while PMV are only one-third as accurate in predicting the thermal sensation of occupants. The reason for this low accuracy is that the PMV is used to predict [20]. The reason for this low accuracy is that the PMV is used to predict thermal comfort for the entire occupant group rather than for individuals. Currently the dominant view of thermal comfort analyses is constructed based on thermophysiological feedback from the human body (e.g., thermal images of the face, skin temperature, heart rate variability, etc.). However, these indexes are not used in the field of evaluating thermal environments and productivity. Second, limitation of previous studies on thermal environment and productivity is the reliance on analytical methodologies such as statistical and regression analyses, such as analysis of variance (ANOVA), multivariate regression analysis (MRA), logistic regression analysis, structural equations, etc. This method analyses the link between thermal environment, thermal comfort and productivity only in a statistical perspective and may not be accurate when used for every office worker. Third, there is the lack of consideration of other factors affecting productivity, such as gender, age, individual basal metabolism and other basic personal attributes. However, these attributes also affect individual productivity to some extent. Fourth, while the type of work in an office is varied, e.g., reading, paperwork, phone calls, meetings, etc. However, current research generally modelled only a single job of varying intensity, ignoring the differences between types of work. The degree to which productivity is affected by the thermal environment varies between different types of work, and this will also affect the results of the analysis.

In addition to the influence of thermal environments on work performance, the demands of the tasks themselves can also shape thermal perception and related psychological responses. Research in cognitive ergonomics and occupational physiology indicates that higher task difficulty can elevate cognitive workload, increase metabolic rate, and alter thermoregulatory processes such as peripheral blood flow. These physiological adjustments can, in turn, influence skin temperature (ST), electrodermal activity (EDA), heart rate (HR), and ultimately thermal sensation and satisfaction. Furthermore, mental workload and stress, often quantified through tools such as the NASA Task Load Index (NASA-TLX) and the State-Trait Anxiety Inventory (STAI) stress scale, are rarely examined in the context of combined thermal and task-related demands.

When thermal environment and task difficulty act together, potential interaction effects emerge. Moderately heat- or cold-stressed conditions may amplify discomfort when combined with high task loads, while thermally neutral environments might partially mask thermal sensations due to heightened task engagement. Despite their relevance to real-world work settings, empirical studies systematically manipulating both environmental parameters and task difficulty levels, while simultaneously collecting physiological, performance, and subjective data, remain scarce. This dual influence of environmental and task-related factors underscores the need for a more integrated approach to understanding and predicting human responses in complex indoor environments. Therefore, despite a growing body of work, several critical knowledge gaps remain, as outlined in the following section.

1.2 Problem statement

Despite substantial progress in the study of indoor thermal comfort, work performance, and human–environment interactions, several important research gaps persist.

1. Limited consideration of individual variability in thermal comfort

Thermal environment assessments still predominantly rely on group-based models such as the PMV, which are not designed to capture individual differences in thermal sensation and satisfaction. Such models often fail to account for personal variables such as gender, age, body mass index (BMI), and basal metabolic rate, leading to reduced predictive accuracy for individuals.

2. Narrow task paradigms in performance studies

Existing research on thermal environments and performance typically focuses on a single task type and often under a limited number of environmental conditions. This approach oversimplifies the diverse demands of real work contexts, where task difficulty and cognitive workload vary substantially.

3. Under exploration of task difficulty effects on thermal perception

While task difficulty has been shown to influence cognitive workload and physiological responses, its specific impact on thermal sensation, satisfaction, and related psychological states (e.g., stress) remains insufficiently studied, particularly in combination with controlled variations in thermal environment.

4. Insufficient integration of physiological and multi-modal data

Physiological parameters (e.g., ST, EDA, HR) and thermal facial imaging capture dynamic responses to both environmental and task demands. However, these signals are rarely integrated into thermal comfort and performance prediction models, limiting our understanding of complex, real-time human–environment interactions.

5. Methodological constraints in modelling

Most previous studies apply traditional statistical analyses such as ANOVA or regression, which may not fully capture non-linear, high-dimensional relationships between environmental, personal, physiological, and psychological factors. More advanced analytical approaches, such as Friedman test, Bayesian modelling, machine learning, and deep learning, are underutilised in this domain.

In light of these gaps, this research aims to:

1. Experimentally examine the effects of both thermal environments and task difficulty on thermal sensation, thermal satisfaction, work performance, workload perception, and stress levels.

2. Collect a multi-modal dataset encompassing environmental metrics, personal characteristics, task parameters, subjective assessments, physiological signals, and thermal images.

3. Apply advanced statistical and machine learning techniques to develop predictive models that account for individual variability and enable the simultaneous prediction of multiple human response outcomes.

1.3 Research aim and objectives

1.3.1 Research aim

The aim of this study is to investigate and model the combined effects of thermal environments, task difficulty, and personal factors on thermal sensation, thermal satisfaction, work performance, workload perception, and stress. This will be achieved through controlled laboratory experiments, Friedman statistical analysis, and multi-modal predictive modelling that incorporates physiological parameters and thermal facial imaging alongside environmental, personal, and task-related variables.

1.3.2 Research objectives

1. Systematically test the effects of three PMV-defined thermal environments and three levels of task difficulty on subjective thermal perception, work performance outcomes, workload measures, and stress levels.

2. Apply Friedman statistical analysis to quantify the impact of task difficulty on thermal sensation, thermal satisfaction, and stress within the same thermal environment, and the impact of thermal environment on work performance, workload perception, and stress when task difficulty is held constant.

3. Develop and evaluate machine learning models to predict each outcome variable (thermal sensation, thermal satisfaction, work performance, workload perception, stress) using different feature sets, including personal characteristics, environmental parameters, task difficulty levels, and physiological parameters (ST, EDA, HR).

4. Build predictive models capable of forecasting the five outcome variables simultaneously (multi-output prediction), compare baseline models using environmental, personal, and task variables with models incorporating physiological parameters, and assess the influence of physiological data on predictive accuracy, generalisation, and computational efficiency for each outcome and for the joint prediction task.

5. Investigate the contribution of thermal facial imaging features, in combination with environmental, personal, task-related, and physiological data, to the performance of

multi-output deep learning models, and evaluate whether the inclusion of thermal images improves prediction accuracy for all five outcome variables.

1.4 Research questions

Based on the research aim and objectives, the following questions guide this study:

RQ1. How do different levels of task difficulty and PMV-defined thermal environments influence thermal sensation, thermal satisfaction, work performance, workload perception, and stress?

RQ2. Within the same thermal environment, how does task difficulty affect thermal sensation, thermal satisfaction, and stress, and, within the same task difficulty level, how does the thermal environment affect work performance, workload sensation, and stress?

RQ3. To what extent can personal characteristics, environmental parameters, and task difficulty levels predict each of the five outcome variables (thermal sensation, thermal satisfaction, work performance, workload sensation, stress) individually?

RQ4. Can a single predictive model accurately forecast all five outcome variables simultaneously?

RQ5. Does the integration of physiological signals and thermal facial imaging features with environmental, personal, and task-related information improve the performance of multi-output deep learning models, and how does this integration affect predictive accuracy for all five outcome variables compared with models without thermal images?

1.5 Outline

This chapter (Chapter 1) presents a brief description of the background of the study, the research questions and objectives, and the outline of this report.

Chapter 2 provides the literature review which includes the previous study of the thermal comfort and workload sensation; factors related to the thermal comfort and workload sensation; thermal comfort perception; and workload sensation/productivity evaluation. Followed by the summary and research gaps.

Chapter 3 provides the methodology that will be used in the thesis, which includes a brief introduction to the methodology; data collection, including the setup of the experiment, the equipment needed for the experiment, the design of the experiment, the flow of the experiment; the preparation of the data; the statistical way of analysing the data; and the development of the predictive model.

Chapter 4 presents the results and analyses of the study, combining both statistical analysis and machine-learning model outcomes. It reports the main experimental findings, highlights significant effects and interactions between thermal environment and task difficulty, and demonstrates how different modelling approaches captured these relationships. This chapter integrates empirical evidence with predictive results, establishing the basis for subsequent discussion.

Chapter 5 provides the discussion of the findings. It interprets the combined statistical and modelling outcomes within the context of existing literature, explains their theoretical and practical implications, and reflects on how the present results advance understanding of thermal-cognitive interaction. The chapter also addresses methodological advantages, limitations, and suggestions for future work.

Chapter 6 concludes the thesis by summarizing the main findings, contributions, and applications of the research. It restates the study's objectives and how they have been achieved, reflects on the overall significance of the work, and outlines potential directions for future research and real-world implementation in intelligent, human-centred thermal-environment design.

Chapter 2 Literature Review

2.1 Introduction to the research field

2.1.1 Overview of the research field

The study of thermal comfort and work-related outcomes in occupational environments represents a critical intersection of environmental ergonomics, occupational health, and organisational psychology [21]. Thermal comfort, defined as the state of mind that expresses satisfaction with the surrounding thermal environment [7, 22], is not only a determinant of physical well-being but also a factor influencing cognitive efficiency, task performance, and emotional stability in the workplace [23, 24]. Inadequate thermal conditions can impair concentration, slow decision-making, elevate perceived workload, and increase stress or stress [25], thereby reducing productivity and potentially compromising safety in high-risk occupations [26].

The importance of this topic extends beyond individual health and comfort. In modern workplaces, ranging from climate-controlled offices to industrial facilities, thermal management is closely tied to operational efficiency, employee satisfaction, and energy consumption [27, 28]. Historically, research has been guided by models such as Fanger's Predicted Mean Vote (PMV) and Predicted Percentage of Dissatisfied (PPD) [29], as well as adaptive comfort frameworks [30], which establish quantifiable links between environmental conditions and subjective comfort. However, in occupational settings, these models must also be understood in relation to task demands, job characteristics, and personal factors that shape how thermal environments influence work outcomes [31, 32].

Recent years have seen a shift from purely comfort-oriented studies toward investigations of how thermal environments interact with work-related variables such as task difficulty, cognitive workload, and emotional states [33, 34]. Empirical evidence shows that even within comfort-compliant temperatures, variations in thermal conditions can affect work speed, accuracy, and error rates [35], especially in cognitively demanding or safety-critical tasks [36].

Advances in wearable physiological monitoring, including continuous skin temperature tracking, electrodermal activity (EDA), and heart rate (HR) measurement, have enabled researchers to detect subtle changes in physiological state that correlate with both thermal perception and work performance [37]. Thermal facial imaging has emerged as a complementary, non-intrusive method for assessing thermal and emotional states [38], offering spatial temperature data that can be integrated with other modalities [39]. Increasingly, studies are adopting multi-modal approaches that combine environmental parameters, personal characteristics, physiological signals, and task performance metrics to better capture the complexity of workplace thermal–performance interactions [40].

Despite these advances, several gaps remain. First, the interaction between thermal comfort and work performance is context-dependent, with results varying across task types, performance metrics, and individual differences [41]. Second, much of the literature isolates one dimension, either comfort or performance, without examining their dynamic interplay over time [42]. Third, laboratory studies often fail to replicate the variability and multitasking demands of real-world work environments [43]. Furthermore, integrative models that simultaneously account for thermal sensation, thermal satisfaction, work performance, workload perception, and stress, while incorporating both physiological and thermal imaging data, are still scarce [44].

Given the multidisciplinary nature and practical importance of this field, it is essential to conduct a systematic literature search to map the range of research linking thermal comfort and work-related outcomes in occupational contexts. This review adopts a structured approach to capture studies spanning environmental ergonomics, occupational performance, physiological monitoring, and thermal imaging. In addition to thematic synthesis, bibliometric mapping using VOSviewer is employed to visualise the thematic structure of the field [45], identify dominant research clusters, and reveal under-explored intersections. The findings of this search and analysis are presented in the next section, providing an empirical foundation for the thematic discussions that follow.

2.1.2 Literature search

A targeted literature search was conducted to capture studies examining the relationships between thermal environments, task demands, individual differences, and human performance or psychological states. The search strategy balanced breadth and precision by combining core concepts (e.g., thermal comfort, PMV, task difficulty, cognitive performance, physiological measures, thermal imaging) with synonymous and related terms, while applying exclusion criteria to remove studies outside this scope. The resulting dataset provided a representative foundation for bibliometric mapping and keyword co-occurrence analysis, which informed the thematic framework developed in this review.

The search was conducted in Scopus, selected for its extensive multidisciplinary coverage and consistent indexing of peer-reviewed journal articles. It targeted publications from 2015 to 2025, written in English, and classified as journal articles. A modular keyword strategy was used, in which search terms were organised into three core conceptual domains and a set of exclusion terms (Table 2.1). Within each domain, synonymous or closely related terms were connected with the Boolean operator OR, while the domains were combined with AND to ensure that all retrieved articles addressed all focal concepts. Exclusion terms were applied using the NOT operator to filter out literature outside the study's scope.

Table 2.1 Concept groups and associated search terms.

Concept group	**Search terms**
Thermal comfort	"Thermal comfort", "thermal sensation", "thermal satisfaction", PMV, PPD
Work and performance	"Task difficulty", "work performance", "workload sensation", productivity, "task performance", "cognitive workload", "mental workload", ergonomics, "human factors"
Parameters	"environmental parameters", "indoor environmental quality", age, gender, BMI, "physiological parameters", "heart rate", HR, "skin temperature", ST, "electrodermal activities", EDA, "thermal imaging", "wearable sensors"
Exclusion terms	"urban heat island", "outdoor thermal comfort"

The exact Boolean search string implemented in Scopus was as follow:

TITLE-ABS-KEY ("thermal comfort" OR "thermal sensation" OR "thermal satisfaction" OR PMV OR PPD)

AND

TITLE-ABS-KEY ("task difficulty" OR "work performance" OR "workload sensation" OR productivity OR "task performance" OR "cognitive workload" OR "mental workload" OR "human factors")

AND

TITLE-ABS-KEY ("environmental parameters" OR "indoor environmental quality" OR age OR gender OR BMI OR "physiological parameters" OR "heart rate" OR HR OR "skin temperature" OR ST OR "electrodermal activities" OR EDA OR "thermal imaging" OR "wearable sensors")

AND NOT

TITLE-ABS-KEY ("urban heat island" OR "outdoor thermal comfort")

AND (LIMIT-TO (DOCTYPE, "ar"))

AND (LIMIT-TO (LANGUAGE, "English"))

This search returned 296 records. All results were exported in both .csv and .ris formats for subsequent processing. Duplicate records were first removed. The remaining records then underwent manual screening of titles and abstracts based on clearly defined exclusion principles. Specifically, studies were excluded if they:

- Studies published more than 10 years prior to the search date, ensuring that the review reflects recent developments (approximately 2015 onward).
- Focused on animal experiments or agricultural production environments (e.g., livestock housing, poultry farms).

- Were conducted primarily in outdoor settings or field environments unrelated to indoor occupational or educational contexts.
- Investigated tasks unrelated to work or learning, such as sports competitions, endurance training, leisure activities, bathing, or purely medical/clinical interventions without direct workplace relevance.

Applying these criteria reduced the dataset to 181 publications that met all inclusion requirements and were relevant to the study's objectives. To characterise the research landscape, bibliometric analyses were conducted to examine the temporal evolution of publications, geographical distribution of research activity, key source titles, and thematic structure based on author keywords. Prior to thematic mapping, generic or non-informative terms (e.g., "male", "human", "article") were manually removed from the keyword list to ensure the focus remained on concepts directly related to the study's scope. The results of these analyses are presented in Section 2.1.3.

2.1.3 Bibliometric Analysis

To further characterise the research landscape on indoor environmental conditions and human performance, the final dataset of 181 publications was analysed in terms of temporal evolution, geographical distribution, publication sources, and thematic structure. This quantitative overview provides insight into how the field has developed over time, where the main research hubs are located, and which journals serve as the principal outlets for dissemination. Publication counts by year, country, and source title were compiled from Scopus metadata, while thematic mapping was conducted using VOSviewer (version 1.6.20) based on author keywords. Together, these analyses offer a multi-dimensional perspective on the scope and focus of the literature, forming the basis for identifying research trends and gaps in subsequent sections.

Figure 2.1 presents the annual distribution of the 181 publications included in the dataset from 2015 to 2025. The number of publications shows a generally upward trajectory over the past decade, increasing from 5 papers in 2015 to a peak of 32 in 2024. The early period from 2015 to 2019 exhibits steady but modest growth (5 to 13 publications per year). A more pronounced rise occurred between 2020 and 2021, with the annual output increasing from 15 to 23 papers. Although there was a temporary dip in 2022 (19 publications), the trend quickly rebounded, culminating in the record high of 2024. The partial count for 2025 (19 publications) suggests sustained activity in the field, though the final total for the year is yet to be determined. Overall, the data indicate accelerating research interest in indoor environmental conditions and human performance, particularly in the last five years.

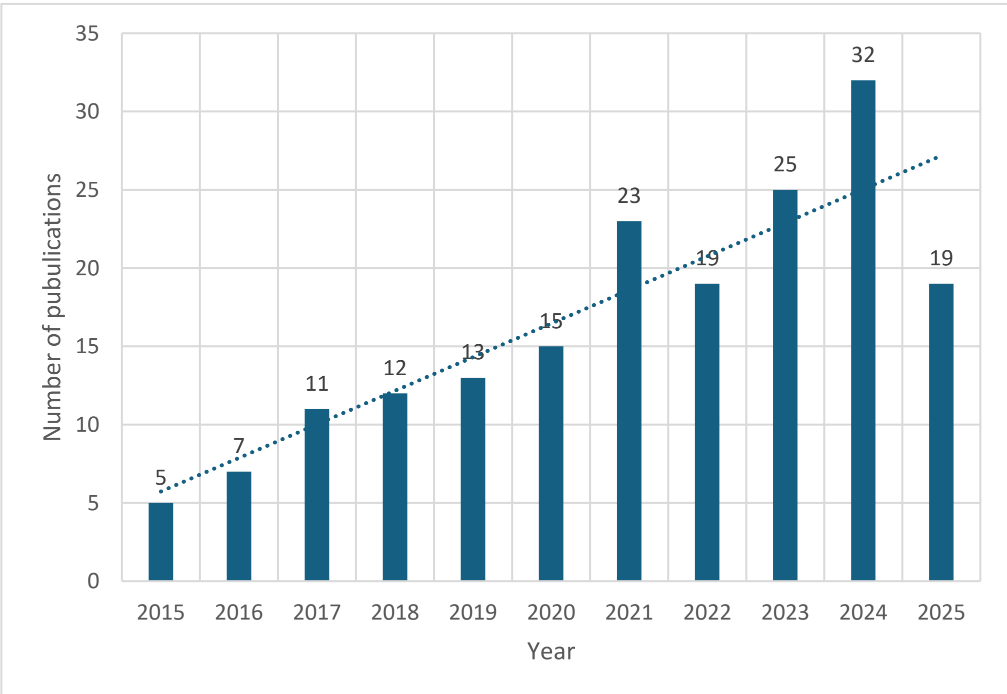


Figure 2.1 Annual publication trends from 2015 to 2025 based on the bibliometric dataset (n = 181).

Table 2.2 illustrates the geographical distribution of publications based on the first author's affiliation. China leads with 31 publications (17.1% of the total), followed by the United States (21, 11.6%) and Italy (14, 7.7%). Other prominent contributors include South Korea (13, 7.2%), Australia (10, 5.5%), Japan (9, 5.0%), Hong Kong (8, 4.4%), Brazil (6, 3.3%), Malaysia (6, 3.3%), and the Netherlands (5, 2.8%). Collectively, these top 10 countries account for 67.9% of all publications, while the remaining 42 countries and regions together make up the remaining 32.0%. This distribution reflects a strong concentration of research activity in East Asia, North America, and parts of Europe, with emerging contributions from other regions.

Table 2.2 Top 10 countries/regions by number of publications, with remaining countries aggregated as "Other".

Countries	**Number of publications**
Building and Environment	31
USA	21
Italy	14
South Korea	13
Australia	10
Japan	9
Hong Kong SAR	8
Brasil	6
Malaysia	6
Netherlands	5
Other countries/regions	58

Table 2.3 shows the distribution of publications by source title. Building and Environment is the most prolific journal, accounting for 33 papers (18.2%), followed by Energy and Buildings (18, 9.9%) and Journal of Building Engineering (14, 7.7%).

Other notable outlets include Sustainability (Switzerland) (11, 6.1%), Indoor Air (7, 3.9%), Building Research and Information (5, 2.8%), Buildings (5, 2.8%), Applied Ergonomics (2, 1.1%), Atmosphere (2, 1.1%), and Applied Sciences (Switzerland) (2, 1.1%). Collectively, these top 10 journals publish 54.7% of the identified works, indicating a moderately concentrated publication landscape dominated by a small number of specialised journals in building science, environmental engineering, and ergonomics.

Table 2.3 Top 10 countries/regions by number of publications, with remaining countries aggregated as "Other journals".

Countries	**Number of publications**
Building and Environment	33
Energy and Buildings	18
Journal of Building Engineering	14
Sustainability	11
Indoor Air	7
Building Research and Information	5
Buildings	5
Applied Ergonomics	2
Atmosphere	2
Applied Sciences	2
Other journals	82

Following the analysis of publication sources, author keywords from the 181 publications were examined to map the thematic structure of research on indoor environmental conditions and human performance. Keywords were extracted from the Scopus records and processed using VOSviewer. A minimum occurrence threshold of seven was applied, which initially yielded 90 keywords. To improve interpretability, nine generic or non-informative terms (e.g., male, adults, article) were manually removed, resulting in a final set of 81 keywords for analysis. The minimum cluster strength parameter was set to seven to achieve a balance between network density and thematic clarity. The resulting keyword co-occurrence network is presented in Figure 2.2, where node size reflects the frequency of occurrence and link strength indicates the extent of co-occurrence between terms.

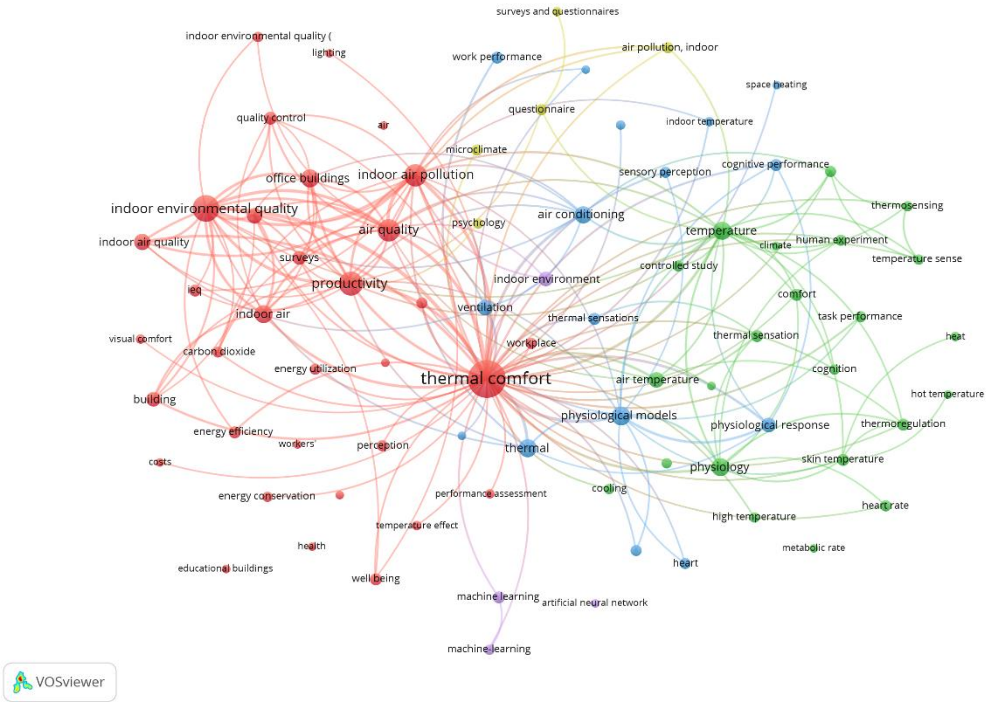


Figure 2.2 Keyword co-occurrence network

Analysis of the network revealed five main thematic clusters. Cluster 1 focuses on indoor environmental quality and building-related factors, encompassing terms such as indoor environmental quality (59 occurrences), air quality (41), indoor air pollution (41), productivity (46), office buildings (27), and ventilation (19). Cluster 2 centres on thermal and physiological parameters, including temperature (28), thermal sensation (13), humidity (10), skin temperature (13), heat (8), and air temperature (21). Cluster 3 emphasises performance assessment and thermal comfort, with keywords such as thermal comfort (115), cognitive performance (12), physiological response (18), work performance (12), and performance assessment (8). Cluster 4 addresses psychological and perceptual aspects of indoor environments, represented by terms like perception (10), psychology (9), microclimate (10), surveys (16), and questionnaire (11). Cluster 5 relates to modelling and computational approaches, including machine learning (12), machine-learning (9), artificial neural network (7), physiological models (31), and indoor environment (17). It should be noted that these clusters are generated automatically by VOSviewer based on keyword co-occurrence patterns and do not correspond directly to the thematic categories used in the subsequent qualitative synthesis.

The most frequent keyword was thermal comfort (115 occurrences), followed by indoor environmental quality (59), productivity (46), air quality (41), and indoor air pollution (41). These terms reflect the dominant focus of the field on comfort assessment, environmental quality, and productivity outcomes. In terms of temporal trends, emerging topics include thermal (average publication year 2023.96), machine-learning (2023.56), machine learning (2023.08), performance assessment (2023.50), and artificial neural network (2023.29), indicating a growing interest in predictive modelling and advanced analytical methods. Foundational topics such as temperature (2020.86), thermal sensation (2020.15), and ventilation (2020.95) have been consistently studied over a longer period. Citation analysis further shows that high-impact topics include sick building syndrome (average normalised citation 2.27), space heating (2.08), indoor temperature (2.09), intelligent buildings (1.94), and work performance (1.69), suggesting that both health-related issues and energy strategies hold significant influence in the literature.

Overall, the bibliometric mapping derived from 81 keywords illustrates a structured thematic landscape comprising five interconnected clusters: (1) indoor environmental quality and building-related factors, (2) thermal and physiological parameters, (3) performance assessment and thermal comfort, (4) psychological and perceptual aspects, and (5) modelling and computational approaches. These clusters reveal clear intersections between different research streams but are generated automatically by VOSviewer based on co-occurrence patterns. As such, they may not fully capture conceptual overlaps or align precisely with the analytical focus of this study. Therefore, in the next section, the keywords are re-classified manually using a content-driven approach to provide a more targeted thematic synthesis.

2.1.4 Manual Thematic Classification

While the five clusters generated automatically by VOSviewer in Section 2.1.3 provided a useful overview of co-occurrence patterns in the literature, their purely statistical grouping does not necessarily align with the analytical framework of this study. In particular, conceptually related topics may be split across clusters, and some clusters may combine terms that are not directly relevant to the research questions.

To ensure that the thematic structure reflects the analytical focus of this study, the automatically generated clusters were reorganised through manual classification. Although automated co-occurrence clustering effectively reveals statistical associations between keywords, it does not always yield conceptually coherent groupings for domain-specific analysis. The manual classification was therefore guided by the substantive relationships between concepts, rather than by statistical proximity alone. Specifically, keywords were reassigned according to their relevance to six predefined analytical dimensions: (1) the effects of thermal environment; (2) the effects of task factors; (3) the interaction between environmental and task influences; (4) the influence of individual differences; (5) physiological, behavioural, and imaging indicators; and (6) analytical and experimental methodologies. This approach integrated domain expertise, semantic interpretation of keywords, and an understanding of the methodological context of the underlying studies. As a result, a single automated cluster could contribute to multiple themes, and each theme could incorporate terms from several clusters.

Table 2.4 shows the six manual themes and their corresponding sources in the automated clustering, along with representative keywords.

Table 2.4 Manual themes and their sources in the VOSviewer clustering

Manual theme	**Source in automated clustering**	**Representative keywords**
Effects of Thermal Environment	Indoor environmental quality and building factors; Thermal and physiological parameters; Performance assessment and thermal comfort	indoor environmental quality, air quality, temperature, humidity, thermal sensation, thermal comfort, thermal satisfaction, heat stress, productivity, physiological response
Effects of Task Factors	Performance assessment and thermal comfort; Psychological and perceptual aspects	task difficulty, task complexity, cognitive workload, performance accuracy, reaction time, perception, attention, error rate
Interaction Between Thermal	Thermal and physiological parameters; Performance	thermal sensation × task difficulty, combined effects, dual-task performance,

Environment and Task Factors	assessment and thermal comfort	workload–environment interaction, cognitive performance change, fatigue
Influence of Individual Differences	Psychological and perceptual aspects	gender, age, body mass index (BMI), thermal adaptability, acclimatisation, personal sensitivity, prior exposure
Physiological, Behavioural, and Imaging Indicators	Thermal and physiological parameters; Modelling and computational approaches	ST, EDA, HR, respiration rate, eye movement, facial expression analysis, posture change, thermal imaging, facial thermography
Analytical and Experimental Methodologies	Modelling and computational approaches	statistical analysis, ANOVA, regression, Bayesian modelling, machine learning, deep learning, experimental design, signal processing, image analysis

1. Effects of Thermal Environment

Encompasses the full range of impacts that thermal conditions, whether hot, cold, or neutral, can exert on individuals. This includes effects on thermal comfort and satisfaction, work performance (accuracy, reaction time, error rates), perceived workload (NASA-TLX), State-Trait Anxiety Inventory (STAI), and physiological states. Both acute and prolonged exposure scenarios are considered, along with short-term fluctuations and recovery patterns.

2. Effects of Task Factors

Covers the independent influence of task-related characteristics, particularly task difficulty, complexity, and time constraints, on thermal sensation, thermal satisfaction, cognitive performance, perceived workload, and emotional responses. These effects are considered irrespective of environmental variation, highlighting the role of task design in shaping both subjective and objective outcomes.

3. Interaction Between Thermal Environment and Task Factors

Investigates how thermal conditions and task demands jointly influence outcomes, including cognitive performance, workload perception, thermal comfort, and psychological states. This theme captures synergistic and antagonistic effects, bidirectional influences (e.g., heat stress impairing performance, high task load altering thermal perception), and mechanisms explored through factorial designs, interaction-term regression, and other analytical approaches.

4. Influence of Individual Differences

Examines how personal characteristics, such as gender, body mass index (BMI), thermal adaptability, and prior acclimatisation, modify the effects of thermal environment and task factors. This theme highlights that identical environmental or workload conditions can have markedly different impacts across individuals due to physiological, psychological, or experiential factors, and considers these differences in both single-task and multi-task contexts.

5. Physiological, Behavioural, and Imaging Indicators

Covers objective measures used to assess thermal comfort, satisfaction, work performance, workload, and emotional states. This includes physiological signals such as skin temperature (ST), electrodermal activity (EDA), heart rate (HR), respiration rate; behavioural indicators such as eye movements, facial expressions, and posture changes; and imaging-based approaches including facial thermography and other thermal or optical imaging techniques. Integration of multimodal physiological, behavioural, and imaging data is also considered.

6. Analytical and Experimental Methodologies

Encompasses the range of analytical and methodological approaches applied in the reviewed studies. This includes statistical techniques (e.g., ANOVA, regression, Bayesian analysis), machine learning and deep learning models, experimental design strategies (e.g., within-subject, between-subject, crossover designs), multimodal feature integration, signal processing, and image analysis methods. Emphasis is placed on approaches that enable robust prediction and interpretation of thermal comfort and performance outcomes.

2.1.5 Summary

Section 2.1 provided a comprehensive bibliometric overview of the research domain concerning the thermal environment, task factors, and their interactions. In Section 2.1.1, the scope of the bibliometric study was defined, clarifying the research questions and analytical objectives. Section 2.1.2 detailed the data acquisition process, including database selection, search strategy, and screening procedures, which ensured that the dataset was both comprehensive and relevant to the study's aims.

In Section 2.1.3, keyword co-occurrence analysis using VOSviewer was conducted to reveal the structure of the research field. This analysis produced five statistically derived clusters, each representing a distinct pattern of keyword associations. While informative, these automated clusters did not always align with conceptually coherent categories required for the present investigation.

To address this, Section 2.1.4 applied manual thematic classification, re-organising keywords into six analytically meaningful themes: Effects of Thermal Environment; Effects of Task Factors; Interaction Between Thermal Environment and Task Factors; Influence of Individual Differences; Physiological, Behavioural, and Imaging Indicators; and Analytical and Experimental Methodologies. This reclassification combined statistical evidence from the bibliometric analysis with domain expertise and semantic interpretation, thereby producing a thematic framework that is directly aligned with the objectives of this research.

Together, these steps established both a quantitative mapping of the literature and a qualitative thematic structure. This dual perspective offers a robust foundation for the subsequent literature review, which will explore each theme in depth, synthesising empirical findings, identifying methodological trends, and highlighting knowledge gaps. The insights gained will directly inform the experimental design, measurement strategies, and analytical methods described in the following chapters.

2.2 Effects of thermal environment

The thermal environment is a fundamental component of the indoor climate that directly shapes human experience and performance. A growing body of research demonstrates that temperature, humidity, air movement, and related thermal factors influence not only subjective comfort but also physiological functioning, cognitive processes, and emotional states. This section reviews these effects across four main dimensions. Section 2.2.1 clarifies the definition and scope of the thermal environment. Section 2.2.2 discusses its impacts on thermal comfort and satisfaction from a subjective perspective. Section 2.2.3 examines physiological and health-related outcomes, while Section 2.2.4 addresses cognitive performance. Emotional and affective responses are considered in Section 2.2.5. Finally, Section 2.2.6 summarizes the key findings of this part.

2.2.1 Definition and scope of thermal environment

The thermal environment is a central concept in the study of indoor environmental quality (IEQ) because it directly influences human thermal balance, comfort, and performance [18]. In most technical standards and academic literature, the thermal environment is described as the set of environmental parameters that govern heat exchange between the human body and its surroundings. These parameters typically include air temperature, relative humidity, air velocity, and mean radiant temperature [46]. Collectively, they determine the energy balance of the human body, shaping whether an individual experiences thermal comfort, neutrality, or discomfort.

According to ASHRAE Standard 55: Thermal Environmental Conditions for Human Occupancy, acceptable thermal environments are defined in terms of ranges of these parameters that are expected to satisfy a specified percentage of occupants [7]. Although the standard does not provide a single-sentence definition of "thermal environment," it implicitly frames the concept as the combination of physical conditions that affect thermal sensation and comfort. Similarly, ISO 7730: Ergonomics of the thermal environment identifies air temperature, mean radiant temperature, air velocity, and humidity as the key environmental parameters, while also highlighting the importance of personal parameters such as metabolic rate and clothing insulation [47]. This broader framing recognizes that thermal experience emerges from the interaction between environmental conditions and human factors.

In academic research, the scope of the thermal environment has been extended beyond purely physical variables. Scholars emphasize that thermal conditions are also shaped by psychological, cultural, and contextual dimensions, since individuals interpret and evaluate the same physical conditions differently depending on expectations, prior experiences, and adaptive behaviours [48]. For example, studies have shown that acceptable thermal ranges vary across climatic regions, building types, and even between seasons within the same location [49]. This perspective highlights that the thermal environment is not static but dynamic, and that individual and contextual variability complicate the establishment of universal comfort standards.

In addition to the definition and parameters, the temporal and spatial patterns of exposure are essential for describing the scope of the thermal environment. Short-term or acute exposure to heat or cold can elicit immediate perceptual and physiological responses, while long-term exposure may result in acclimatization, adaptation, or cumulative health effects [50]. Furthermore, laboratory experiments often assume constant thermal conditions, whereas real-world environments are more likely to be fluctuating, with daily or seasonal variations in temperature, humidity, and air movement [51]. Recognizing these exposure patterns is critical, as fluctuating and long-term conditions may produce different comfort perceptions and health outcomes compared to short-term, stable exposures.

Research on the thermal environment has been conducted across a wide range of application contexts. Laboratory-based studies, often carried out in climatic chambers, enable precise control of thermal variables and allow researchers to isolate their effects [52]. Field studies in offices, classrooms, residential buildings, and hospitals provide insights into how occupants experience and adapt to thermal conditions under everyday circumstances [53]. Beyond typical indoor settings, investigations have also focused on extreme or specialized environments, such as industrial workplaces, vehicles and aircraft, and regions with severe climatic conditions [54]. These diverse contexts illustrate the broad relevance of the thermal environment and its significance for both comfort and performance.

To address these complexities, international guidelines (e.g., ASHRAE, ISO) typically define acceptable thermal environments in terms of comfort zones that represent conditions under which most occupants are expected to be satisfied [7]. However, these guidelines are often criticized for being based on population averages and controlled laboratory conditions, which may not fully capture inter-individual variability or real-world adaptability [52]. As a result, the modern understanding of the thermal environment increasingly incorporates a human-centred perspective, viewing it not only as a set of physical parameters but also as a multidimensional construct shaped by physiology, psychology, and social context [55].

2.2.2 Effects on thermal comfort and satisfaction (subjective impacts)

Thermal comfort is commonly defined as *"the condition of mind that expresses satisfaction with the thermal environment"*, a definition that has been formalized in ASHRAE Standard 55 [7]. This definition highlights that comfort is inherently a subjective psychological state rather than a purely physiological reaction. According to ASHRAE, six primary factors determine thermal comfort: four environmental parameters, air temperature, mean radiant temperature, air velocity, and relative humidity, and two personal parameters, metabolic rate and clothing insulation [55]. The four environmental parameters together constitute the thermal environment, which governs heat exchange between the human body and the surrounding space [56]. Variations in these parameters directly influence whether individuals report feeling cold, neutral, or hot.

To quantify the influence of the thermal environment on comfort, Fanger developed the Predicted Mean Vote (PMV) model, which estimates the average thermal sensation of a large group of people on a seven-point scale ranging from −3 (cold) to +3 (hot) [29]. Building on this, the Predicted Percentage of Dissatisfied (PPD) index was introduced to estimate the proportion of occupants likely to feel thermally dissatisfied under specific conditions [57]. These indices were later incorporated into ASHRAE Standard 55 and have since become the primary benchmarks for defining acceptable indoor thermal environments [58]. PMV and PPD thus provide a standardized, quantitative framework that links measurable environmental parameters to subjective comfort outcomes. However, subsequent field studies have revealed important limitations. Because PMV and PPD were derived primarily from controlled laboratory experiments with homogeneous groups, they often fail to capture the diversity of real-world comfort experiences [30]. In particular, the models typically do not account for psychological expectations, cultural differences, or adaptive behaviours that strongly influence comfort perception in practice. For instance, in naturally ventilated buildings, occupants frequently report acceptable comfort under conditions that PMV would classify as unacceptable [59]. Such discrepancies underscore the need to move beyond aggregated indices and examine subjective dimensions of thermal comfort more directly. As a result, researchers have increasingly focused on measures such as thermal sensation, thermal preference, thermal acceptability, and thermal satisfaction, which provide a more nuanced understanding of how the thermal environment influences comfort at the individual level [60]. In practice, these subjective dimensions are typically assessed through point-in-time ("right-now") surveys, which capture occupants' thermal perceptions at a specific moment under actual environmental condition [61].

Thermal sensation (TSENS)

Thermal sensation refers to an individual's immediate perception of thermal state, typically expressed on a bipolar scale ranging from cold to hot [62]. It is most commonly measured through the Thermal Sensation Vote (TSENSV), a seven-point Likert-type scale originally developed by Fanger and subsequently adopted in ASHRAE Standard 55 as the standard method for assessing subjective thermal response [7]. TSENSV has been widely used in both experimental and field studies, and it serves as the cornerstone of predictive models such as the PMV [29]. The seven scale points and their descriptors are shown in Table 2.5

Table 2.5 Thermal Sensation Vote (TSENSV, 7-point)

Scale	**Descriptor**
+3	Hot
+2	Warm
+1	Slightly warm
0	Neutral
-1	Slightly cool

-2	Cool
-3	Cold

The thermal environment influences TSENSV primarily through four environmental parameters. Air temperature is the dominant factor, with higher temperatures shifting votes toward the warm side and lower temperatures toward the cool side. Mean radiant temperature (MRT) exerts a similar effect, particularly noticeable in spaces with large, glazed façades or radiant systems. Air velocity modifies convective heat transfer, providing cooling relief in warm conditions but potentially causing draft discomfort in cooler environments. Relative humidity generally plays a secondary role, though high humidity in hot climates can suppress evaporative cooling while very dry air in cold conditions may increase discomfort [63].

A substantial body of research has examined how environmental parameters shape TSENSV. Early chamber experiments established systematic links between temperature, air movement, and reported sensation, forming the basis of the PMV model [64]. Later investigations highlighted important deviations in real buildings: for example, occupants in naturally ventilated offices often reported neutral sensations under conditions that PMV predicted as slightly warm or cool [64]. Large cross-climatic surveys such as the ASHRAE RP-884 project further showed that acceptable sensation ranges vary across regions, reflecting adaptation and cultural expectations [65]. Studies in more extreme environments, including industrial workplaces and outdoor settings, also reported departures from model predictions, underscoring the role of contextual and behavioural factors [66].

In summary, TSENS represents the most direct subjective response to the thermal environment. While strongly determined by physical parameters such as air temperature, MRT, air velocity, and humidity, TSENSV is also moderated by adaptation, prior exposure, and psychological expectations. Evidence from field and cross-climatic research indicates that TSENS cannot be fully explained by physical conditions alone, underscoring the need to integrate subjective and contextual dimensions when evaluating thermal comfort [67].

Thermal preference (TPREF)

Thermal preference refers to an occupant's expressed desire for a change in the thermal environment, indicating whether conditions should become cooler, warmer, or remain unchanged [68]. It is typically measured using a three-point Likert-type scale, which has been widely applied in ASHRAE 55 thermal comfort surveys [7]. In addition to temperature preference, many field studies, most notably those compiled in the ASHRAE RP-884 database, also included a parallel three-point scale for air movement preference, where occupants indicate whether they would prefer the air to be slower, no change, or faster [65]. These scales are summarized in Table 2.6.

Table 2.6 Thermal Preference Vote (TPREFV) and Air Movement Preference Vote scales

	Aspects	
	Temperature preference	**Air movement preference**
Prefer to be	Cooler	More air movement
	No change	No change
	Warmer	Less air movement

TPREF is strongly shaped by the thermal environment, with air temperature exerting the clearest influence: in warm conditions occupants typically prefer cooler environments, while in cold conditions they prefer warmer ones [68]. Similar effects are observed for radiant surfaces, humidity, and air movement, with airflow often functioning as an independent dimension of preference rather than merely a modifier of temperature [69].

Past studies have examined TPREF across a wide range of contexts. Research has shown that temperature and radiant conditions systematically influence whether occupants want warmer or cooler environments [70]. Other studies highlighted the role of air movement, confirming that airflow is often evaluated as a distinct comfort dimension [71]. Large survey projects such as RP-884 have demonstrated that preferences vary across climates, with stronger demands for cooling in hot–humid regions and more balanced or warmer preferences in cooler climates [72]. Additional work has explored how factors such as clothing, cultural norms, and age shape preference patterns, indicating that TPREF is not solely a physical response but also influenced by social and behavioural contexts [73].

In summary, TPREF captures the directional desire for change in both temperature and air movement. While closely tied to thermal environment parameters, preference responses vary across climates, seasons, and building types, reflecting not only physical conditions but also adaptive opportunities available to occupants. The integration of temperature and air movement preference scales, as implemented in RP-884, provides a richer basis for understanding adaptive comfort in real-world contexts.

Thermal acceptability (TAC)

Thermal acceptability (TAC) reflects whether occupants consider the thermal environment to meet a minimum standard of comfort, regardless of whether it feels slightly warm, slightly cool, or neutral [74]. It is often assessed with a binary question, "Is this environment acceptable?", which underpins the 80% acceptability criterion in ASHRAE Standard 55 [7]. To capture more nuanced responses, TAC can also be measured on a 7-point scale ranging from very unacceptable to very acceptable, as summarized in Table 2.7.

Table 2.7 Thermal Acceptability Vote (TACV, 7-point)

Scale	**Descriptor**
+3	Very acceptable
+2	Moderately acceptable
+1	Slightly acceptable
0	Just acceptable (neutral)
-1	Slightly unacceptable
-2	Moderately unacceptable
-3	Very unacceptable

TAC is primarily determined by how environmental parameters interact with occupants' tolerance limits. Air temperature is the dominant factor, with acceptability declining rapidly once conditions deviate from the comfort range. Radiant asymmetry (e.g., cold windows or hot ceilings) can reduce acceptability even when average temperatures are comfortable. Air movement may improve acceptability in warm environments but cause draft complaints in cooler ones. Humidity generally plays a secondary role, though high levels in hot climates and excessive dryness in cold climates both contribute to lower acceptability [75].

Studies on TAC emphasize that acceptability is sensitive not only to overall temperature but also to localized discomfort. Research has shown that cold window surfaces, hot ceilings, or drafts can significantly reduce the proportion of occupants who judge conditions acceptable, even when average thermal conditions are within comfort ranges [76]. Other studies highlight that acceptability improves when occupants have access to simple adaptive measures, such as shading devices or personal fans, which allow them to mitigate discomfort [77]. Investigations have also noted that prolonged exposure to marginally uncomfortable conditions tends to lower acceptability over time, suggesting that tolerance is dynamic rather than fixed [78]. These findings underline that TAC reflects how people weigh both environmental conditions and the opportunities available to manage them.

In summary, TAC offers a population-based criterion that complements individual measures such as sensation and preference. By quantifying the proportion of occupants who find conditions acceptable, it provides a practical metric for performance evaluation and compliance (e.g., the 80% rule in ASHRAE Standard 55) [7]. TAC also serves as a bridge between research and practice, translating subjective experience into a simple tool for design and operation while highlighting the diversity of occupant responses.

Thermal satisfaction (TAST)

Thermal satisfaction (TSAT) captures occupants' overall evaluation of indoor thermal conditions, expressing how satisfied or dissatisfied they feel rather than simply whether conditions are acceptable [79]. Unlike TAC, which reflects tolerance thresholds, TSAT represents a broader appraisal of comfort expectations over time. It is widely applied in post-occupancy evaluation surveys, typically measured with a 7-point scale (Table 2.8).

Table 2.8 Thermal Satisfaction Vote (TSATV, 7-point)

Scale	**Descriptor**
+3	Very satisfied
+2	Moderately satisfied
+1	Slightly satisfied
0	Just satisfied (neutral)
-1	Slightly unsatisfied
-2	Moderately unsatisfied
-3	Very unsatisfied

TSAT is strongly shaped by environmental conditions. High or low air temperatures can both reduce satisfaction, and radiant asymmetry, such as cold windows or overheated surfaces, often lowers ratings even when average temperatures are acceptable [80]. Air movement usually improves satisfaction in warm settings but may cause draft complaints in cooler ones, while humidity plays a smaller role, with excessive moisture or dryness contributing to discomfort [81]. In addition, psychological factors such as expectations and perceived control can modify these responses, highlighting that satisfaction reflects both physical and contextual influences.

Research on TSAT has consistently shown that satisfaction is not fully explained by thermal conditions alone. Experimental studies found that subjects often expressed dissatisfaction when thermal transitions were abrupt, even if final conditions fell within comfort standards [82]. Field observations indicated that perceived control plays an important role: occupants with access to windows, fans, or thermostats generally reported higher satisfaction than those without such options [83].

Longitudinal studies have also suggested that satisfaction is influenced by exposure duration, with prolonged periods of mild discomfort leading to greater dissatisfaction than short-term exposures at comparable conditions [84]. Together, these findings highlight that TSAT reflects both physical parameters and the broader experiential context of thermal exposure.

In summary, TSAT reflects occupants' overall appraisal of the thermal environment, extending beyond sensation, preference, or acceptability. As a key indicator in POE and IEQ assessments, it links environmental conditions and psychosocial factors to well-being and building performance [85].

Other aspects

In addition to sensation, preference, acceptability, and satisfaction, several other subjective dimensions have been employed to capture the richness of occupants' thermal experience. These include the thermal comfort vote (TCV), which directly rates comfort or discomfort; thermal expectation (TEXP), which reflects anticipatory beliefs shaped by culture, climate, and building type; perceived control (PCON), which assesses the extent to which occupants feel able to influence their environment; and overall comfort (OCOM), a holistic evaluation that integrates the thermal domain with other aspects of indoor environmental quality. Each of these dimensions adds a complementary perspective, emphasizing cognitive, behavioural, or integrative aspects that extend beyond immediate thermal sensation.

Past research has demonstrated that these dimensions are sensitive to both environmental and contextual factors. Laboratory studies have shown that TCV does not always align with predicted PMV values, underscoring the limits of purely physical models [86]. Field surveys highlight how expectations and perceived control strongly mediate satisfaction: occupants in naturally ventilated buildings often report higher comfort and tolerance, even at conditions outside standard comfort zones [87]. Similarly, OCOM has been widely used in post-occupancy evaluation and IEQ research, with thermal conditions consistently emerging as one of the strongest predictors of global comfort, though moderated by acoustics, lighting, and air quality[88] .

Taken together, these additional dimensions reveal that thermal experience is not only a physiological response to environmental parameters such as temperature, humidity, and air velocity, but also a cognitively and socially mediated appraisal. By incorporating expectation, control, and overall evaluations, research has advanced from a narrow focus on temperature balance to a broader understanding of how thermal environments shape human comfort, well-being, and adaptive behaviour.

Summary

This section has reviewed the main subjective indicators of thermal experience. Core constructs such as TSENS, TPREF, TAC, and TSAT capture different levels of occupants' thermal responses, while additional dimensions including TCV, TEXP, PCON, and OCOM extend the scope to cognitive, behavioural, and integrative aspects.

Together, these indicators highlight that thermal comfort and satisfaction are multidimensional outcomes shaped not only by physical parameters of the environment but also by expectations, perceived control, and broader evaluations. Past studies consistently show that subjective impacts cannot be fully explained by environmental conditions alone, but rather emerge from the interaction between environmental, psychological, and contextual factors. This overview provides the basis for the following sections on measurement methods and thermal comfort models, where these subjective dimensions are operationalized into tools for building design and evaluation.

By situating these dimensions within a coherent framework, this section provides a foundation for subsequent discussions of measurement approaches and thermal comfort models, where the challenge is to translate subjective evaluations into reliable tools for design, assessment, and operation of indoor environments.

2.2.3 Effects on physiological and health-related functions (physiological impacts)

Thermal environments exert fundamental effects on the human body, shaping not only how occupants feel but also how their physiological systems function and adapt. While thermal comfort is usually framed as a subjective state of mind, it is underpinned by biological mechanisms that maintain thermal homeostasis. Deviations from the thermoneutral zone activate regulatory processes designed to preserve core temperature, but these responses inevitably place additional load on cardiovascular, metabolic, and other organ systems [89]. If exposures are prolonged or extreme, the consequences extend beyond temporary strain to measurable health risks, including impaired sleep, reduced functional capacity, and increased morbidity and mortality [90].

A substantial body of evidence across laboratory experiments, occupational studies, and epidemiological research has underscored these risks. Laboratory chamber studies provide precise insights into the thresholds and dynamics of thermoregulatory responses [91]. Occupational research in agriculture, construction, and manufacturing demonstrates how heat or cold exposures compromise endurance, hydration, and safety [92]. Public health investigations, particularly those examining heat waves and under-heated housing, reveal strong links between thermal stress and seasonal mortality, especially among vulnerable groups such as the elderly, children, and individuals with chronic illness [93].

Understanding these physiological effects is crucial for two reasons. First, they provide the biological substrate upon which subjective comfort experiences (Section 2.2.2) are formed, meaning that perceptions of comfort cannot be separated from bodily strain. Second, they highlight the safety and health dimension of thermal environments: even conditions that occupants consider “tolerable” may silently increase cardiovascular load or elevate long-term health risk [94]. In the following subsections, three interrelated domains are reviewed: thermoregulatory responses, cardiovascular and metabolic strain, and health outcomes.

Thermoregulatory responses

The human thermoregulatory system represents the first line of defense against thermal stress. In hot environments, vasodilation, sweating, and increased cutaneous blood flow promote heat dissipation, whereas in cold environments, vasoconstriction, shivering, and heightened metabolic heat production conserve core temperature [95]. Classic climatic chamber studies, which manipulated temperature and humidity under controlled conditions, established systematic thresholds for sweating onset, shivering, and changes in skin temperature [96].

Later research highlighted the role of personal factors such as metabolic rate and clothing insulation. Individuals with higher activity levels or heavier clothing can delay the onset of thermal strain but may also accumulate greater metabolic or cardiovascular load in the process [44]. Acclimatization adds another layer: populations in tropical climates initiate sweating at lower ambient temperatures than those in temperate or cold climates, reflecting adaptive shifts in thermoregulatory sensitivity [97]. Similarly, seasonal acclimatization affects thermal tolerance, with cold-acclimatized individuals showing blunted shivering responses compared to those unaccustomed to low temperatures.

Recent advances have extended these findings into applied contexts. For instance, wearable sensors in field studies have documented real-time changes in skin temperature and sweating among outdoor workers, linking physiological strain to productivity loss [44]. In indoor settings, subtle thermoregulatory responses such as peripheral vasoconstriction can occur even at temperatures occupants report as "acceptable," suggesting that comfort surveys may underestimate hidden physiological stress [98].

Cardiovascular and metabolic strain

Beyond immediate thermoregulation, thermal environments impose notable demands on cardiovascular and metabolic systems. Heat exposure elevates heart rate and redistributes blood flow toward the skin, while cold exposure raises blood pressure and stimulates metabolic heat production [99]. Controlled experiments have quantified these effects, showing that even moderate deviations from thermoneutrality increase cardiovascular load, oxygen consumption, and energy expenditure [100].

Epidemiological studies provide compelling evidence of these physiological burdens in real-world contexts. Seasonal mortality peaks, particularly in winter, are strongly associated with cold-induced cardiovascular events such as myocardial infarction and stroke [101]. Conversely, heat waves correlate with spikes in hospital admissions for dehydration, arrhythmias, and renal stress [102]. Occupational research confirms similar patterns: construction and agricultural workers facing high heat stress exhibit reduced endurance, higher dehydration risk, and increased accident rates [103].

Importantly, subclinical strain often precedes overt illness. Elevated blood pressure in cold environments or sustained tachycardia in hot conditions can persist without immediate symptoms, yet still contribute to long-term health deterioration. Chronic exposure to such environments may exacerbate pre-existing conditions, particularly among elderly individuals or those with compromised cardiovascular function. This underscores the need to evaluate not only comfort but also the hidden physiological costs of thermal environments.

Health outcomes

When thermoregulatory and cardiovascular responses are overwhelmed, more severe health consequences can arise. Acute risks include heat exhaustion, heat stroke, dehydration, hypothermia, and frostbite, all of which are well documented in both field investigations and clinical case records [104]. Public health studies consistently demonstrate excess mortality during extreme heat events, with the elderly, socially isolated individuals, and those with chronic illness being disproportionately affected [105].

Cold environments also pose substantial risks. Research on under-heated housing in colder regions links low indoor temperatures to higher rates of respiratory infections, cardiovascular events, and excess winter deaths [106]. Chronic cold exposure has been associated with increased arterial stiffness and hypertension, further highlighting the compounding effects of suboptimal thermal environments.

Even in less extreme conditions, inadequate indoor climate control can impair health and daily functioning. Studies in hot and humid climates show that suboptimal cooling disrupts sleep quality, contributes to chronic fatigue, and reduces cognitive recovery [107, 108]. Similarly, insufficient heating during winter has been linked to musculoskeletal discomfort and exacerbation of chronic respiratory illnesses [109]. These findings highlight that "everyday" thermal exposures, not just extremes, shape long-term health trajectories.

Summary

In summary, the physiological impacts of thermal environments encompass thermoregulatory adjustments, cardiovascular and metabolic strain, and a spectrum of health outcomes ranging from acute to chronic. Evidence from laboratory, occupational, and epidemiological studies converges to show that these effects are substantial, context-dependent, and often underappreciated in comfort-focused research. Together, they demonstrate that thermal comfort studies must integrate physiological and health evidence alongside subjective evaluations (Section 2.2.2), ensuring that indoor environments are designed not only for comfort but also for safety, resilience, and long-term well-being.

2.2.4 Effects on cognitive performance (cognitive impacts)

Thermal environments affect not only thermal sensation and physiological strain but also the efficiency of human cognition. Cognitive performance encompasses a spectrum of processes, attention, memory, reasoning, and decision-making, that are fundamental for everyday functioning, learning, and occupational safety. Unlike physiological strain, which can often be directly measured, cognitive impairment may remain unnoticed until it manifests in errors, slowed responses, or reduced productivity [110]. This makes cognitive outcomes both subtle and consequential.

Evidence from psychology, ergonomics, and education has shown that deviations from thermal neutrality can degrade performance even under moderate, everyday conditions. Heat stress is commonly associated with lapses of vigilance, reduced working memory, and impaired decision-making [111]. Cold exposure, while sometimes producing short bursts of alertness, typically undermines consistency, fine motor control, and problem-solving efficiency [112]. Importantly, these effects scale up: from individual errors in attention to collective impacts on learning outcomes or workplace productivity.

To capture the breadth of this evidence, the following subsections examine four domains of cognitive impact: attention and vigilance, memory and learning, executive function and decision-making, and productivity in real-world contexts. Each domain highlights a distinct pathway through which thermal environments shape mental performance.

Attention and vigilance

Attention is often the first cognitive function to deteriorate when thermal stress is present. Sustained vigilance requires stable cortical arousal, yet both heat and cold disrupt this balance. High temperatures accelerate fatigue, increasing the likelihood of attentional lapses, while cold discomfort competes for cognitive resources, producing variability in reaction times [113].

Experimental evidence supports these mechanisms. Subjects performing simple reaction time or signal-detection tasks in hot conditions typically respond more slowly and commit more errors, especially during prolonged exposure [114]. Cold stress produces a different pattern: initial hyper-arousal may briefly sharpen responses, but as discomfort accumulates, performance becomes inconsistent, and error rates rise. Psychophysiological studies using EEG measures confirm these patterns, showing reduced alertness markers in heat and fluctuating arousal under cold [115].

Applied research highlights the practical consequences. In transportation, drivers in overheated vehicles demonstrate delayed hazard detection and longer braking times [116]. Surveillance operators in poorly conditioned control rooms report more frequent lapses when monitoring monotonous video feeds [117]. Such findings underscore that vigilance tasks, especially monotonous or sustained ones, are disproportionately vulnerable to thermal stress.

Notably, not all individuals are equally affected. Children, whose attentional systems are still developing, and older adults, who often have reduced cognitive resilience, appear more susceptible to distraction under thermal stress [118]. This suggests that building thermal environments should be designed with particular sensitivity to populations engaged in sustained attention tasks, whether in classrooms, offices, or safety-critical operations.

Memory and learning

Memory and learning are longer-term cognitive processes with profound educational and developmental implications. Heat stress has been shown to impair the encoding and retrieval of information, while cold exposure diverts attention toward discomfort, reducing cognitive capacity for memory consolidation [112].

Laboratory studies provide clear evidence. In controlled conditions, subjects exposed to elevated temperatures recall fewer items in working-memory tasks and perform worse on tests of verbal recall and numerical manipulation [119]. Neuroimaging research suggests that heat reduces activation in prefrontal and hippocampal regions associated with memory processing [120]. Cold exposure, by contrast, often produces fragmented recall patterns, reflecting divided attention rather than direct memory impairment [121].

Educational research illustrates the cumulative impact of these effects. Large-scale analyses of standardized test performance show that students taking exams during heat waves achieve significantly lower scores, even after adjusting for socioeconomic status [122]. Longitudinal studies further demonstrate that repeated exposure to overheated classrooms contributes to widening academic disparities over time [123]. Conversely, insufficient winter heating has been linked to reduced concentration and poorer recall, with children in under-heated schools reporting greater difficulties in sustained study [124].

These findings carry equity implications. Students from low-income households are more likely to attend schools with inadequate climate control, meaning that thermal stress can exacerbate existing educational inequalities. Thus, the cognitive impacts of thermal environments extend beyond individual performance, shaping broader patterns of educational opportunity and achievement.

Executive function and decision-making

Executive functions, such as planning, reasoning, inhibition, and decision-making, represent higher-order cognitive capacities that are particularly sensitive to stressors. Heat reduces inhibitory control and increases reliance on heuristic shortcuts, while cold tends to slow cognitive flexibility and impair problem-solving efficiency [112].

Laboratory research has demonstrated that subjects in hot conditions are more likely to make impulsive choices, favouring immediate but suboptimal rewards, while cold exposure prolongs deliberation and reduces accuracy in multi-step reasoning tasks [125]. Complex dual-task paradigms further show that both heat and cold reduce coordination efficiency, reflecting diminished executive control [126].

Applied studies underscore the stakes of these impairments. Military research indicates that soldiers operating in hot environments commit more tactical planning errors and respond more slowly to dynamic threats [127]. In healthcare, staff working in overheated wards show increased documentation mistakes and slower diagnostic decisions, issues with direct consequences for patient safety [128]. Industrial operators and pilots under thermal stress also experience reduced situational awareness, increasing the risk of errors in high-stakes operations [129].

Executive function impairments highlight the importance of thermal design in domains where decision quality is paramount. Unlike attention or memory, where deficits are often incremental, failures in executive control can produce disproportionate risks, from misdiagnoses in medical settings to accidents in industrial or transportation contexts.

Productivity and real-world outcomes

While the previous sections focus on specific cognitive functions, their combined effects converge in broader productivity outcomes. Declines in vigilance, memory, and executive control translate into measurable reductions in efficiency, accuracy, and safety. These effects are not confined to individuals but accumulate across organizations and societies.

Office studies show that workers in thermally stressful environments type more slowly, make more errors, and report higher fatigue, even when they subjectively describe conditions as "acceptable" [130]. Inadequate cooling during summer or insufficient heating during winter both reduce output, with meta-analyses estimating significant productivity losses attributable to suboptimal indoor climates [131].

Educational contexts reveal similar patterns. Poor classroom conditions reduce not only test performance but also day-to-day engagement, participation, and motivation [132]. Over time, these effects contribute to disparities in educational attainment, with implications for workforce readiness and social equity.

In industrial and agricultural settings, the stakes are even higher. Heat stress has been linked to reduced work pace, increased accident rates, and greater incidence of errors, combining human risks with economic costs [133]. At the societal level, cross-national estimates suggest billions of dollars in annual productivity losses due to thermal stress, reinforcing the urgent need to integrate cognitive considerations into building standards, occupational safety regulations, and public health planning.

Summary

Thermal environments shape cognition in multifaceted ways: undermining vigilance in monotonous tasks, disrupting memory and learning in educational contexts, impairing executive control in complex operations, and reducing productivity at organizational and societal scales. Unlike physiological strain, which is often visible, cognitive impairments may remain hidden until they manifest in errors, accidents, or long-term achievement gaps. Taken together with the subjective (Section 2.2.2) and physiological (Section 2.2.3) dimensions, these findings establish cognitive performance as a critical yet often underappreciated component of thermal comfort research. Designing thermal environments that safeguard cognitive functioning is therefore essential for supporting well-being, educational opportunity, productivity, and public safety.

2.2.5 Effects on emotional and affective states (emotional impacts)

Thermal environments influence not only subjective comfort (Section 2.2.2), physiological strain (Section 2.2.3), and cognitive functioning (Section 2.2.4), but also the emotional and affective states that shape human experience and behaviour. Emotions are central to psychological well-being, motivation, and social interaction. They mediate how people perceive their surroundings, respond to stressors, and engage with others. Unlike discrete cognitive outcomes, emotional impacts are often diffuse and multifactorial, emerging from both direct physiological pathways and indirect psychological mechanisms.

Research across environmental psychology, occupational health, and social sciences shows that deviations from thermal neutrality can alter mood, increase irritability, elevate perceived stress, and influence social dynamics. These emotional outcomes matter because they affect not only individual well-being but also interpersonal relationships, organizational culture, and even community resilience during extreme climate events.

The following subsections highlight four interrelated domains of emotional impact: general mood and affect, stress and irritability, motivation and engagement, and social interaction.

Mood and affect

Thermal conditions have a direct influence on general mood states. Exposure to heat is consistently associated with negative affect, including feelings of fatigue, tension, and reduced vigour [26]. Cold exposure can elicit similar declines in positive affect, often accompanied by discomfort and withdrawal tendencies [134]. Laboratory studies using standardized mood inventories have demonstrated that both heat and cold shift affective profiles toward greater negativity, even when subjects do not consciously report dissatisfaction with the environment [135].

These findings align with broader theories of environmental affect, which suggest that discomfort acts as a background stressor, subtly colouring emotional tone and reducing positive affective states. Real-world evidence supports this: in classrooms, students in overheated rooms report feeling less motivated and more tired, while office workers in under-heated spaces describe decreased enthusiasm and higher levels of discontent [136]. Such affective shifts have implications for both well-being and productivity.

Stress and irritability

Thermal stress is not only a physical strain but also an emotional one. Heat exposure elevates physiological arousal, which can manifest emotionally as irritability, frustration, or heightened stress perception [137]. Cold environments, while less often associated with aggression, can increase tension and reduce emotional stability when discomfort persists [138].

Experimental research shows that subjects in hot conditions report greater irritability and lower tolerance for frustration during problem-solving tasks [139]. Field studies further reveal that crime rates and aggressive incidents tend to rise during heat waves, suggesting a link between thermal stress and social aggression [140]. In workplaces, overheated conditions have been linked to more interpersonal conflicts and reduced patience among staff.

Importantly, these effects are not solely physiological. Psychological appraisal plays a role: when individuals attribute discomfort to external, uncontrollable factors (e.g., poor building design), emotional responses are amplified. This illustrates how thermal stress interacts with perceptions of control and fairness, shaping affective outcomes.

Motivation and engagement

Thermal environments also influence motivational states and willingness to engage in tasks. Heat often reduces intrinsic motivation, leading to perceptions of effortfulness and diminished persistence [141]. Cold exposure, while sometimes producing short bursts of alertness, can reduce willingness to sustain engagement, leading to task withdrawal [142].

In educational settings, overheated classrooms have been shown to lower student participation and reduce self-reported motivation to learn [143]. In workplaces, employees in thermally uncomfortable environments are more likely to disengage, avoid optional tasks, or postpone effortful activities [144]. These motivational effects may partly explain why productivity losses under thermal stress cannot be fully accounted for by cognitive impairments alone: emotional disengagement is an additional pathway.

Vulnerable groups, such as children or individuals with pre-existing mental health conditions, may be particularly sensitive to these motivational impacts, underscoring the importance of designing thermal environments that support not only cognitive but also affective engagement.

Social interaction and interpersonal dynamics

Emotional responses to thermal conditions extend beyond individuals, shaping social behaviour and interpersonal relations. Heat stress has been linked to increased aggression and reduced prosocial behaviour, while cold environments may promote withdrawal and reduced social interaction [145].

Field studies during heat waves report more frequent conflicts in communities and elevated rates of interpersonal violence [140]. In organizational contexts, overheated offices are associated with higher levels of irritability and reduced cooperation among colleagues [146]. Conversely, thermally comfortable environments tend to foster more positive interactions, greater group cohesion, and enhanced satisfaction with social exchanges [147].

These findings point to a broader social dimension of thermal comfort: emotional and affective states induced by the thermal environment can scale up to influence group dynamics, workplace culture, and community resilience. In settings where teamwork and collaboration are essential, maintaining emotional stability through adequate thermal management becomes a critical design consideration.

Summary

In summary, thermal environments affect not only cognition and physiology but also emotional and affective states. Heat and cold can alter mood, increase irritability, reduce motivation, and shape social interactions. These effects have implications for individual well-being, group dynamics, and broader societal outcomes such as aggression rates and community harmony. Emotional impacts represent a key but sometimes overlooked dimension of thermal comfort research. Addressing them is essential for creating environments that support not only physical health and cognitive performance but also psychological well-being and positive social relations.

2.2.6 Summary

Thermal environments exert multifaceted influences on human beings, extending well beyond simple sensations of heat or cold. As outlined in the preceding subsections, these impacts encompass subjective comfort, physiological regulation, cognitive functioning, and emotional well-being, each contributing to the overall human experience of thermal conditions.

First, subjective responses (Section 2.2.2) reflect how individuals consciously perceive their environment, shaping satisfaction, preference, and behavioral adaptation. These evaluations are immediate and intuitive, but they are also highly variable across individuals and contexts. Second, physiological outcomes (Section 2.2.3) highlight the body's continuous regulatory efforts. Heat and cold stress disrupt thermal balance, imposing cardiovascular and metabolic strain. These responses are not only indicators of physical health but also directly influence comfort and performance. Third, cognitive impacts (Section 2.2.4) demonstrate that thermal conditions affect higher mental processes, from vigilance and memory to executive control. Even when occupants report acceptable comfort, subtle decrements in attention, accuracy, and decision-making can occur, with implications for learning, productivity, and safety. Fourth, emotional and affective states (Section 2.2.5) reveal another critical dimension: thermal stress can alter mood, increase irritability, reduce motivation, and shape social interactions. These affective responses link the thermal environment to interpersonal dynamics, organizational climate, and even community-level outcomes.

Taken together, these findings show that the effects of thermal environments are multidimensional and interdependent. Subjective comfort, physiological strain, cognitive performance, and emotional states interact dynamically: for example, discomfort can heighten stress, which in turn undermines attention and social cohesion. Recognizing these interconnections is essential for developing a holistic understanding of thermal comfort and its broader consequences. From a research perspective, this synthesis indicates that thermal environments influence not only thermal comfort but also the task cognitive–performance domain, encompassing both subjective workload sensation and objective task performance. This dual influence underscores the need for integrated approaches that jointly consider comfort and performance, thereby providing the foundation for the subsequent discussion of task factors and their interaction with thermal conditions.

2.3 Effects of task factors

Task factors are another fundamental dimension shaping human experience and performance in indoor and occupational settings. Whereas the thermal environment influences comfort, physiology, cognition, and emotional states, task characteristics determine the demands placed on the body and mind and the consequences of human errors. A growing body of research demonstrates that features such as workload, complexity, duration, pacing, and risk level affect not only physical strain but also cognitive functioning, motivation, fatigue, and safety outcomes. This section reviews these effects across five main dimensions. Section 2.3.1 defines and clarifies the scope of task factors. Section 2.3.2 examines their impacts on physiological load and physical strain. Section 2.3.3 discusses effects on cognitive functioning and task performance. Section 2.3.4 considers emotional and motivational states, while Section 2.3.5 focuses on fatigue and recovery. Section 2.3.6 addresses safety and risk consequences. Finally, Section 2.3.7 summarizes the key findings of this part.

2.3.1 Definition and scope of task factors

Task factors refer to the intrinsic characteristics of the activities that individuals perform, which determine the nature and magnitude of the demands placed on the body and mind during task execution. Unlike environmental conditions, which originate from external physical surroundings, task factors are embedded in the structure and requirements of the work itself: what must be done, how it is organized, and under what constraints. They directly shape the allocation of physiological and psychological resources, influencing not only immediate performance but also long-term health, safety, and well-being.

A number of attributes are commonly identified as central to task factors. Workload and intensity describe the degree of physical exertion or metabolic demand associated with a task. For example, heavy manual labour generates sustained muscular activity and elevated energy expenditure, while sedentary clerical work imposes minimal physical load [148]. Complexity and cognitive demand capture the information-processing requirements of tasks, distinguishing between simple, repetitive operations and multifaceted activities that involve decision-making, problem-solving, or multitasking [149]. Duration and pacing relate to the temporal structure of work: some tasks are short and self-paced, allowing intermittent rest, while others are extended or externally paced, leaving little opportunity for recovery [150]. Criticality and risk level concern the seriousness of potential consequences when errors occur, minor mistakes in routine office tasks may be tolerated, whereas lapses in surgical procedures, air-traffic control, or nuclear plant operations can have severe outcomes [151]. Finally, affective and motivational characteristics reflect how tasks are experienced emotionally: monotonous and repetitive activities may induce boredom and disengagement, whereas challenging and meaningful work can foster motivation and satisfaction [152].

The scope of task factors considered in this section is restricted to those features that exert direct, short- to medium-term impacts on human performance and well-being. Broader organizational and social aspects of work design, such as management structures, leadership practices, or compensation systems, are important in their own right but fall outside the present discussion, as their impacts are more indirect and long-term [153]. Instead, the emphasis here is on how the immediate features of tasks themselves influence physiological load, cognitive functioning, emotional states, fatigue dynamics, and safety outcomes.

Recognizing the scope of task factors is important for two reasons. First, these factors help explain why individuals exposed to the same environmental conditions may experience very different levels of strain and performance outcomes: a worker engaged in heavy lifting will perceive and respond to the indoor environment differently from a colleague engaged in data entry at the same workstation [154]. Second, they highlight that human performance and well-being are shaped not only by environmental quality but also by the interaction between task demands and individual capacities [155]. Understanding the definition and scope of task factors therefore provides the foundation for systematically examining their multifaceted effects in the subsequent subsections.

2.3.2 Effects on physiological load and physical strain

Task factors exert some of their most immediate and tangible effects on the body by altering physiological load and physical strain. Unlike environmental influences, which act externally, task demands generate endogenous loads: they determine how much energy the body must produce, how muscles and joints are engaged, and how long recovery opportunities are available [156]. These internal demands interact with thermoregulation, circulation, and musculoskeletal systems, thereby shaping both immediate comfort and long-term health outcomes [157].

Workload and metabolic rate

The overall workload of a task is a primary determinant of physiological load. High-intensity tasks such as lifting, carrying, or repetitive manual work can drive metabolic rates to several times the resting level, resulting in elevated heart rate, oxygen uptake, and body heat production [100]. By contrast, sedentary tasks such as computer operation or clerical work demand little metabolic energy, but may still impose strain through static muscular activation and reduced circulation. The metabolic cost of different task categories is so fundamental that occupational standards often classify work intensity in metabolic equivalents (met) for health and ergonomics evaluations [158].

Experimental and field studies consistently show that tasks with higher metabolic rates accelerate fatigue and increase the likelihood of heat strain in warm conditions, while sedentary tasks heighten vulnerability to localized cold stress or discomfort in cooler settings [90]. For example, construction workers engaged in continuous heavy lifting may experience rapid rises in core temperature and sweating, whereas office workers in the same environment may complain of cold extremities due to low heat production. This illustrates how workload shapes baseline physiological states and modulates environmental sensitivity.

Posture, movement, and musculoskeletal strain

The biomechanics of task execution, posture, frequency of movement, and force exertion, are another critical source of strain. Prolonged sitting requires low energy expenditure but results in static loading of the lumbar and cervical spine, contributing to back pain and neck stiffness [148]. Similarly, tasks requiring prolonged standing can cause venous pooling in the legs, swelling, and discomfort. These outcomes occur even when the metabolic workload is modest, demonstrating that musculoskeletal strain is not solely determined by energy cost.

Repetitive tasks, such as assembly line operations or typing, place cumulative stress on specific muscle groups and tendons. Without sufficient variation or recovery, these repetitive motions can lead to localized muscle fatigue, inflammation, and eventually chronic musculoskeletal disorders [150]. Occupational health research has long documented elevated rates of wrist, shoulder, and lower-back problems in jobs that combine repetitive actions with awkward postures. Thus, posture and motion patterns constitute an independent dimension of task-related physical strain.

Duration and pacing of physical effort

The temporal pattern of work must be sustained and at what pace, further determines physiological load. Continuous high-intensity activity rapidly depletes energy reserves, elevates cardiovascular strain, and increases core temperature, heightening the risk of exhaustion and dehydration [157]. Even at moderate intensity, if sustained for extended periods without breaks, physiological systems accumulate strain that manifests as fatigue, reduced performance capacity, and discomfort.

Pacing strongly influences these outcomes. Self-paced tasks allow the worker to adjust effort according to their capacity, facilitating natural micro-breaks and partial recovery. In contrast, externally paced tasks, such as conveyor-belt work or time-pressured services, constrain recovery opportunities and amplify strain, even when absolute workload is moderate [150]. Field studies have shown that workers on tightly scheduled assembly lines report higher fatigue and musculoskeletal complaints than those in comparable but self-paced jobs. Thus, both duration and pacing critically shape how physical strain accumulates over time.

Health outcomes and long-term risks

The cumulative impact of task-related physical demands extends beyond immediate comfort to long-term health consequences. High physical workloads, particularly when combined with unfavorable postures or insufficient recovery, are linked to increased incidence of musculoskeletal disorders, cardiovascular strain, and chronic fatigue syndromes [150]. Occupations involving heavy manual handling, for example, display persistently higher rates of back injuries and early work disability compared with sedentary roles.

Conversely, extremely low workloads, characteristic of prolonged sedentary office work, have their own risks. Low energy expenditure reduces metabolic health, contributing to obesity, insulin resistance, and elevated risk of cardiovascular disease [159]. Epidemiological studies increasingly recognize sedentary behavior itself as an independent health hazard, even among individuals who meet physical activity guidelines. Thus, both extremes of task demand, excessively high and excessively low, can undermine health, albeit through distinct physiological mechanisms.

Summary

In sum, task factors shape physiological load and physical strain through multiple, interacting pathways: the energy cost of the task, the biomechanics of posture and movement, and the temporal structure of effort. These influences manifest acutely as variations in fatigue, discomfort, and thermal sensitivity, and chronically as musculoskeletal and metabolic health risks. Importantly, they establish the baseline physiological state upon which environmental conditions act, thereby conditioning how individuals perceive and respond to thermal environments. Understanding task-related physical load is therefore essential for any comprehensive assessment of comfort and performance.

2.3.3 Effects on affective states, motivation, and fatigue

Task factors not only shape physical and cognitive demands but also strongly influence affective experience, motivational engagement, and the dynamics of fatigue and recovery. These dimensions are closely intertwined: affective states such as boredom or stress affect motivational drive; motivation in turn determines the level of effort invested; and sustained effort leads to fatigue and the need for recovery. Together, they determine not only how people feel about their tasks but also how effectively they can perform them under varying conditions [17].

Task monotony, boredom, and disengagement

Monotonous or repetitive tasks often lead to boredom, a state characterized by under-stimulation, lack of interest, and psychological discomfort. Boredom reflects a mismatch between the individual’s need for stimulation and the low informational content of the task. It reduces intrinsic motivation, undermines vigilance, and creates a sense of time dragging, which can make even simple tasks feel effortful [17].

Empirical studies in industrial and office settings show that monotonous work is consistently associated with higher error rates, slower responses, and increased reports of mental fatigue compared with varied or stimulating tasks [160]. Workers in repetitive assembly line jobs, for instance, often report greater dissatisfaction and higher turnover intention. Monotony can also magnify awareness of unfavorable environmental conditions: individuals who are bored or disengaged are more likely to notice and complain about thermal discomfort than those absorbed in meaningful activities [161].

Challenge, motivation, and engagement

In contrast to monotony, tasks that provide an optimal level of challenge tend to foster engagement and motivation. When work is perceived as meaningful and appropriately demanding, individuals experience heightened focus, persistence, and satisfaction [162]. This motivational state supports sustained performance and can buffer the perception of external discomfort, as motivated individuals often exhibit greater tolerance toward adverse conditions.

However, the benefits of challenge are conditional. When demands exceed available resources, tasks that are too difficult or time-pressured may shift from motivating to stressful. Research shows that excessive challenge elevates error rates, increases stress hormones, and accelerates the onset of mental fatigue [119]. The relationship between challenge and motivation thus resembles an inverted-U: insufficient challenge leads to boredom, moderate challenge maximizes engagement, and excessive challenge generates stress and disengagement.

Stress, stress, and affective strain

Some tasks carry inherently high stakes, such as surgical operations, military missions, or emergency response activities. These contexts often elicit stress and stress, activating physiological responses like elevated heart rate, blood pressure, and cortisol release. While acute stress can sharpen attention and promote rapid responses, sustained stress undermines working memory, narrows attentional focus, and increases susceptibility to errors [119].

Chronic exposure to high-stress tasks contributes to long-term affective strain, which is associated with burnout, sleep disturbances, and mental health issues. For instance, healthcare workers and first responders frequently report high levels of emotional exhaustion and depersonalization, reflecting the cumulative toll of stress-laden work [152]. Thus, task-related stress has both immediate and enduring consequences for individual well-being and organizational safety.

Fatigue accumulation and recovery

Sustained cognitive effort and emotional strain inevitably lead to fatigue, a multidimensional state involving reduced alertness, slower reaction times, impaired decision-making, and subjective tiredness. Fatigue reflects the depletion of psychological and physiological resources, and its onset is accelerated by tasks that are monotonous, highly demanding, or emotionally stressful [152].

Recovery processes are therefore critical in managing fatigue. Short breaks, task variation, and supportive motivational contexts have been shown to restore attentional capacity and reduce subjective fatigue. Conversely, insufficient recovery leads to cumulative fatigue, where deficits carry over across tasks or work shifts [152]. Over time, chronic fatigue increases the risk of burnout, cardiovascular health problems, and diminished job performance. Effective task design thus requires balancing demands with opportunities for rest and recovery to sustain both well-being and productivity.

Summary

Affective states, motivation, and fatigue represent interconnected dimensions through which task factors shape human performance. Monotony induces boredom and disengagement; optimal challenge enhances motivation and engagement; excessive demands create stress and stress; and prolonged effort produces fatigue that necessitates recovery. These processes influence not only task performance but also subjective perceptions of comfort and tolerance of environmental conditions. In the context of thermal environments, affective and motivational factors may moderate how individuals perceive thermal stressors and how long they can maintain performance under suboptimal conditions. Incorporating these dimensions into the analysis of task factors is therefore vital for a comprehensive understanding of human performance in occupational and indoor settings.

2.3.4 Effects on fatigue and recovery

Fatigue is one of the most widely documented consequences of sustained task demands. It can be defined as a multidimensional state of reduced capacity, expressed in diminished alertness, slower responses, impaired decision-making, and subjective tiredness [163]. Recovery, by contrast, denotes the restoration of depleted resources, enabling individuals to regain functional capacity. Together, fatigue and recovery form a dynamic process that mediates the relationship between task factors, human performance, and well-being [164].

Mechanisms of fatigue development

Research distinguishes between physiological, cognitive, and emotional components of fatigue. Physiological fatigue results from sustained muscular activity, metabolic strain, or cardiovascular load, leading to reduced efficiency and endurance [90].

Cognitive fatigue arises when attentional and working memory resources are continuously taxed, producing declines in accuracy and vigilance [111]. Emotional fatigue develops under prolonged exposure to stress and affectively demanding tasks, draining motivational resources and contributing to exhaustion [152]. Although analytically distinct, these forms of fatigue often co-occur, confirming that fatigue represents a systemic outcome of task demands.

Empirical evidence of fatigue accumulation

Experimental research consistently demonstrates that performance deteriorates with increasing time-on-task. Vigilance studies report progressive declines in detection accuracy, longer reaction times, and reduced subjective alertness during extended monitoring tasks [111]. Cognitive experiments similarly show that sustained working memory or problem-solving tasks lead to measurable decrements in accuracy and efficiency after prolonged engagement [111].

Field studies provide convergent evidence. Investigations of extended work hours across diverse occupational contexts show that fatigue accumulation is associated with higher error rates, reduced productivity, and increased accident risk [165]. Longitudinal analyses further indicate that repeated exposure to demanding schedules without sufficient recovery exacerbates performance decline and elevates health risks [152]. Together, these findings establish fatigue accumulation as a robust and generalizable effect of sustained task demands.

Recovery processes and moderating factors

Recovery processes counteract fatigue by replenishing depleted resources. Empirical evidence shows that short breaks restore attentional capacity and reduce subjective fatigue [42, 152]. Task variation and alternating between different forms of activity have also been found to mitigate declines in performance by engaging distinct resource pools [166]. Sleep has consistently been identified as the most effective recovery mechanism, restoring both physiological and cognitive functioning [108].

The effectiveness of recovery is moderated by task characteristics and individual factors. Research indicates that monotonous tasks require more frequent breaks to sustain performance, while tasks that are highly engaging may delay fatigue onset but subsequently result in more pronounced exhaustion [152]. Individual differences such as age, health, and prior training also influence recovery trajectories [167]. Insufficient recovery leads to cumulative fatigue, a condition documented in longitudinal studies as a predictor of chronic health problems, burnout, and long-term performance decrements [152].

Summary

In sum, fatigue and recovery represent dynamic, interdependent processes through which task factors influence both short-term and long-term outcomes. Fatigue develops through physiological, cognitive, and emotional pathways, and has been consistently observed across experimental and field studies. Recovery processes, ranging from short breaks to sleep, are critical in restoring depleted resources, but their effectiveness depends on task conditions and individual characteristics. Without adequate recovery, cumulative fatigue emerges, undermining performance and well-being over time. These insights highlight the importance of incorporating fatigue and recovery into any comprehensive analysis of task factors and their interaction with environmental conditions.

2.3.5 Effects on safety and risk consequences

Safety and risk outcomes represent a critical dimension of task factors, as the characteristics of work directly influence both the probability of human error and the severity of its consequences. Human error has been identified as a dominant cause of accidents across a wide range of domains, and task demands are consistently highlighted as primary determinants of error occurrence [152]. By shaping workload, attentional allocation, stress, and fatigue, task factors act as proximal mechanisms through which safety is either maintained or compromised.

Error likelihood and task demands

A consistent theme in the literature is that error likelihood follows a U-shaped relationship with task demand. On one hand, under-load conditions, such as monotonous or low-stimulation tasks, are associated with vigilance decrements and attentional lapses. Controlled vigilance experiments demonstrate that detection accuracy declines steadily as task monotony increases, with error rates rising significantly over time-on-task [111]. Field studies further confirm that low cognitive demand tasks are accompanied by higher rates of attentional failures and omissions [42].

On the other hand, over-load conditions elevate the risk of mistakes by exceeding cognitive capacity. Laboratory research on complex decision-making tasks shows that as informational load grows, subjects' error rates increase and decision quality deteriorates [168]. Time pressure studies similarly demonstrate that compressed time windows reduce accuracy, as individuals rely on heuristics rather than deliberate reasoning [150]. Meta-analyses of workload and human error suggest that both excessively low and high demands lead to suboptimal safety outcomes, supporting the notion of an optimal workload zone for minimizing errors [165].

Consequence severity and task criticality

The severity of safety outcomes is determined not only by the frequency of errors but also by the criticality of the tasks in which they occur. Research on human error emphasizes that the same type of lapse can produce trivial or catastrophic consequences depending on task context [42]. Studies in safety-critical domains consistently report that even minor attentional failures or misjudgements can escalate into severe outcomes when tasks involve high responsibility or risk [42].

Comparative analyses across task categories support this differentiation. Empirical investigations show that error tolerance is higher in routine, low-stakes tasks, where mistakes can be corrected without major repercussions [42]. Conversely, in complex, high-risk systems, error tolerance is minimal, and the same magnitude of error has disproportionately greater consequences [151]. This body of research underscores that the safety implications of task factors cannot be evaluated solely in terms of error rates, but must also account for the potential severity of consequences embedded in task criticality.

Stress, fatigue, and safety outcomes

Stress and fatigue have been identified as key mediators linking task characteristics to safety risks. Experimental evidence shows that acute stress can initially enhance vigilance, but sustained or excessive stress impairs working memory, narrows attentional focus, and increases reliance on habitual responses, all of which elevate error likelihood [151]. Studies of cognitive performance under induced stress consistently demonstrate higher error rates and reduced flexibility in problem-solving [119].

Fatigue research provides equally strong evidence. Laboratory vigilance tasks reveal that prolonged time-on-task produces systematic declines in accuracy and response speed [111]. Field studies across multiple sectors consistently associate extended work hours with increased accident incidence and reduced operational safety [111, 169]. Longitudinal research further shows that cumulative fatigue, resulting from insufficient recovery across days or weeks, predicts higher injury rates and greater susceptibility to errors [152]. Reviews of fatigue and safety conclude that accumulated resource depletion is one of the most robust predictors of accidents, regardless of task domain [165].

Summary

In summary, task factors shape safety and risk through three primary pathways. First, they influence error likelihood, with both under-load and over-load conditions producing elevated error rates. Second, they determine consequence severity, as the criticality of tasks dictates whether errors remain minor or escalate into accidents with serious outcomes. Third, they operate through mediating mechanisms of stress and fatigue, which amplify risks by degrading attention, decision-making, and response capacity. The convergence of experimental, field, and longitudinal evidence underscores that safety cannot be understood in isolation from task characteristics. These insights provide an essential backdrop for exploring how task demands may interact with environmental stressors, such as thermal conditions, to shape overall safety and well-being.

2.3.6 Assessment of workload

At the same time, to systematically evaluate how task factors influence affective states, motivation, and fatigue, researchers have developed a range of workload assessment approaches. These methods can be broadly categorized into objective physiological measures and subjective self-report scales. Existing approaches to assessing workload are mainly categorised into objective and subjective approaches. The objective approach is primarily a measurement of human physiological parameters. Current research has found that many physiological parameters can be used to assess workload or mental effort. Heart activity has been regarded as one of the most common parameters which reflect workload [1]. It has been demonstrated that spectral measures of HRV are sensitive to the gap between task and rest by examining the literature [2]. Besides, skin parameters such as EDA and ST can also reflect workload or stress level [3], [4]. Although these studies have confirmed the use of physiological parameters as a reliable and direct measure of workload [5]. However, physiological parameters may not only respond to workload but may also vary as a result of the work itself, and it ignores the subjective experience of workload.

Subjective workload feelings differ from objective workload evaluation in that they can be incorporated into the subjective experience of the workload, which is a crucial part of the workload. And this subjective workload sensation can only be assessed by self-report. The most widely used research is the NASA task load index (NASA TLX) [6]. NASA TLX has been created by the Human Performance Group at NASA's Ames Research Centre for assessing the work stress level of astronauts. It assesses workload in six dimensions: mental demand, physical demand, temporal demand, effort, performance and frustration. Nowadays it is used far beyond its initial application areas (aviation, aerospace) and this evaluation system has been wildly used in different areas including healthcare, industrial production, etc [7], [8], [9]. For example, Hoonakker et al. applied the NASA TLX to evaluate the workload of ICU nurses [10]. Akyeampong et al. used this system to assess the impact of excavator's design on the workload [11]. There have also been studies using the NASA TLX to assess the cognitive workload of the driving environment in which the car is travelling on the driver [12]. Furthermore, since the NASA TLX is designed primarily for astronauts, there are also studies that have been adapted for specific environments. For instance, surgery task load index (SURG_TLX) designed for the evaluation of surgery activity [13]. And the driving activity load index (DALI) designed for the workload of driving [14]. Expect NASA TLX, there are a few other ways of assessing subjective workload, such as subjective workload assessment technique (SWAT) and workload profile (WP). However, all of these assessment methods are inferior to the NASA TLX application because of some limitations. For example, SWAT is not very sensitive to low-brain workloads [15]. WP is too sensitive to assessments for different task operations [16].

Similarly, in the field of thermal comfort research, the subjective workload evaluation has been used extensively when it comes to the effects of thermal comfort and thermal discomfort on mental workload, especially using the NASA TLX. For instance, In the study designed by Akimoto et al., NASA TLX was used to assess worker workloads in a study of the effects of the building thermal environment on productivity [17]. Belyamani et al. studied the local wearable cooling device's improvement in thermal comfort and workload cognition which applied NASA TLX to evaluation the workload cognition [18]. In Lan et al. study, the NASA TLX was used to assess the effects of temperature on office worker performance, productivity and stress levels [19]. There was also research used the NASA TLX for assessing the impact of office thermal discomfort on job performance [20].

2.3.7 Summary

Task factors exert broad and multifaceted influences on human functioning, extending beyond the immediate execution of work. As outlined in the preceding subsections, these influences span physiological, cognitive, affective, and safety-related domains, each shaping how individuals experience and respond to their tasks.

First, task demands impose physiological load and strain (Section 2.3.2), reflected in metabolic rate, musculoskeletal stress, and long-term health outcomes. Second, they generate cognitive and attentional demands (Section 2.3.3), which determine information-processing efficiency, vigilance, and error likelihood. Third, tasks influence affective and motivational states as well as fatigue dynamics (Section 2.3.4), with monotony, challenge, stress, and recovery processes governing engagement and resource depletion. Fourth, task characteristics are closely tied to safety and risk outcomes (Section 2.3.5), as both error probability and consequence severity are shaped by workload, criticality, and fatigue.

Taken together, these findings demonstrate that task factors are multidimensional determinants of human experience and performance. They affect the body’s physiological resources, the mind’s cognitive capacity, and the individual’s motivational and emotional states, with direct implications for productivity, safety, and well-being. Importantly, evidence also indicates that task factors influence not only task cognitive–performance domain outcomes but also thermal comfort, as workload, attentional focus, and motivational states alter how individuals perceive and tolerate thermal conditions. Recognizing this dual role establishes the rationale for examining the combined effects of thermal environments and task factors, which will be the focus of the following section.

2.4 Thermal environment and task factors: combined effects on thermal comfort, task cognitive–performance domain (TCPD) and stress

The preceding sections have reviewed the influences of the thermal environment (Section 2.2) and task factors (Section 2.3) when considered independently. Each of these domains has been shown to exert substantial effects on human physiology, cognition, affect, and safety. Yet in real-world contexts, thermal conditions and task demands seldom occur in isolation. Instead, they operate simultaneously and often interact in shaping how people feel, think, and perform. A comprehensive understanding of human responses therefore requires examining their combined effects, rather than treating the thermal environment and task factors as separate influences.

To address this need, the present section organizes the literature review around three outcome dimensions that together capture the range of human responses to the interplay between thermal environment and task demands. The first dimension is thermal comfort, which encompasses individuals' subjective evaluations of thermal sensation and thermal satisfaction. Thermal comfort reflects not only the direct influence of environmental conditions but also the way in which ongoing tasks shape perceptual salience and tolerance of thermal stimuli. The second dimension is the task cognitive–performance domain (TCPD), an integrative construct introduced in this review to capture both subjective workload sensation (e.g., perceived effort, mental demand, attentional load) and objective task performance (e.g., accuracy, speed, error rate, vigilance). By jointly considering workload sensation and performance outcomes, TCPD provides a comprehensive framework for assessing how task demands, and thermal conditions influence cognitive functioning and behavioural effectiveness. The third dimension is stress, treated here as a distinct outcome rather than a subcomponent of comfort or performance. Stress reflects both psychological strain (e.g., subjective pressure, tension) and physiological stress responses (e.g., cardiovascular activation, cortisol release). Importantly, stress can arise directly from thermal strain, directly from task demands, or from their interaction, and in turn it may mediate effects on thermal comfort and TCPD.

Together, these three dimensions, thermal comfort, TCPD, and stress, offer a structured lens for analysing how thermal environment and task factors combine to influence human experience and functioning. The following subsections review empirical evidence for each of these outcome dimensions, discuss integrative mechanisms, and identify research gaps that motivate the present study.

2.4.1 Combined effects on thermal comfort

Thermal comfort is not solely determined by environmental parameters such as air temperature, humidity, and air movement; it is also shaped by the characteristics of the task being performed [12]. Task-related factors such as workload intensity, posture, and cognitive engagement influence how individuals perceive thermal sensation and satisfaction [41]. For instance, tasks with higher metabolic demands elevate internal heat production and accelerate the onset of heat discomfort [47], whereas sedentary or low-intensity tasks generate less metabolic heat and therefore increase the likelihood of reporting cold discomfort [130]. Body posture further modifies thermal perceptions: prolonged sitting has been associated with reduced peripheral blood circulation and localized cooling in the extremities, which contributes to complaints of cold discomfort in office settings, while standing or moving tasks often facilitate greater heat dissipation but can also increase the perception of heat in warm environments [41]. These physiological interactions underline that the body's thermal balance is not only dependent on environmental exposure but also on the metabolic and postural requirements of the task at hand.

Beyond physical load, cognitive demands also modulate thermal perception. Several experimental studies have demonstrated that when individuals are highly engaged in demanding cognitive tasks, their attention to thermal sensations decreases, a phenomenon often described as the "distraction effect" [31]. In such cases, discomfort signals from the thermal environment are present but receive less cognitive processing, leading subjects to report lower levels of dissatisfaction even under suboptimal conditions. This attentional diversion is particularly evident in time-pressured or high-stakes tasks, where performance monitoring consumes a large portion of cognitive resources [42]. In contrast, monotonous or repetitive tasks tend to heighten self-awareness and internal focus, thereby amplifying the salience of thermal discomfort [41]. Workers engaged in repetitive assembly line tasks, for example, often report that minor deviations in temperature or airflow become more noticeable and bothersome than they are for workers in more varied, stimulating roles [12]. Thus, motivation and attentional allocation appear to regulate the extent to which individuals consciously register and report thermal sensations.

Empirical evidence supports this moderating effect of task factors on thermal comfort. Laboratory studies have shown that subjects performing vigorous or physically demanding activities report greater heat dissatisfaction compared with those engaged in light sedentary work under the same environmental conditions [163]. Conversely, office workers engaged in prolonged low-intensity tasks are more prone to cold complaints, particularly in air-conditioned environments [146, 163]. Controlled experiments on cognitive load further indicate that thermal dissatisfaction is less pronounced during periods of high attentional demand [12]; for example, subjects solving complex reasoning or memory tasks often tolerate warmer or cooler conditions with fewer complaints. By contrast, during monotonous monitoring tasks, subjects not only report greater discomfort but also attribute it more strongly to the thermal environment [12]. This finding is consistent with the argument that when tasks are insufficiently stimulating, individuals become more sensitive to environmental stressors, leading to greater salience of thermal discomfort. Field studies in occupational contexts also corroborate these laboratory findings: in manufacturing environments, heat complaints are more frequent during repetitive, assembly-based tasks, whereas in knowledge work settings, cold complaints dominate during extended computer use in air-conditioned offices [46]. These converging lines of evidence suggest that thermal comfort reflects an interaction between environmental conditions, physical workload, and attentional engagement, rather than a purely environmental response.

In summary, task factors act as moderators of thermal comfort. Physical demands alter the body’s heat balance and influence the rate of thermoregulatory strain [90], while cognitive and attentional demands shape the salience of thermal sensations through mechanisms of distraction or heightened awareness [31]. Monotony, motivation, and task engagement further determine whether discomfort is downplayed or amplified [152]. As a result, individuals engaged in different types of tasks may experience and report thermal comfort in systematically different ways, even under identical environmental conditions [41]. Recognizing this interaction is essential for understanding subjective comfort in occupational and indoor settings, and it highlights the need to jointly consider both thermal and task contexts when evaluating environmental satisfaction [31]. Such an integrated perspective is particularly relevant for workplace and building design, where thermal conditions and task characteristics are tightly coupled in shaping both individual well-being and organizational productivity.

2.4.2 Combined effects on task cognitive–performance domain (TCPD)

The task cognitive–performance domain (TCPD) refers to the combined outcomes of subjective workload sensation and objective task performance. Research has increasingly demonstrated that thermal environments do not only influence how comfortable individuals feel, but also how demanding they perceive a task to be and how well they are able to perform it [31]. Importantly, subjective workload and objective performance do not always change in parallel: thermal discomfort often manifests first in increased self-reported workload before measurable performance deterioration becomes apparent [31]. This highlights TCPD as a multidimensional construct that captures both the immediate cognitive–affective strain and the behavioral consequences of combined thermal and task demands.

On the subjective side, the NASA Task Load Index (NASA LX) has been widely used to assess workload sensation under varying thermal conditions. Studies consistently show that workers in warm environments report higher ratings of mental demand, effort, and frustration compared with those in neutral conditions [31]. Cool environments, meanwhile, often lead to increased ratings of effort and dissatisfaction linked to difficulties in maintaining alertness and comfort during sedentary tasks [31]. Field studies demonstrate that thermal interventions such as localized cooling devices or personal fans reduce NASA TLX scores in the dimensions of effort and frustration, underscoring the sensitivity of subjective workload to thermal variation [31]. These findings suggest that workload sensation is a particularly sensitive indicator of thermal strain, capturing changes that may not yet be observable in performance metrics.

On the objective side, numerous experiments have examined how thermal stress influences accuracy, attention, reaction time, and error rates. Heat exposure has been found to impair vigilance, slow reaction times, and elevate error frequency in complex or time-pressured tasks [111]. Cold exposure, by contrast, is associated with longer response latencies, reduced working memory capacity, and declines in sustained attention [121]. The magnitude of these effects depends strongly on task characteristics: when tasks are simple or intrinsically engaging, individuals are often able to maintain stable performance despite reporting greater workload, reflecting compensatory effort [163]. However, when tasks are monotonous, cognitively complex, or under strict time constraints, thermal stress accelerates cognitive fatigue and produces measurable declines in task performance [111].

Evidence from applied contexts further illustrates these interactions. Driving simulator studies have shown that elevated cabin temperatures impair lane-keeping accuracy and increase attentional lapses, particularly during monotonous highway driving [116]. In healthcare, surgeons exposed to elevated operating room temperatures demonstrate reduced precision and higher error rates in fine motor tasks, effects that are exacerbated by the high cognitive load of surgical decision-making [170]. Industrial studies similarly report that both heat and cold stress degrade accuracy in precision manufacturing tasks, highlighting the vulnerability of safety-critical work to combined thermal and task demands [163]. Collectively, these findings reveal that TCPD outcomes are shaped by a dual pathway: thermal stress elevates subjective workload sensation, often preceding performance decline, and under certain task conditions it also directly undermines objective task performance.

In summary, TCPD provides a valuable framework for understanding how thermal and task factors interact to influence human functioning. Thermal discomfort heightens workload perception, as captured by NASA TLX, and under specific task conditions it impairs objective performance in ways that compromise accuracy, vigilance, and safety [130]. Recognizing this duality underscores the importance of integrating both subjective and objective measures in studies of human performance in thermal environments, and it highlights the need for task-sensitive approaches when designing occupational and indoor settings.

2.4.3 Combined effects on stress

While thermal comfort and TCPD capture the immediate subjective and behavioural outcomes of combined thermal and task exposures, they do not fully account for the broader psychophysiological strain that arises under such conditions. In practice, individuals often experience these influences not only as discomfort or increased workload, but also as stress, a multidimensional state encompassing physiological arousal, perceived overload, and emotional strain [165]. Stress therefore represents both an independent outcome and a mediating process: it emerges from the same thermal and task demands that shape comfort and performance, yet it also contributes to longer-term consequences such as fatigue, burnout, and reduced well-being [152, 165]. For this reason, research has increasingly examined stress as a distinct domain in its own right, complementing thermal comfort and TCPD to provide a more comprehensive understanding of human responses to combined environmental and task challenges.

Stress represents a multidimensional response encompassing physiological arousal, psychological strain, and behavioural consequences, and it arises when perceived demands exceed available resources [152]. Both thermal environments and task demands are established sources of stress, and when they occur simultaneously, their effects often accumulate rather than substitute. Thermal discomfort can act as an environmental stressor that amplifies the strain imposed by demanding or monotonous tasks, while task pressure can heighten sensitivity to thermal stressors. Thus, stress emerges as a key outcome domain reflecting the combined influence of thermal and task factors.

Physiologically, thermal stress interacts with task stress to increase autonomic and endocrine responses. Heat exposure elevates core temperature, heart rate, and sweat production, while cold exposure induces vasoconstriction, shivering, and altered cardiovascular activity [89]. When these thermal loads are paired with demanding tasks, the resulting strain is intensified. For example, high-temperature environments combined with cognitively complex or time-pressured tasks have been shown to increase heart rate variability indices of stress and elevate cortisol levels beyond the effects of either factor alone [37]. Similarly, cold environments exacerbate stress responses during prolonged vigilance tasks, leading to heightened sympathetic activation and impaired thermoregulation [121]. These physiological interactions underline how thermal and task demands jointly tax the body's adaptive systems.

Psychologically, the combination of thermal discomfort and demanding tasks amplifies perceived stress and emotional strain. Individuals performing tasks under hot or cold conditions consistently report higher scores on self-report stress scales compared with neutral environments [171]. Task complexity and time pressure further magnify these perceptions, creating a sense of overload that exceeds the sum of environmental, and task demands considered separately [42]. In monotonous tasks, thermal discomfort often heightens frustration and irritability, while in high-stakes contexts such as surgery or emergency response, thermal stress compounds performance pressure and increases stress [130]. Over extended periods, this cumulative stress contributes to fatigue, reduced motivation, and burnout, linking short-term stress reactions to longer-term health and well-being outcomes.

Evidence from applied settings illustrates the significance of these combined effects. In occupational environments such as manufacturing plants, workers exposed to high heat while meeting production quotas report elevated stress, greater emotional exhaustion, and higher turnover intentions [133]. Healthcare professionals working in thermally suboptimal hospital wards experience heightened stress and reduced job satisfaction, particularly when combined with heavy workloads and long shifts [46]. Driving studies similarly reveal that both heat and cold stress increase driver strain, with subjects reporting higher frustration and reduced tolerance for errors under challenging traffic conditions [116]. These findings highlight that stress is not only an outcome in itself but also a mechanism: elevated stress mediates the pathway through which thermal and task factors jointly impair performance, safety, and well-being.

In summary, stress reflects the convergence of thermal discomfort and task demands at both physiological and psychological levels. Thermal and task stressors interact to elevate autonomic arousal, hormonal responses, perceived overload, and emotional strain [31]. These combined effects contribute to immediate decrements in attention and decision-making, as well as long-term risks such as fatigue and burnout. Recognizing stress as both an outcome and a mediating process is essential for understanding the broader consequences of combined thermal and task exposures, and it underscores the importance of integrated approaches to managing environmental and task design in occupational and indoor contexts [46].

2.4.4 Integrative mechanisms and models

The findings reviewed in Sections 2.4.1–2.4.3 demonstrate that the combined influence of thermal environments and task factors manifests in three major domains: thermal comfort, TCPD, and psychological stress. However, these domains should not be considered isolated outcomes. Instead, they reflect different levels of a dynamic system in which environmental and task stressors interact through direct and indirect mechanisms. In this regard, thermal environments exert their most immediate impact on subjective sensations of comfort and satisfaction, while task factors such as complexity, duration, and workload requirements exert their most immediate impact on TCPD. Beyond these direct effects, cross-domain influences are also evident: thermal discomfort can heighten subjective workload and impair task performance, whereas task demands can shape the perception of thermal comfort by either masking or amplifying discomfort depending on attentional allocation and motivational state [31]. Ultimately, both comfort and workload converge in the experience of stress, which captures the cumulative psychological cost of environmental and task demands and may itself feed back into both comfort and performance over time [165].

The most robust direct pathway is the influence of the thermal environment on thermal comfort. This link has been consistently supported across laboratory and field studies, which show that deviations from thermoneutral conditions directly reduce comfort ratings[172]. Similarly, task demands constitute the primary determinant of TCPD, as reflected in subjective workload assessments (e.g., NASA-TLX) and objective performance outcomes [31]. Indirect effects, however, are equally important. A growing body of evidence indicates that thermal discomfort can increase perceived workload, even when objective task performance is preserved through compensatory effort [130]. At the same time, task demands can exert feedback effects on thermal comfort. When tasks are demanding and attentionally engaging, individuals may report reduced perception of discomfort due to attentional distraction. Conversely, when workload is excessive or motivation is low, discomfort is intensified, leading to lower thermal satisfaction [79]. Stress represents the point at which the effects of thermal comfort and TCPD intersect. Both environmental discomfort and high workload contribute to stress perception. Importantly, stress functions both as an outcome and as a mediator: discomfort may elevate stress, which in turn undermines cognitive resources and increases workload perception, while sustained workload can elevate stress, which subsequently heightens sensitivity to thermal discomfort [31]. This dual role positions stress as both a consequence and a feedback mechanism within the integrative system. These interconnections are summarized in Figure 2.3, which illustrates the direct, indirect, and feedback relationships among thermal environment, task factors, thermal comfort, TCPD, and stress.

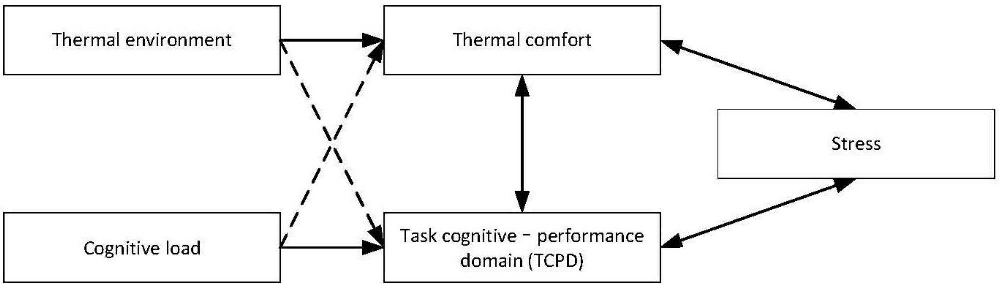


Figure 2.3 Conceptual model illustrating the relationships between thermal environment, task factors, thermal comfort, task-related cognitive performance and demand (TCPD), and stress.

Note: Solid arrows indicate direct effects, while dashed arrows represent indirect or feedback pathways.

Several theoretical perspectives provide useful, though partial, explanations of these interrelationships. The stress–strain framework conceptualizes thermal and task demands as external stressors that elicit physiological and psychological strain, offering a general lens for understanding additive effects [165]. The workload × environment model further specifies the compounding impact of task and environmental demands on performance [165]. The compensatory effort model highlights why performance can remain stable in the short term despite increased thermal discomfort and workload, by attributing stability to elevated compensatory resource allocation [165]. Beyond these, broader models from occupational health and cognitive psychology enrich the analysis. The Job Demands–Resources (JD-R) model positions thermal and task demands as job demands whose effects are moderated by available resources such as motivation or coping capacity [165]. Multiple Resource Theory (MRT) explains how thermal discomfort consumes attentional and cognitive resources that would otherwise be available for task execution, intensifying subjective workload [31]. From a long-term health perspective, the allostatic load model illustrates how repeated or chronic exposure to thermal and task stressors accumulates physiological and psychological costs, thereby explaining distinctions between short-term compensatory stability and long-term fatigue [173]. Similarly, the Person–Environment Fit model emphasizes mismatches between individual capacities and environmental demands, highlighting how personality, motivation, and thermal tolerance shape the comfort–workload–stress nexus [165]. The Effort–Recovery model underscores the necessity of recovery periods, warning that sustained exposure to combined thermal and task demands without sufficient rest can lead to cumulative stress and deteriorating performance [152]. Finally, environmental stress models from environmental psychology conceptualize thermal load as one of several prototypical environmental stressors, acting through both direct sensory discomfort and indirect cognitive pathways to influence well-being and performance [12].

2.4.5 Research gaps and future directions

While Section 2.4.4 outlined how existing models conceptualize the combined effects of thermal environments and task demands, it also highlighted the fragmented nature of current evidence. Despite progress, several important gaps remain at the level of this specific research domain. Addressing these gaps is vital for building more comprehensive and predictive models of how environmental and task stressors interact to influence comfort, performance, and well-being.

Limited integration of theoretical frameworks and empirical validation

Current models such as stress–strain, workload × environment, and compensatory effort provide valuable insights into specific mechanisms, but they have rarely been tested as part of a unified causal structure [165]. Studies often examine comfort, TCPD, or stress in isolation, leaving their interdependencies underexplored. Without systematic empirical validation, it remains unclear whether these processes are additive, multiplicative, or mediated through feedback loops.

Predominant reliance on short-term laboratory designs

Much of the literature is based on highly controlled experiments of limited duration [12]. While useful for isolating variables, these designs fail to capture the cumulative effects of repeated or chronic exposure, such as fatigue, motivation decline, or long-term health impacts. The distinction between short-term compensatory stability and long-term strain, emphasized in theories like the allostatic load and effort–recovery models, remains insufficiently addressed in empirical work.

Insufficient modeling of feedback and dynamic processes

Although theoretical accounts acknowledge bidirectional influences, such as stress amplifying discomfort, or workload shaping comfort perception, few studies have explicitly modeled these dynamics [33]. Most analyses remain static, limiting their ability to explain how responses evolve over time or across repeated exposures.

Underrepresentation of individual differences

Variability in responses to combined thermal and task stressors is substantial, shaped by factors such as thermal tolerance, personality traits, motivation, and coping strategies [12]. Yet these moderators are rarely integrated into models, leading to oversimplified predictions and limited applicability across populations.

Narrow measurement approaches

Research remains dominated by subjective rating scales (e.g., comfort votes, NASA-TLX) [12]. While these capture valuable perceptions, they are prone to bias and may not reflect underlying physiological strain. Objective performance measures are often simplistic, and physiological indicators are inconsistently applied. As a result, the full scope of comfort–workload–stress interactions is not adequately captured.

Limited ecological scope

Most studies focus on office-like tasks and student samples, often within Western cultural contexts [28]. This narrow scope raises questions about generalizability to diverse occupational settings (e.g., manufacturing, healthcare, outdoor work) and cultural environments where thermal expectations and stress appraisals differ.

Future directions

To address these gaps, future research should pursue several complementary directions. First, integrative models should be explicitly specified and empirically validated, capturing both direct effects (thermal → comfort; task → TCPD) and cross-domain pathways (thermal → TCPD; task → comfort). Second, dynamic perspectives are needed, distinguishing short-term compensatory adjustments from long-term strain accumulation, ideally through longitudinal and field studies. Third, individual differences such as motivation, personality, and thermal tolerance should be systematically integrated as moderators, enabling more personalized predictions of vulnerability to combined stressors. Fourth, multimodal measurement protocols that combine subjective, behavioral, and physiological indicators should be adopted to enhance validity and reliability. Finally, research must broaden its ecological scope, extending beyond laboratory contexts to diverse occupational and cultural settings, thereby ensuring the generalizability and practical relevance of findings.

In summary, while current research provides important building blocks, advancing this domain requires moving from fragmented and static studies toward integrative, dynamic, and ecologically valid approaches. Such progress will allow for more accurate prediction of comfort, performance, and stress outcomes, and will lay the groundwork for designing environments and tasks that safeguard both productivity and well-being under thermal stress.

2.5 Personal Factors and Individual Differences

Thermal comfort and productivity are not uniform experiences; they are mediated by personal characteristics that shape how individuals perceive, respond to, and cope with environmental conditions. While predictive models often rely on "average" assumptions of metabolic rate or clothing insulation, real populations exhibit wide variation in physiology, behaviour, and socio-cultural background. These differences are critical for understanding the thermal comfort–productivity–discomfort (TCPD) relationship, as they determine not only subjective comfort but also cognitive performance, resilience to stress, and long-term health outcomes. This section reviews major categories of individual differences, age and gender, clothing insulation, body composition and metabolism, and other contextual factors, that influence TCPD outcomes.

2.5.1 Age and gender

Age and gender are central determinants of how individuals experience thermal comfort and how discomfort translates into productivity losses and stress. Physiologically, ageing is associated with reduced sweat gland activity, diminished vasomotor responsiveness, and slower thermoregulatory adjustments [174]. These changes make older adults more vulnerable to both heat and cold, often perceiving discomfort earlier and recovering more slowly. In occupational settings, such discomfort is not merely subjective; it frequently results in reduced concentration, slower task completion, and higher fatigue, thereby linking age directly to the thermal comfort–productivity–discomfort (TCPD) pathway [175].

Gender differences also shape thermal responses. Women commonly report feeling colder than men in identical environments, a difference attributed to lower basal metabolic rates, higher body fat distribution, and hormonal influences [176]. These physiological distinctions mean that, in cooler offices, women may experience discomfort and reduced performance, whereas men may be relatively unaffected. Conversely, in warmer environments, men's higher metabolic heat production can lead to earlier onset of fatigue and stress, while women may remain closer to their comfort zone [47]. Such disparities highlight that thermal environments optimised for one demographic may inadvertently disadvantage another, with direct implications for productivity and well-being.

The interaction between age, gender, and task demands further complicates TCPD outcomes. In physically demanding tasks, older workers often experience greater cardiovascular and thermal strain, leading to faster onset of fatigue and stress compared with younger counterparts [99]. In cognitively demanding tasks, thermal discomfort has been shown to impair memory, attention, and decision-making, with gender differences in tolerance thresholds influencing error rates and persistence[31, 99]. For example, women may experience sharper declines in performance in cold offices, while men may be more susceptible to productivity losses in hot conditions, particularly during tasks requiring sustained vigilance.

Different thermal environments also magnify these variations. In stable, moderate climates, age and gender differences may be less pronounced, but in extreme or fluctuating conditions, disparities become significant. Older adults are less able to adapt to rapid temperature changes, while younger individuals can recover more quickly from thermal stress [50]. Gendered responses are similarly context-dependent: colder environments exacerbate discomfort and stress among women, whereas hotter environments elevate strain among men.

2.5.2 Clothing insulation

Clothing insulation is among the most immediate and modifiable personal factors shaping thermal comfort, productivity, and stress. By altering heat transfer between the body and the environment, clothing directly influences whether individuals perceive an environment as tolerable or stressful. Inadequate insulation in cold environments leads to shivering, impaired dexterity, and cognitive distraction, while excessive insulation in warm conditions elevates core temperature, cardiovascular strain, and fatigue [47]. Both scenarios undermine productivity and increase discomfort, making clothing a critical mediator in the thermal comfort–productivity–discomfort (TCPD) pathway.

The role of clothing is highly dependent on the thermal environment. In stable, moderate climates, individuals can often self-regulate by adjusting layers, maintaining both comfort and adequate performance. However, in extreme or fluctuating conditions, clothing choices become more consequential. High insulation in hot environments can trap heat, producing rapid onset of stress, while insufficient insulation in cold, draughty spaces reduces concentration and endurance [163]. The inability to adapt clothing to changing environments, such as during outdoor labour or in poorly controlled offices, exacerbates discomfort and accelerates productivity losses.

Clothing also interacts with task factors. For physically demanding tasks, restrictive or insulating garments (e.g., protective equipment, uniforms) significantly increase metabolic strain, leading to faster fatigue and heightened stress [47]. In cognitively demanding but sedentary tasks, inadequate insulation in cool offices has been linked to reduced typing speed, increased error rates, and lower task persistence [31, 177]. Long-duration tasks intensify these effects, as prolonged exposure amplifies discomfort and impairs sustained attention. Conversely, short-duration tasks may allow individuals to tolerate suboptimal conditions without major productivity decline.

Cultural and organisational constraints further complicate the picture. In many workplaces, clothing choices are limited by dress codes or safety standards, restricting adaptive behaviour. For example, formal office attire may prevent employees, particularly women, from adjusting clothing to suit cooler environments, while industrial workers may be required to wear protective gear that increases heat burden [47]. Such constraints elevate stress by removing personal control, a key factor in coping with thermal strain.

2.5.3 Body mass index (BMI) and basal metabolic rate (BMR)

Body composition and metabolic rate are fundamental determinants of how individuals experience thermal environments and how discomfort translates into productivity losses and stress. Higher body mass index (BMI), often associated with greater adipose tissue, reduces heat dissipation and increases insulation, predisposing individuals to greater heat strain in warm environments [178]. Conversely, lower BMI may heighten susceptibility to cold stress, leading to discomfort and impaired task performance in cooler settings [50]. Basal metabolic rate (BMR), reflecting the body's endogenous heat production, also shapes thermal responses. Individuals with higher BMR generate more metabolic heat, which can accelerate fatigue and stress in hot environments but may provide an advantage in cold conditions [47].

The influence of body composition and metabolism varies across thermal environments. In hot climates, high BMI individuals often report earlier onset of discomfort, higher cardiovascular strain, and reduced tolerance for sustained activity [99]. At the same time, those with high BMR may experience faster rises in core temperature, increasing stress and limiting endurance. In cold environments, individuals with greater adiposity may enjoy greater insulation and comfort, while leaner individuals with lower BMR experience faster heat loss and reduced manual dexterity [50]. These differences underscore that thermal comfort cannot be assumed uniform across populations.

Task factors further moderate these relationships. During physically demanding tasks, higher metabolic rates amplify heat production, exacerbating discomfort and accelerating productivity decline under warm conditions [47]. Sedentary cognitive tasks reveal a different pattern: individuals with lower BMR may struggle to maintain comfort in cool offices, experiencing distraction, slower performance, and heightened stress [47]. Long-duration tasks magnify these effects, as sustained exposure compounds physiological strain, whereas short-duration tasks may allow individuals to tolerate suboptimal conditions without major productivity impairment.

Despite the clear physiological rationale, research in this domain faces limitations. BMI remains a coarse measure that does not distinguish between fat and muscle mass, obscuring important differences in thermoregulatory capacity. BMR is often estimated using generalised formulae rather than measured directly, reducing accuracy [172]. Moreover, most studies are conducted in controlled laboratory settings, with limited insight into how body composition and metabolism interact with clothing, task demands, or adaptive behaviours in real-world workplaces.

2.5.4 Other personal factors

Beyond age and gender, clothing insulation, and body composition, a variety of other personal factors influence thermal comfort, productivity, and stress outcomes. Acclimatisation history is one such determinant: individuals accustomed to hot or cold climates often display altered thermal tolerance and expectations, which affect how discomfort translates into stress and performance [68]. Health status also plays a role, as chronic conditions such as cardiovascular disease, diabetes, or respiratory disorders can impair thermoregulatory capacity, making some individuals more vulnerable to thermal strain [90].

Psychological and behavioural traits further complicate TCPD dynamics. People differ in thermal sensitivity, coping strategies, and motivation levels, which shape how discomfort is perceived and how it impacts task persistence [12]. Cultural and socio-economic backgrounds also matter, influencing clothing practices, thermal expectations, and access to adaptive technologies [48]. Finally, circadian rhythm and chronotype may determine when individuals are most susceptible to thermal stress during the working day, with potential implications for productivity across shifts [179].

While these factors are clearly relevant, they are not the primary focus of the present study. Instead, this review concentrates on age and gender, clothing insulation, and body composition and metabolism, as these represent the most consistently studied and quantifiable personal determinants of thermal comfort, productivity, and stress. Other factors are acknowledged here for completeness, but they are considered peripheral to the core scope of this work.

2.5.5 Summary

Personal factors are central to the way thermal conditions are experienced and evaluated. In terms of thermal comfort, individual differences explain why the same physical environment can be perceived as tolerable for some but uncomfortable for others. Variations in age, gender, clothing habits, and body composition mean that thermal sensation is never uniform, and subjective comfort thresholds differ considerably across populations.

Within the TCPD (thermal comfort–productivity–discomfort) framework, these personal differences translate into measurable outcomes. Age and gender shape how discomfort impairs physical endurance and cognitive performance under different environmental and task conditions. Clothing insulation offers a means of adaptation, but when restricted, it can directly undermine productivity. Body composition and metabolism influence how heat is generated and dissipated, altering the balance between comfort, efficiency, and fatigue. Together, these factors highlight the strong link between personal variation and workplace or task productivity.

From the perspective of stress, personal determinants moderate both physiological strain and emotional responses. Older workers often experience faster stress accumulation during exposure to thermal extremes; women may find lower temperatures particularly stressful when adaptive clothing is constrained; and individuals with high metabolic rates may enter heat stress more quickly. Such stress reactions compound the effects of discomfort on both well-being and performance.

Other factors, such as acclimatisation, health status, chronotype, and psychological or cultural differences, may also shape thermal comfort, TCPD dynamics, and stress responses. However, this review emphasises three main dimensions (age and gender, clothing insulation, body composition and metabolism) because they are most consistently reported in the literature and most directly quantifiable.

While useful, existing models tend to oversimplify personal variation, treating individuals as interchangeable averages. This misses crucial diversity in comfort thresholds, productivity trajectories, and stress responses. To move forward, integrative methods are needed that connect subjective evaluations with objective evidence. Accordingly, the next section (2.6) examines physiological factors and imaging indicators, which offer measurable insights into how the human body responds to thermal environments. These indicators serve as essential tools for refining TCPD understanding and capturing stress responses beyond self-report alone.

2.6 Physiological parameters and imaging indicators

The study of thermal environments has long relied on subjective reports of thermal comfort, yet these alone cannot fully capture the body's multidimensional responses. Thermal sensation may be moderated by age, gender, clothing, or body composition, but physiological signals such as skin temperature, heart rate, and sweat rate provide more direct evidence of how the body interacts with its environment [47]. By integrating physiological indicators, researchers can move beyond perception to a more comprehensive understanding of comfort thresholds.

Within the TCPD framework, physiological monitoring makes it possible to link environmental stressors with quantifiable productivity outcomes. For example, core temperature elevation or cardiovascular strain can correspond to measurable changes in task accuracy, reaction time, or endurance [180]. Imaging technologies, such as infrared thermography, extend this capability by mapping thermal distribution across the body, thereby connecting discomfort reports with functional impairments in real-world settings.

From the perspective of stress, physiological markers provide critical evidence that cannot be derived from subjective assessments alone. Elevations in heart rate variability, cortisol levels, or thermal asymmetries reveal the onset of strain even before it is consciously perceived [44]. Imaging indicators allow non-invasive tracking of these stress responses, bridging the gap between self-reported discomfort and objective physiological condition.

Taken together, physiological factors and imaging indicators supply a necessary complement to the personal differences discussed in Section 2.5. They offer a pathway to more precise, inclusive, and predictive models of human response to thermal environments, enabling researchers to ground subjective comfort and productivity reports in measurable biological substrates.

2.6.1 Physiological parameters

Physiological parameters serve as the most direct indicators of how the human body interacts with thermal environments. Unlike subjective thermal sensation votes, which are prone to bias and adaptation effects, these measures provide objective evidence of thermal load, circulatory adjustments, and autonomic activity. From the perspective of thermal comfort, parameters such as skin and core temperature map closely onto sensations of warmth, neutrality, and cold. Within the TCPD framework, they capture the physiological mechanisms through which thermal strain contributes to productivity losses and fatigue. Regarding stress, cardiovascular, metabolic, and sweat responses reflect activation of the autonomic nervous system and endocrine pathways triggered by heat or cold exposure.

By examining these parameters systematically, researchers can clarify how bodily responses complement subjective perception, helping to build more robust and predictive models of thermal comfort, productivity, and stress. The sections below highlight the most relevant physiological measures, including core and skin temperature, sweat rate, cardiovascular indicators, and biochemical markers, each discussed in relation to comfort, TCPD, and stress outcomes.

Core body temperature (T_c)

T_c represents the body's internal thermal state and is maintained within a narrow range of approximately 36.5–37.5 °C under normal conditions. This relative constancy is achieved through the integration of metabolic heat production and heat exchange mechanisms, including vasodilation, sweating, vasoconstriction, and shivering [95]. In practice, T_c can be measured using several approaches: rectal and esophageal thermistors remain the research gold standard due to their accuracy, while ingestible telemetric capsules offer less invasive monitoring for field studies [181]. More recently, wearable devices using tympanic, oral, or gastrointestinal sensors have been applied in occupational and sports settings, although their precision can vary depending on the environment [182]. Because T_c is tightly controlled, even subtle shifts in its level or rate of change provide important insights into how the body experiences comfort, productivity, and stress.

In terms of thermal comfort, changes in T_c are closely associated with subjective sensations of warmth or cold. Controlled climatic chamber studies demonstrate that an increase of only 0.3–0.5 °C above baseline T_c is sufficient for subjects to report moving from a thermally neutral state to warm discomfort [183]. The magnitude of discomfort is exacerbated when T_c rises quickly, indicating that the velocity of change can be more salient than the absolute deviation. Conversely, a gradual reduction of T_c by as little as 0.2–0.4 °C consistently results in reports of chilliness or cold discomfort, despite the absolute values still lying within a clinically normal range [82]. These observations highlight the sensitivity of thermal comfort to modest deviations in internal temperature, underscoring the body's reliance on T_c as a physiological anchor for subjective perception.

Within the TCPD framework, T_c dynamics provides a mechanistic link between environmental exposure and productivity outcomes. Elevated T_c is a strong predictor of physical performance loss: industrial and military studies have shown that once T_c exceeds around 38–38.5 °C, endurance declines sharply, work output decreases, and individuals are forced to slow pace or rest [184]. Cognitive productivity also suffers, with experimental findings indicating that even smaller rises in T_c, around 0.5 °C, correlate with slower decision-making, impaired memory recall, and increased error rates in tasks demanding concentration [31]. Similarly, cold-induced reductions in T_c compromise neuromuscular performance. Drops of approximately 0.5–1.0 °C reduce finger dexterity, reaction speed, and fine motor coordination, leading to measurable declines in tasks such as electronic assembly, data entry, or surgical precision [185]. Together, these results demonstrate that T_c variation is a critical determinant of both physical and cognitive productivity performance.

From the perspective of stress, T_c fluctuations act as powerful physiological stressors. An upward shift in T_c under heat exposure activates the sympathetic nervous system, producing elevated heart rate, increased ventilation, and redistribution of blood flow, often accompanied by rises in cortisol and other stress hormones [99]. These responses, while adaptive, accumulate as perceived strain, with workers reporting irritability, tension, and fatigue under sustained heat loads [37]. In the opposite direction, declining T_c under cold exposure triggers shivering thermogenesis and catecholamine release, protective mechanisms that nonetheless generate metabolic cost and elevate systemic stress [89]. Further studies suggest that repeated episodes of T_c deviation, whether through occupational heat stress or prolonged cold environments, contribute to chronic stress load and reduced resilience over time [186]. Thus, T_c emerges as a central biomarker not only for momentary discomfort and productivity loss but also for the broader physiological stress burden induced by thermal environments.

Skin temperature (T_{sk})

Skin temperature (T_{sk}) reflects the thermal state at the body surface and is the primary channel through which individuals perceive warmth or coolness. Unlike T_c, which is buffered by homeostatic regulation, T_{sk} responds more rapidly to environmental variations, making it a sensitive but less stable indicator of thermal state [43]. Measurement is commonly performed using thermocouples or thermistors attached to representative sites such as the chest, back, hands, and feet, with mean skin temperature often calculated from weighted averages [44]. Infrared thermography has also become increasingly popular for non-contact assessment of T_{sk} in both laboratory and field studies [38]. These methods reveal that fluctuations in T_{sk} are strongly tied to subjective sensation, productivity outcomes, and stress responses.

With respect to thermal comfort, variation in T_{sk} provides the most immediate physiological correlate of subjective warmth and coolness. Numerous studies demonstrate that increases in mean T_{sk} toward 34–35 °C correspond with sensations of warm discomfort, while decreases below 30–31 °C produce reports of cold discomfort [183]. Importantly, local skin temperatures also matter. For example, cooling of the extremities such as hands and feet leads to disproportionate reports of discomfort compared to similar cooling of the trunk [187]. Furthermore, the rate of change in T_{sk}

strongly influences comfort perception, with rapid cooling generating acute discomfort even when T_c remains stable [82]. T_{sk} thus serves as the most direct physiological mediator between the external environment and subjective comfort assessment.

Within the TCPD framework, changes in T_{sk} have been shown to predict both physical and cognitive productivity outcomes. Elevated T_{sk} under heat stress is associated with reduced vigilance and slower reaction time, undermining efficiency in cognitive tasks [111]. Conversely, localized reductions in T_{sk}, particularly in the hands and fingers, substantially degrade manual dexterity and tactile sensitivity, leading to performance loss in precision-oriented work such as typing, assembly, or laboratory tasks [96]. Long-duration exposures intensify these effects, as sustained deviations in T_{sk} accumulate into discomfort-driven fatigue and concentration lapses [31]. These findings demonstrate that not only average T_{sk} but also its distribution across the body plays a crucial role in shaping task performance.

In relation to stress, alterations in T_{sk} are recognized as reliable markers of autonomic nervous system activity. A rapid drop in peripheral T_{sk}, commonly observed in cold exposure, reflects vasoconstriction orchestrated by sympathetic activation and is perceived as stressful well before T_c falls significantly [134]. Similarly, facial thermal asymmetries measured through changes in Tsk distribution have been linked to psychological stress, suggesting that T_{sk} is sensitive to both environmental and emotional challenges [38]. In hot environments, sustained elevations in T_{sk} can also contribute to sympathetic strain, as the body mobilizes sweating and circulatory adjustments to facilitate heat loss [34]. Over repeated episodes, these T_{sk} fluctuations accumulate to higher stress loads, bridging physiological strain with subjective discomfort and reduced resilience.

Sweating parameters

Sweating is one of the most important mechanisms by which the body regulates heat balance. It facilitates evaporative cooling, delaying rises in T_c and stabilizing thermal equilibrium under hot conditions. Unlike T_c and T_{sk}, which can be expressed directly in degrees Celsius (˚C), sweat production itself is less straightforward to quantify. Measures such as sweat rate (ml/min·cm²) and skin humidity provide insights into moisture dynamics but are technically demanding, requiring specialized capsules, absorbent patches, or localized humidity sensors [188]. In practice, a more accessible and robust proxy is electrodermal activity (EDA), also referred to as galvanic skin response (GSR). EDA records changes in skin conductance (μS) caused by sweat gland activity, typically measured via electrodes placed on the fingers or palms, or by wearable devices with integrated sensors [189]. Because sweating is responsive to both thermal and psychological challenges, EDA has become a central physiological signal linking thermal comfort, productivity, and stress.

In relation to thermal comfort, sweating patterns recorded through EDA provide a close physiological correlate of subjective sensations. Because EDA reflects sudomotor activity in eccrine sweat glands, particularly those concentrated on the palms, soles, and face, its rise directly accompanies increased perceptions of warmth, dampness, or clamminess. In hot-dry environments, sweat evaporates efficiently, so subjects may report tolerable comfort despite higher sweat production. By contrast, in hot-humid environments, limited evaporation causes skin wetness to persist, and EDA values remain elevated for longer durations, aligning with consistently lower comfort ratings [75]. Local differences are also important: sweat accumulation on the palms and face disproportionately drives discomfort and distraction compared to sweating on the trunk [27]. By attaching electrodes at localized skin sites, EDA can quantify these spatial variations, offering an objective marker that captures the subjective perception of stickiness and thermal dissatisfaction before T_c shifts significantly.

Within the TCPD framework, EDA illustrates the dual role of sweating in physical and cognitive performance. On the one hand, moderate sweating helps defend core temperature stability, delaying rapid rises in T_c and allowing sustained physical performance during strenuous work or exercise [89]. On the other hand, when sweating becomes excessive, reflected as prolonged elevations in EDA, negative outcomes appear. In physical tasks, excess surface moisture compromises grip stability, fine motor control, and tool manipulation, with higher error rates documented in assembly-line and precision work as EDA increases [190]. At the same time, fluid and electrolyte losses associated with heavy sweating reduce muscular endurance, creating longer-term productivity declines. In cognitive tasks, sustained high EDA values during heat exposure are predictive of lapses in sustained attention, slower reaction speed, and decreased accuracy in memory recall and vigilance testing [112]. These findings demonstrate that while sweating initially functions as a protective mechanism, continuous elevations in EDA reflect the tipping point where thermoregulation ceases to be beneficial and begins impairing both physical precision and cognitive efficiency.

From the perspective of stress, EDA is recognized as one of the most robust and sensitive indicators of sympathetic nervous system activity. Under thermal stress, elevated EDA corresponds directly to sudomotor activation driven by sympathetic cholinergic fibers, reflecting intensified thermoregulatory effort and strain [191]. Importantly, EDA is equally responsive to psychological stressors independent of thermal load: mental workload, time pressure, or emotional arousal can all elicit sharp phasic increases in palmar or plantar EDA, even in otherwise comfortable environments [44]. This dual sensitivity underscores EDA's role as a marker that integrates both thermal strain and psychological load in real time. Moreover, longitudinal studies show that repeated or persistently elevated EDA, whether due to high humidity, prolonged heat stress, or continuous cognitive demand, contributes to cumulative strain and reduced resilience [192]. In this way, EDA provides a uniquely direct index of how sweating links comfort, productivity, and stress in everyday contexts.

Cardiovascular parameters

The cardiovascular system undergoes immediate adjustments when exposed to thermal or psychological challenges, redistributing blood flow to balance heat exchange and maintain core stability. A range of indicators can be obtained: heart rate (HR) and blood pressure (BP) reflect overall circulatory load; inter-beat interval (IBI) describes the time spacing between successive heartbeats; blood volume pulse (BVP) indicates peripheral vascular dynamics; and heart rate variability (HRV), including the low-frequency/high-frequency (LF/HF) ratio, reveals autonomic balance between sympathetic and parasympathetic inputs [44]. Among these, the most widely adopted and practically useful in both laboratory and field research are HR, BP, and HRV, which together provide a comprehensive yet feasible picture of cardiovascular regulation [44].

For measurement, electrocardiography (ECG) remains the gold standard for extracting HR, IBI, and HRV with high temporal resolution [44]. However, photoplethysmography (PPG), an optical technique that detects blood volume fluctuations in tissues has become the mainstream approach in wearable and real-time monitoring devices, allowing convenient assessment of HR, HRV, IBI, and BVP [193]. BP is often measured via Oscillo metric cuffs, though continuous tracking can also be achieved with advanced PPG-based finger sensors [194].

With respect to thermal comfort, cardiovascular responses provide important though indirect signals. Under heat exposure, skin vasodilation increases blood pooling in peripheral vessels, reducing venue return and stroke volume. To maintain cardiac output, the body compensates with an elevated HR and shortened IBI, a shift that subjects often perceive as palpitations, higher effort at rest, or feelings of warmth-driven fatigue [99]. In cold conditions, intense vasoconstriction raises BP, while diminished peripheral perfusion reduces BVP amplitude at the skin level, contributing to sensations of cold extremities and stiffness [134]. Importantly, these changes usually precede large T_c deviations, meaning people can feel uncomfortable even though core homeostasis is still preserved. Furthermore, HRV offers a sensitive window: lower overall HRV and an elevated LF/HF ratio indicate reduced parasympathetic modulation, a physiological state consistent with subjective reports of strain, restlessness, or unease [44] Taken together, subtle alterations in HR, BP, BVP, and HRV provide early cardiovascular correlations of comfort or discomfort that complement temperature-based measures.

Within the TCPD framework, cardiovascular indicators link directly to performance capacity. For physical productivity, elevated HR during heat initially helps sustain oxygen delivery and muscle perfusion, but over time continuous tachycardia reduces efficiency, accelerates cardiovascular fatigue, and heightens dehydration risk [99]. Workers in construction or industrial tasks have shown productivity declines as HR exceeds individually sustainable thresholds [133]. In cold, vasoconstriction and elevated BP conserve core heat but compromise microcirculation in extremities; reduced BVP amplitude in fingers or toes is strongly correlated with impaired dexterity, slower fine-motor tasks, and higher error rates in manual work [134]. For cognitive productivity, HRV is a critical signal: both heat and cold exposures reduce parasympathetic activity and raise LF/HF, patterns associated with loss of attentional stability, decreased working memory accuracy, and longer reaction times in vigilance tests [112]. Thus, cardiovascular metrics not only describe systemic thermoregulatory strain but also predict tangible losses in both physical stamina and cognitive efficiency.

From the perspective of stress, cardiovascular parameters, particularly HR, BP, and HRV, are among the most validated autonomic markers across thermal and psychological domains. Under heat stress, elevated HR and suppressed HRV reveal sympathetic predominance, often co-occurring with heightened cortisol release and subjective strain [195]. Cold stress generates parallel strain through increases in BP, vasoconstriction, and reduced HRV, reinforcing the overlap of cardiovascular and perceptual stress responses [134]. Importantly, non-thermal stressors such as mental workload, performance pressure, or emotional arousal elicit highly similar cardiovascular patterns: accelerated HR, shortened IBI, elevated BP, and reduced HRV with higher LF/HF ratio [163]. These converging responses indicate that cardiovascular adjustment is a domain-general mechanism of stress reactivity. Methodologically, the widespread adoption of PPG-based monitoring devices, which allow noninvasive, continuous recording of HR, HRV, IBI, and BVP, has expanded opportunities to assess stress across real-world settings. Compared with traditional

ECG systems, PPG offers higher usability and scalability, making cardiovascular parameters practical biomarkers of cumulative stress load, resilience, and recovery.

Metabolic parameters

Metabolic indicators describe the body's energy turnover and substrate utilization, both of which strongly influence thermoregulation. Oxygen uptake (VO_2) provides a direct measure of aerobic metabolism, while energy expenditure represents the total energetic cost of sustaining activity and thermoregulatory processes. Blood lactate concentration serves as a marker of anaerobic metabolism and indicates when energy demand exceeds aerobic capacity. The respiratory exchange ratio (RER), calculated as the ratio of carbon dioxide production to oxygen consumption, reflects the balance between carbohydrate and fat oxidation. Together, these parameters capture how the body generates and mobilizes energy under different thermal and stress conditions. The most widely adopted method is indirect calorimetry, where VO_2 and carbon dioxide output are monitored to estimate both energy expenditure and RER [89]. This can be performed with laboratory-based metabolic carts or with portable spirometry devices in field settings. Blood lactate is typically obtained through fingertip or earlobe capillary sampling and analyzed with handheld lactate analyzers [196] In addition, wearable metabolic sensors that integrate ventilation and accelerometer data are increasingly used to provide continuous, less intrusive estimates of energy expenditure in natural environments [44].

In relation to thermal comfort, metabolic activity governs how much heat is produced endogenously and, in turn, how individuals perceive thermal states. Oxygen uptake (VO_2) and energy expenditure directly reflect metabolic heat production: higher VO_2

during physical activity translates into greater internal heat load. Even at rest, thermal conditions alter basal metabolism. In cold environments, metabolism rises through non-shivering thermogenesis, increasing VO_2 to generate additional heat. Prolonged cold also elevates the RER due to shifts toward carbohydrate utilization[197]. In heat, metabolic activity may be suppressed to minimize excess endogenous heat, but the supporting adjustments in circulation and sweat production still impose a greater subjective load even under moderate workloads [43]. Rising VO_2 and energy expenditure are consistently associated with perceptions of strain or overheating, while efficient metabolic regulation corresponds with greater comfort.

Within the TCPD framework, metabolic parameters highlight trade-offs in physical and cognitive productivity. For physical work, elevated VO_2 initially supports sustained output and is beneficial in cold, but excessive energy expenditure accelerates glycogen depletion and promotes early fatigue in heat [89]. Accumulated lactate signals reliance on anaerobic metabolism and marks the physiological boundary where work capacity is compromised, especially when combined with rising T_c [157]. For cognitive tasks, high metabolic cost indirectly impairs performance by diverting resources toward thermoregulation rather than neural processing. Studies show that elevated VO_2 and lactate during heat stress correlate with slower response speed, impaired working memory, and reduced problem-solving accuracy [111]. Metabolic measures thus delineate when the energetic burden of thermoregulation transitions from adaptive to detrimental for productivity.

From the perspective of stress, metabolic responses reflect systemic energy mobilization under both thermal and psychological challenge. Acute heat or cold exposure alters VO_2, RER, and lactate, signaling intensified energy turnover.

Psychological stressors, including workload anticipation, performance stress, and emotional arousal, similarly elevate resting metabolic rate via sympathetic–adrenal activation, increasing VO_2 and shifting substrate use [163]. Chronically elevated energy expenditure represents cumulative allostatic load, whereby persistent thermoregulatory or psychological strain reduces resilience and recovery capacity [198]. Viewed in this way, metabolic measures capture not only the physical cost of maintaining thermal homeostasis but also the energetic dimension of stress, providing a key complement to cardiovascular and EDA indices.

Other physiological parameters

A number of less frequently applied physiological indicators, such as blood glucose, blood oxygen saturation (SpO_2), cortisol, and blood pH, offer supplementary perspectives beyond the mainstream cardiovascular and metabolic measures. Compared with HR or T_c, these parameters are generally more difficult to assess in routine studies, since most require invasive sampling, laboratory assays, or specialized monitoring systems.

In terms of thermal comfort, fluctuations in glucose and cortisol underpin subjective experiences of energy depletion or arousal during heat and cold exposure, while minor changes in SpO_2 or pH may contribute to discomfort in extreme environments [89]. Within the TCPD framework, instability in glucose availability can compromise endurance and cognitive clarity, cortisol elevations can reduce long-term productivity by signaling chronic strain, and shifts in oxygen balance can further constrain attention and fine motor control under demanding thermal conditions [163]. From the perspective of stress, these additional parameters are particularly relevant, since both thermal and psychological stressors mobilize glucose, alter endocrine markers such as cortisol, and affect acid–base equilibrium, together reflecting the cumulative systemic burden of strain [44].

Looking forward, the incorporation of these additional parameters into thermal–psychological research remains limited but promising. Advances in noninvasive sensing technologies (e.g., sweat or saliva assays, wearable optics for tissue oxygenation) hold the potential to make glucose, cortisol, and oxygen dynamics more accessible in field studies. Integrating these indices with established metrics such as T_c, HR, and EDA could provide a more comprehensive understanding of how thermal, metabolic, and neuroendocrine systems interact to shape comfort, productivity, and resilience in complex environments.

Summary

Taken together, the physiological parameters discussed above illustrate multiple complementary layers of how the body responds to thermal and psychological challenges. T_c and T_{sk} offer direct readouts of thermal load, while EDA captures sweat gland activity and its immediate perceptual correlates. Cardiovascular measures such as HR, BP, and HRV reflect autonomic regulation and circulatory adjustments, and metabolic indicators like VO_2, lactate, and RER quantify the energetic cost of maintaining balance. In addition, more specialized markers including blood glucose, SpO_2, and cortisol extend the perspective to substrate availability, oxygen supply, and endocrine pathways, even though their measurement is relatively less practical in real-world contexts.

Across these domains, three themes emerge consistently. First, each parameter provides a physiological basis for thermal comfort and discomfort, helping explain why subjective sensations can diverge from core temperature alone. Second, within the TCPD framework, they undermine or support productivity, whether by constraining physical endurance, impairing dexterity, or limiting cognitive function. Third, they converge as indicators of stress, with shared autonomic and metabolic signatures that accumulate under both thermal and psychological loads.

While these physiological measures form the traditional foundation of thermal–psychological research, their application is often constrained by intrusiveness, instrumentation, or measurement continuity. To complement these approaches, recent years have seen increasing attention to image-based measurements, such as thermal infrared imaging and computer vision analysis, that promise unobtrusive, noncontact, and scalable alternatives for monitoring human thermal states. The following section introduces these image-oriented methods as a growing frontier in the assessment of comfort, productivity, and stress.

2.6.2 Image-based parameters

Traditional physiological monitoring approaches, including thermocouples for skin temperature, electrodes for cardiovascular or electrodermal activity, and mouthpieces for metabolic assessment, provide accurate information on human thermal and stress responses. However, these methods often require direct sensor contact, regular calibration, or even invasive sampling, which limits their feasibility in large-scale studies and everyday environments. In real-world contexts such as workplaces, classrooms, or public spaces, there is a clear demand for unobtrusive, scalable, and continuous monitoring techniques that can minimize user burden while still capturing physiologically meaningful data.

In response to this need, image-based approaches have become increasingly attractive. Rather than relying on physical sensors attached to the body, they use optical and neuroimaging techniques to provide indirect yet informative measures of human states. Infrared thermography makes it possible to track skin surface temperature across the face or body without contact, thereby revealing localized vasomotor and thermoregulatory activity. Facial video analysis with standard RGB cameras can detect subtle variations in expression and skin color, which have been linked to discomfort, fatigue, and emotional reactivity. Electroencephalography and functional near-infrared spectroscopy, although still dependent on wearable electrodes or optodes for signal acquisition, generate spatial topographic maps that are used to visualize changes in brain activity across regions. These image-like readouts extend the scope of image-based parameters, especially for understanding cognitive and affective processes under thermal challenge.

Taken together, these imaging approaches provide new opportunities to evaluate the three central dimensions of this review. They contribute to the study of thermal comfort by mapping bodily and facial thermal signatures, to the TCPD framework by detecting signs of physiological fatigue and cognitive decline, and to the assessment of stress through the visualization of emotional and neural activation patterns. With the rapid progress of computer vision and machine learning, image-based parameters are likely to become an essential complement to traditional physiological measurements, offering richer and less intrusive assessments of human responses in thermally and psychologically challenging conditions.

Infrared thermography (IRT)

Infrared thermography is one of the most widely applied image-based techniques for noncontact physiological monitoring. It relies on the detection of infrared radiation naturally emitted from the skin surface, converting these emissions into high-resolution thermal images that represent temperature distribution across the body [38]. Because skin temperature is tightly regulated through vasomotor adjustments and peripheral blood flow, IRT provides a sensitive window into thermoregulatory processes. Compared to conventional temperature probes, it allows the visualization of heterogeneous regional responses, for instance, distal extremities cooling faster than the trunk, without requiring multiple points of contact or invasive devices [39]. This spatial richness has made IRT a preferred method for both laboratory experiments and applied studies in indoor environments and occupational settings [199].

In relation to thermal comfort, thermal imaging is valuable for identifying localized patterns that underpin subjective sensations. Facial thermography in particular has demonstrated that cooling of the nasal tip, cheeks, or hands is consistently associated with reports of cold discomfort, while elevated temperatures in the forehead, periorbital, or neck regions correspond with perceptions of heat strain [39]. IRT can detect these shifts early, often preceding noticeable changes in core body temperature, thereby providing an objective link between localized skin temperature dynamics and whole-body comfort assessments [44]. Furthermore, thermal asymmetries across facial or limb regions have been used to explain why individuals report heterogeneous comfort sensations within the same ambient environment [39].

Within the TCPD framework, IRT contributes to understanding both physical and cognitive dimensions of productivity. On the physical side, whole-body thermograms can reveal progressive heat storage that predicts earlier onset of fatigue, reductions in muscular endurance, and declining precision in manual tasks [163]. In occupational contexts, elevated skin temperatures measured by IRT have been correlated with productivity losses in construction, manufacturing, and agriculture [133]. On the cognitive side, changes in facial thermal distributions, such as reductions in frontal or nasal temperatures under both heat and cold exposure, have been linked to attentional decrements, slower reaction times, and a higher incidence of errors in vigilance and memory tasks [112]. This evidence shows that thermal imaging can serve as a noninvasive early warning indicator of cognitive and physical performance decline under thermal stress [37].

From the perspective of stress, IRT has been widely employed to identify autonomic signatures of psychological activation. A frequently reported pattern is pronounced nasal tip cooling during acute stress, reflecting sympathetic vasoconstriction, in parallel with periorbital warming driven by elevated blood flow during emotional arousal [200]. These signatures have been observed not only in controlled laboratory paradigms but also in applied contexts such as driver monitoring, security screening, and classroom evaluations [119]. Thermal signatures are reproducible, highly sensitive to momentary fluctuations, and independent of verbal self-reports, making them especially valuable for situations where overt behavioural signs of stress are subtle or concealed[38].

Taken together, IRT demonstrates a dual potential: it provides objective, noncontact measurement of thermal comfort through monitoring of surface temperature dynamics, while also capturing physiological and affective dimensions of productivity and stress. Its scalability, combined with recent advances in portable infrared cameras and automated image analysis, suggests that IRT will remain a cornerstone of image-based methods for human thermal–psychological research[201].

Facial video imaging

Facial video analysis using standard RGB cameras has emerged as another important image-based method for monitoring human thermal and psychological responses. Unlike infrared thermography, which primarily captures surface temperature, facial video focuses on subtle visual signals such as micro-expressions, muscle activations, and changes in skin coloration. These signals are closely connected to autonomic activity, vascular responses, and affective states, making them a rich source of information on comfort, productivity, and stress [44]. The advantage of this approach is that it uses widely available video technology and can be integrated into daily environments without the need for specialized thermal sensors[202].

In relation to thermal comfort, facial video imaging can capture expression changes that accompany sensations of heat or cold. Elevated skin redness, particularly around the cheeks and forehead, is associated with increased peripheral blood flow during heat exposure, while pallor or tightening of facial muscles is commonly observed in cold environments [89]. These subtle yet observable expression patterns correspond with subjective comfort ratings and provide indirect but valuable cues on environmental adequacy. Moreover, facial video analysis has been explored in HVAC and indoor climate studies as a nonintrusive way to predict discomfort before individuals consciously report it [203].

Within the TCPD framework, this technique has been applied to monitor productivity in both physical and cognitive domains. For physical activities, video-based analysis can reveal indicators of thermal strain such as facial flushing, sweating sheen, or micro-expressions of exertion, which correlate with declining endurance or impaired manual performance [39]. In cognitive tasks, facial dynamics such as frequent frowning, blinking, or gaze aversion have been linked with attentional lapses, mental fatigue, and reduced task accuracy under thermal load [31, 39]. Advances in computer vision and machine learning have further enabled automatic classification of facial features, allowing the prediction of productivity decline in real-time monitoring systems [204].

From the perspective of stress, facial video imaging provides an accessible and non-invasive way to detect emotional arousal and autonomic activation. Classical studies of micro-expressions have shown that brief, involuntary facial movements reliably reveal stress, stress, or concealment, even when individuals attempt to suppress overt responses [38, 39]. Under thermal or psychological stress, facial cues such as increased blinking, lip compression, or furrowing of the brow can serve as reliable windows into sympathetic activation [38]. Importantly, these features can be analysed at scale, making video imaging particularly relevant for group or organizational contexts where individual sensor attachment is impractical.

Overall, facial video analysis offers a versatile and widely accessible tool for connecting visual signs with thermal comfort, performance outcomes, and stress states. Its integration with automated recognition systems and multimodal data fusion (e.g., combining with speech or posture analysis) further underscores its potential as a component of comprehensive human–environment monitoring frameworks [66].

Electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) topographic mapping

Electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) occupy a slightly different position within image-based approaches. Both techniques still require wearable sensors for signal acquisition, EEG through scalp electrodes and fNIRS through optodes measuring haemodynamic responses in cortical tissue, but their outputs are commonly presented as topographic maps that visualise spatial patterns of neural activity [205]. By providing a regional overview of how brain activity changes under environmental or psychological challenges, these maps align with the broader concept of image-based parameters.

In relation to thermal comfort, EEG and fNIRS mapping have helped to uncover the neural correlates of thermal sensation and dissatisfaction. Alterations in prefrontal and parietal activation have been shown to correspond with subjective reports of discomfort in hot or cold conditions [31]. Such spatially resolved neural data extend the understanding of comfort beyond peripheral physiological measures, shedding light on the central processes that mediate the perception of the thermal environment.

Within the TCPD framework, these mapping techniques are particularly applicable to cognitive productivity. EEG studies demonstrate that exposure to high thermal load is linked to reduced power in alpha rhythms and increased theta activity, both markers of reduced attentional capacity and higher mental fatigue[31]. fNIRS mapping similarly indicates diminished prefrontal oxygenation under heat stress, which is associated with poorer executive function and slower decision-making [206]. These findings underscore the value of neuroimaging-based outputs as indicators of when thermal strain begins to compromise sustained cognitive performance.

From the perspective of stress, EEG and fNIRS topographic representations provide well-established markers of arousal, stress, and emotional load. Stressful states are typically accompanied by increased frontal asymmetry in EEG or elevated prefrontal haemodynamic responses in fNIRS, even when overt behavioural cues are absent [115]. Because these signatures are spatially visualised, they can capture not only the intensity but also the cortical distribution of stress-related activity, offering richer insights than traditional unidimensional measures.

In summary, EEG and fNIRS topographic mapping represent a complementary, image-like class of parameters. Although limited by the need for wearable hardware, their capacity to visualise central nervous system responses positions them as a crucial bridge between peripheral physiological measures and subjective psychological outcomes. As technologies for portable neuroimaging mature, these approaches are expected to become increasingly integrated into multimodal frameworks for assessing comfort, productivity, and stress [115].

Summary

The review of image-based approaches demonstrates that they provide valuable, complementary perspectives on human thermal and psychological responses. Infrared thermography enables the detection of surface temperature dynamics that underpin local thermal sensations and stress-related vasomotor shifts. Facial video imaging captures micro-expressions and subtle colour changes that reflect discomfort, fatigue, and emotional strain in a non-invasive and widely accessible manner. EEG and fNIRS topographic mapping extend the analysis inward, visualising neural activity patterns that relate directly to cognitive performance and stress regulation.

Taken together, these techniques broaden the methodological toolkit for assessing the three central dimensions of interest. For thermal comfort, they map physiological and expressive signatures that precede or accompany subjective reports. Within the TCPD framework, they offer early-warning indicators of declining productivity in both physical and cognitive domains. From the perspective of stress, they uncover autonomic, facial, and neural markers that highlight the convergence of thermal and psychological demands.

Despite clear advantages of being non-contact or image-like in their readout, challenges remain, including sensitivity to environmental conditions, reliance on specialised hardware, and ethical considerations tied to visual monitoring. Nonetheless, with continued advances in computational imaging and machine learning, image-based parameters are expected to evolve into robust, unobtrusive complements to conventional physiological methods, enabling richer assessments of comfort, productivity, and stress in real-world settings.

2.6.3 Summary

The review of both physiological and image-based parameters illustrates the breadth of tools currently available for investigating the interplay between thermal comfort, productivity, and stress. Physiological markers such as core and skin temperature, electrodermal activity, cardiovascular indices, and metabolic measures provide well-established and mechanistically grounded evidence of how the human body responds to environmental and psychological challenges. They yield precise, quantitative data and have a long record of validation across laboratory and occupational studies. However, their use is often limited by issues of invasiveness, the need for sensor attachment, and difficulties in continuous implementation outside controlled settings.

Image-based approaches, including infrared thermography, facial video analysis, and EEG or fNIRS topographic mapping, address some of these constraints by offering non-contact, spatially resolved, or visually interpretable readouts. These methods broaden the capacity to assess comfort, productivity, and stress in social or real-world contexts where attaching multiple sensors is impractical. At the same time, they remain sensitive to technical conditions such as camera resolution, lighting, or motion artefacts, and they raise new challenges in terms of data privacy and ethical use.

Taken together, these two categories of parameters are best viewed as complementary. Physiological measures provide established accuracy and deep mechanistic insight, while image-based parameters supply scalability, unobtrusive monitoring, and spatial richness. A combined, multimodal approach is therefore likely to capture the complexity of human responses more effectively than any single method alone.

On this basis, the next section turns to analytical and experimental methodologies (2.7). If physiological and image-based parameters determine what can be measured, methodologies determine how these measurements are obtained, processed, and interpreted. It is through carefully designed protocols, integrating laboratory accuracy with ecological validity, and combining statistical and computational analysis, that the full potential of these parameters can be realised. Section 2.7 therefore outlines the methodological frameworks that underpin rigorous and meaningful use of these diverse measures in the study of thermal comfort, productivity, and stress.

2.7 Analytical and experimental methodologies

The preceding sections reviewed a wide spectrum of physiological and image-based parameters, each of which provides different vantage points on the complex interrelations among thermal comfort, productivity, and stress. Yet, parameters alone cannot generate knowledge: their value depends upon how they are captured, processed, and interpreted. The methodological dimension is therefore the essential bridge between human responses and scientifically valid conclusions.

This section considers methodologies in a broad sense, encompassing experimental design, analytical techniques, and computational modelling. Experimental methodologies determine the ecological context and degree of control under which data are gathered, ranging from tightly regulated laboratory studies to ecologically valid field investigations. Analytical techniques shape the extraction of meaningful features from complex physiological and image-based signals and underpin the validity of subsequent inference. Finally, modelling approaches, particularly those using machine learning and deep learning, enable the translation of multimodal data into predictive insights, adaptive control strategies, and theoretical advances.

By structuring the review in this order, beginning with experiment, progressing through analysis, and culminating in modelling, the discussion follows the natural trajectory from data collection to knowledge construction. This logic reflects the increasing emphasis on integrative and data-driven research strategies, and provides a foundation for identifying methodological strengths, limitations, and opportunities for future development in the study of thermal comfort, TCPD, and stress.

2.7.1 Experimental methodologies

The methodological design of experiments is central to advancing the understanding of thermal comfort, productivity, and stress. Different approaches, laboratory, field, and hybrid, offer varying balances between environmental control, data richness, and ecological validity. This section elaborates on the major traditions of experimental research, illustrating key features, strengths, and limitations through concrete examples.

Laboratory-based experiments

Laboratory studies are most often conducted in climate chambers or controlled test rooms where temperature, relative humidity, air velocity, and radiant load can be programmed with high precision [172]. Such facilities support both step-change protocols, for instance, exposing subjects to sequential stages of 22 °C, 26 °C, and 30 °C in order to map the trajectory from thermal neutrality to heat stress [82], and steady-state protocols, in which subjects remain under one condition for extended periods to assess adaptation and fatigue [111]. By enforcing strict control over extraneous variables, laboratory experiments allow researchers to isolate the causal effect of thermal conditions on physiological and psychological outcomes and to replicate findings across subjects or sessions.

Typical laboratory tasks are designed to probe multiple aspects of human response. Physical activity protocols such as treadmill walking, cycle ergometry, or load-lifting are used to vary metabolic rate and hence internal heat production [163]. Cognitive testing paradigms, including Stroop interference tasks, N-back working memory loads, and psychomotor vigilance tasks, are employed to characterise attentional control and executive function under different thermal challenges[31]. Subjective feedback is also integral to these sessions, with subjects asked at regular intervals to provide thermal sensation votes, thermal acceptability ratings, or perceived exertion scores, ensuring the alignment of objective measures with phenomenological experience[61]. Multimodal recording is routine: skin and core temperatures, electrodermal activity, heart rate variability, and respiration are complemented by infrared thermography to capture spatial skin temperature distributions, and by neuroimaging techniques such as EEG or fNIRS to explore cortical correlates of thermal comfort and stress. A paradigmatic setup involves subjects completing working memory tasks while under manipulated climate conditions, enabling simultaneous analysis of autonomic stress indicators, neurocognitive performance, and subjective comfort [206].

The strength of laboratory-based methodologies lies in their precision, replicability, and capacity to establish causal inference. They are particularly well suited to mechanistic investigations, model validation, and the exploratory testing of novel sensors. However, they are also constrained by their artificiality, short duration of exposure, and restricted subject pools, often composed of university students rather than representative occupational groups. Consequently, questions concerning ecological validity, generalisation, and long-term behavioural adaptation remain difficult to address within the laboratory alone [30].

Field-based experiments

Field investigations extend the methodological landscape by situating subjects within real-world contexts and monitoring their responses under everyday conditions. Offices, classrooms, healthcare facilities, workshops, construction sites, and transport systems have all served as natural laboratories for research on comfort, productivity, and stress [46]. Rather than manipulating variables in strict increments, field studies embrace naturally occurring fluctuations in climate as well as the presence of co-occurring stressors such as acoustic noise or air pollution. Longitudinal office studies are a common design, tracking employees across seasonal variations and collecting repeated data on thermal perception, physiological state, and performance indicators such as error rates, typing speed, or decision-making efficiency [207]. In education, studies have investigated how classroom temperature and ventilation affect not only students' immediate comfort but also learning outcomes, test scores, and behavioural markers of attention and engagement [208]. Outdoor occupational research has similarly examined construction workers or agricultural labourers during summer heat events, integrating wearable heart-rate monitors, actigraphy, and portable thermal cameras to document heat strain in situ [44].

Field research relies heavily on portable and unobtrusive instrumentation, including wristbands for measuring electrodermal activity, chest straps for heart rate variability, wearable accelerometers for physical activity, and compact cameras for facial thermography. These devices enable continuous or long-duration monitoring, allowing fine-grained temporal data aligned to real tasks and social contexts. Nonetheless, data recorded under field conditions are prone to noise from movement artefacts, variability in sensor positioning, and the confounding influence of unmonitored environmental factors. Ethical considerations are also more prominent, particularly in workplaces where continuous video or biometric surveillance must be carefully governed to protect privacy and autonomy [209]. Despite these challenges, field studies remain indispensable because they provide direct insight into how comfort, productivity, and stress unfold together within the dynamic complexity of daily life.

Hybrid experiments

Recognising the complementary benefits and limitations of laboratory and field strategies, researchers have increasingly developed hybrid designs. A well-established tradition is the use of semi-controlled "living laboratories" that reproduce real-world environments, such as open-plan offices or classrooms, while retaining precise HVAC management and integrated monitoring infrastructure[27]. Within these spaces, subjects pursue typical workflows, typing, meetings, collaborative projects, while researchers adjust temperature, ventilation, or lighting. This setup permits simultaneous measurement of environmental variables, physiological signals, and behavioural outcomes, such as keystroke dynamics, error rates, or group decision-making quality.

A parallel line of innovation is the adoption of virtual reality and immersive simulation technologies. Here, subjects are placed in VR environments that convincingly render classrooms, outdoor plazas, or public transport interiors, while thermal and airflow conditions are modulated in the physical laboratory [135]. Such integration achieves a high degree of ecological plausibility while maintaining sufficient control to permit experimental manipulation. These emerging methodologies are of particular relevance for testing adaptive comfort systems, in which data streams from physiological or image-based sensors feed directly into HVAC control loops to regulate indoor conditions automatically [44]. In this sense, hybrid and ecological approaches not only bridge laboratory and field but also act as testbeds for next-generation technologies, where questions of feasibility, ethical deployment, and user acceptance can be addressed before large-scale real-world implementation.

Summary

Taken together, laboratory, field, and hybrid methodologies trace a continuum from precision to realism. Laboratory experiments excel in allowing the careful isolation of mechanisms but suffer from limited ecological validity. Field studies reflect the true complexity and variability of human–environment interaction but require methodological sophistication to manage confounding influences. Hybrid approaches occupy the middle ground, increasingly highlighted as a methodological frontier for combining external relevance with experimental rigour. A multi-tiered research strategy, drawing systematically upon all three traditions, is therefore needed to build a comprehensive evidence base for the interrelated phenomena of thermal comfort, productivity, and stress [210].

2.7.2 Analytical methodologies

The transformation of experimental observations into reliable knowledge requires rigorous analytical methodologies. Physiological and image-based parameters, though rich in information, are inherently noisy and heterogeneous. Signals differ in scale, temporal resolution, and susceptibility to artefacts, while subjective ratings add an additional layer of variability. To cope with these challenges, analytical techniques in this field have evolved along three main lines: classical statistical modelling, signal processing and feature extraction, and multimodal integration. Together these methods determine the extent to which thermal comfort, TCPD, and stress can be meaningfully characterised and predicted.

Statistical analysis

Statistical methods provide the interpretative backbone for much of the literature. At the most basic level, descriptive statistics summarise distributions of thermal sensation, productivity scores, or physiological responses, while inferential tests such as t-tests and ANOVA quantify differences between groups or conditions[211]. Regression analysis is widely used to explore how environmental variables predict outcomes, for example, modelling the linear association between air temperature and reported comfort, or the nonlinear relationship between humidity and reaction times[212]. In longitudinal or field studies, data often exhibit hierarchical structures, such as repeated measures nested within individuals, or individuals nested within offices or classrooms. Mixed-effects models and multilevel modelling frameworks have therefore become increasingly common, allowing both fixed effects of environmental predictors and random effects of individuals or groups to be estimated simultaneously[12]. These methods enable researchers to account for intra-individual variability and inter-individual heterogeneity, providing a more nuanced understanding of when performance or wellbeing deteriorates. Moreover, statistical approaches underpin threshold identification, for instance estimating neutral temperature zones for comfort, or critical exposure points beyond which stress indicators rise sharply. Their strength lies in transparency and interpretability, though reliance on linearity and distributional assumptions can sometimes limit sensitivity to complex dynamics.

Signal processing and feature extraction

Laboratory and field experiments routinely generate large volumes of physiological and image-based signals that require preprocessing before analysis. Heart rate–derived indices are a canonical example: raw inter-beat intervals are cleaned of artefacts, interpolated to regular sampling rates, then decomposed into time-domain features (e.g. SDNN, RMSSD) and frequency-domain components through Fourier or wavelet transforms, yielding indicators of autonomic balance [213]. Electrodermal activity is detrended and separated into tonic and phasic signals, the latter identifying discrete sympathetic arousals such as skin conductance responses associated with discomfort or attentional shifts [214]. EEG signals are segmented into epochs, filtered, and transformed into spectral power within delta, theta, alpha, and beta bands, with relative ratios often interpreted as markers of workload or relaxation [31]. Nonlinear measures such as approximate entropy or fractal scaling coefficients provide added sensitivity to system complexity, capturing dynamics not visible in mean-based metrics.

Image-based modalities follow a parallel route of feature extraction. Infrared thermograms are often segmented into facial or regional areas of interest, with mean and maximum temperatures, asymmetry indices, or temporal gradients computed for each [39]. Facial video recordings are analysed for colour changes in skin regions related to peripheral blood flow, or for muscle activation patterns expressed as facial action units associated with discomfort and emotional strain [38]. The central goal of these procedures is to engineer a reduced set of reproducible, interpretable features that represent the physiological or expressive correlates of thermal and psychological states. Without this stage of structured feature extraction, downstream analyses would be overwhelmed by dimensionality and artefacts, particularly in studies intending to move towards predictive modelling.

Multimodal integration

Modern ergonomics and environmental psychology studies increasingly employ multimodal measurement frameworks, which combine physiological, behavioural, and image-based indicators. As a result, data integration has become a methodological frontier. Multimodal datasets are challenging because signals differ in units, dynamic ranges, and sampling frequencies; synchronisation of time stamps, interpolation of missing values, and normalisation across modalities are necessary steps before integration [66]. Once harmonised, several strategies are available. Early fusion combines features from different modalities into a single dataset fed into statistical or machine learning models, maximising the chance to detect cross-modal associations. Late fusion aggregates the outputs of modality-specific models at the decision stage, useful when datasets differ in size or quality. Hybrid techniques combine elements of both, sometimes employing representation learning or canonical correlation analysis to detect shared latent factors across modalities [215].

Empirical studies illustrate the benefits of multimodal integration. For example, combining EEG spectral features reflecting cognitive workload, HRV indicators of autonomic balance, and infrared thermal asymmetries of the face has been shown to improve prediction accuracy of task performance under thermal stress compared to any single measure [31]. Similarly, blending electrodermal responses with facial expression analysis can yield sensitive markers of acute stress that would be overlooked by self-report alone. The advantage of these integrative approaches lies in their holistic scope, capturing central nervous, autonomic, and expressive domains simultaneously. However, such richness also introduces risks. Overfitting becomes a salient concern when high-dimensional multimodal data are applied to modest sample sizes, and interpretability decreases as fusion complexity increases. As such, multimodal techniques demand careful experimental validation, rigorous cross-validation routines, and transparent reporting to ensure robustness and reproducibility.

Summary

Analytical techniques thus constitute the indispensable middle layer between data collection and predictive modelling. Statistical methods ground the field in established inferential logic and enable threshold identification. Signal processing and feature extraction transform high-dimensional raw biosignals and images into interpretable indices of state and performance. Multimodal integration opens a systems-level perspective, increasing sensitivity and ecological validity by jointly analysing diverse indicators. Collectively these methods provide the analytical infrastructure for interpreting data from experimental studies and preparing robust inputs for the machine learning and deep learning models described in the next section.

2.7.3 Modelling methodologies

Experimental design and analytical techniques provide the data and features required for understanding human responses, but it is through modelling that these insights are organised into predictive and generalisable frameworks. Modelling approaches enable the translation of multimodal signals into decision-support tools, adaptive control systems, and theoretical advances concerning the interplay between thermal comfort, productivity, and stress. In this area, three main currents can be distinguished: classical machine learning algorithms, deep learning architectures, and multimodal and emerging modelling directions.

Classical machine learning

Traditional machine learning methods have been widely applied in comfort and stress research due to their relative simplicity, robustness under limited sample sizes, and interpretability. Algorithms such as decision trees, random forests, k-nearest neighbours, logistic regression, and support vector machines are particularly effective in classification tasks, for instance distinguishing between comfortable and uncomfortable states or between high and low stress conditions based on physiological features [31]. Regression-based models have been used to predict continuous outcomes such as productivity metrics or predicted mean vote (PMV) scores from sets of environmental and physiological variables [19]. These models are advantageous because they allow clear examination of feature importance, supporting theoretical understanding, and they are computationally efficient, readily deployed in embedded systems or real-time applications. Yet they are less capable of capturing nonlinear and highly complex relationships across large, heterogeneous datasets, limiting their applicability as multimodal and image-based measures gain prominence.

Deep learning models

With the increasing availability of high-dimensional data, deep learning has become a powerful tool, particularly for image-based and sequential physiological data. Convolutional neural networks (CNNs) have been applied to infrared thermography and facial video, enabling the automatic detection of subtle temperature patterns or micro-expressions that correspond to discomfort and stress, without requiring manual feature engineering [38]. Recurrent neural networks (RNNs) and their extensions such as Long Short-Term Memory (LSTM) networks are suited to processing temporal signals such as HRV, EDA, or EEG, capturing time-dependent dynamics of stress onset and thermal adaptation [216]. More recently, transformer-based architectures have been explored for their ability to model long-range dependencies across multimodal physiological sequences [217]. These methods provide remarkable flexibility in uncovering complex and nonlinear patterns, and they can outperform classical models in predictive accuracy. However, they are data-hungry, requiring large, diverse, and labelled datasets, and they demand substantial computational resources. Furthermore, issues of model transparency and interpretability remain ongoing challenges, especially in applied contexts such as workplaces or classrooms where stakeholders require understandable explanations of system decisions.

Multimodal and emerging approaches

Given the multidimensional character of comfort, productivity, and stress, models that integrate multiple data channels are now at the forefront of research. Multimodal modelling strategies include early fusion, where features from physiological, environmental, and image-based signals are combined into a single vector, and late fusion, where independent models for each modality are trained and their decisions aggregated [162]. Hybrid approaches employing attention mechanisms or representation learning seek to exploit both modality-specific and shared latent factors. Empirical studies have, for instance, demonstrated that combining EEG spectral data with heart rate variability and thermal imaging significantly enhances predictive accuracy in identifying cognitive fatigue compared with unimodal models [115]. Other emerging methodologies include transfer learning, in which pretrained models are adapted to new subject groups or environments, thereby addressing the common issue of small and heterogeneous datasets; and personalisation strategies, where models are tuned to individual baselines to improve sensitivity to idiosyncratic responses. In parallel, explainable AI (XAI) frameworks are increasingly integrated with modelling pipelines, offering feature attributions, saliency maps, or local surrogate models to enhance interpretability and trust, which are vital for practical uptake [218].

Summary

Modelling approaches mark the stage where experimental measures and analytical features are consolidated into predictive insights. Classical machine learning provides interpretable and resource-efficient tools, deep learning expands the frontier towards complex and high-dimensional data, and multimodal frameworks allow rich integration across central, autonomic, and expressive domains. The field is now moving towards models that are both powerful and transparent, capable of capturing nonlinearities without sacrificing explainability. Collectively, these approaches extend the methodological toolkit available for investigating comfort, productivity, and stress, and they underpin the development of adaptive systems that adjust environments in real time according to human needs.

2.7.4 Summary

The methodological review presented in this section has illustrated how experiments, analytical procedures, and modelling collectively shape the investigation of thermal comfort, productivity, and stress. Experimental methodologies provide the raw foundation, with laboratory studies ensuring precision and causal inference, field investigations capturing ecological validity, and hybrid or ecological approaches balancing these aims while serving as testbeds for emerging technologies. Analytical techniques transform the resulting biosignals and images into structured descriptors, ranging from classical statistical models that detect thresholds and associations, through sophisticated signal processing that extracts dynamic features, to multimodal frameworks that integrate diverse sources of data. Building upon this layered foundation, modelling approaches consolidate features into predictive systems, enabling classification, regression, and real-time adaptation through both traditional machine learning and advanced deep learning architectures, while multimodal and explainable strategies enhance predictive accuracy and user trust.

Taken together, these methodological components trace a coherent trajectory: from careful experimental design, through rigorous analytical treatments, to powerful, generalisable models capable of adaptive application. This integrated perspective underscores that no single layer is sufficient: precision in experimentation is diminished without adequate analysis, analytical insights remain static without predictive modelling, and models lose salience in the absence of ecologically grounded data. The methodological continuum outlined here thus provides a comprehensive toolkit for advancing the study of comfort, productivity, and stress, and it establishes a direct bridge to the next stage of inquiry, where analytical–experimental integration is connected with real-world implementation and the design of adaptive, human-centred environments.

2.8 Summary and research gaps

This chapter reviewed current evidence on how thermal environment, task characteristics, and individual factors jointly shape thermal comfort, productivity, and stress. Section 2.2 described how deviations from thermal neutrality alter physiological regulation and subjective comfort, often impairing performance. Section 2.3 demonstrated that task factors such as workload intensity, complexity, and duration independently modulate cognitive efficiency and stress responses. Section 2.4 emphasised that thermal and task demands rarely act in isolation: their combined effects on the task cognitive–performance domain (TCPD) and stress are interactive and often nonlinear. Section 2.5 highlighted the moderating role of individual differences in age, sex, BMI, acclimatisation, and psychological dispositions. Section 2.6 examined both physiological and image-based measurement parameters, showing their complementary roles but also their limitations. Section 2.7 outlined methodological practices, from laboratory to field to hybrid experiments, from statistical analysis to multimodal integration, and from classic machine learning to advanced deep learning. Collectively, these reviews present a multi-layered overview of how comfort, productivity, and stress have been studied, and underscore the need for more integrative and comprehensive approaches.

Despite the breadth of existing work, several shortcomings remain:

1. Fragmented focus. Much of the literature considers either the thermal environment or task demands in isolation, while relatively few studies examine their joint effects. Even when combined designs exist, they often underrepresent the role of individual variability, treating it as statistical noise rather than as a meaningful explanatory variable. This limits the capacity to capture the actual interplay among environment, task, and personal factors, which is critical for understanding TCPD and stress in realistic contexts.

2. Experimental limitations. Laboratory studies allow precise control and causal inference but at the cost of external validity, as subjects are typically exposed for short periods under artificial conditions. Field studies, in contrast, capture real settings but introduce uncontrolled variability and confounds, which weaken causal interpretation. Hybrid or ecological laboratory designs offer a solution but remain in their infancy, with limited standardisation and inconsistent application across studies. This leaves a methodological gap between precision and realism.

3. Monomodal measurement. Many studies still rely solely on questionnaires or a single physiological indicator. Such approaches cannot reflect the multidimensional nature of human thermal and psychological responses. Attempts at multimodal measurement are increasing but often lack systematic integration, and feature extraction practices are inconsistent, making comparisons across datasets difficult. As a result, the full potential of multimodal synergies is rarely realised.

4. Analytical constraints. Conventional inferential statistics dominate, focusing on average effects across conditions and subjects. While valuable, these methods often fail to detect dynamic, nonlinear, and cross-level interactions that characterise complex human responses. More advanced signal processing approaches, such as spectral analysis of EEG or decomposition of electrodermal activity, are applied inconsistently across studies and rarely standardised, reducing replicability and comparability of findings.

5. Modelling challenges. Predictive modelling remains underdeveloped in this field. Machine learning methods are usually confined to single-output classifications such as comfort vs. discomfort, and do not incorporate broader outcomes like productivity or stress. Deep learning approaches, although powerful, are constrained by small sample sizes typical of laboratory ergonomics experiments, raising risks of overfitting. Moreover, model outputs are often non-transparent, limiting trust and practical application. Integrated models that jointly predict comfort, work performance, workload, and stress are rarely attempted.

6. Practical and ethical issues. Implementation in real contexts is complicated by concerns about privacy in image-based monitoring, comfort and compliance with wearable sensors, and the acceptability of algorithm-driven environmental adaptation. Without systematic research into these aspects, even technically advanced systems may face resistance in occupational or educational settings.

The present dissertation is designed to address these limitations in an integrated way. Thirty-two subjects are studied under three systematically varied thermal conditions (PMV -1, 0, +1) while performing cognitive tasks of graded difficulty. Multimodal data, including physiological measures, facial thermography, behavioural outcomes, and subjective assessments, are collected under repeated-measures conditions. Bayesian analyses test hypotheses about the effects of environment and task on comfort, productivity, workload, and stress, while machine learning and deep learning models explore predictive strategies. The modelling framework is deliberately layered: beginning with demographic and environmental variables, extending to task complexity, physiological indicators, and finally image-based features, with both single- and multi-output predictions. By combining experimental control with ecologically relevant tasks and multimodal measures with advanced yet interpretable modelling, this research directly responds to the identified gaps and contributes a more holistic, predictive, and human-centred framework for understanding thermal comfort, productivity, and stress.

Chapter 3 Methodology

3.1 Introduction

3.1.1 Scope and structure

This chapter outlines the methodological approach taken to address the research questions identified in Chapter 1. In order to generate meaningful and reliable answers to those questions, the study draws on the insights from Chapter 2, where the literature review revealed key limitations in existing knowledge. In particular, previous studies often examined thermal comfort, cognitive task performance, and stress in isolation, relied heavily on subjective measures, and commonly applied frequentist statistics without fully integrating multimodal response data. These gaps justified the adoption of a controlled experimental design that combines subjective, behavioural, physiological, and imaging measures, analysed using Friedman test inference and developed into predictive machine learning and deep learning models. In doing so, the methodology is directly aligned with the research objectives and strategically targeted at filling the knowledge gaps identified in the literature.

The remainder of this chapter is organised into seven sections. Section 3.1 Introduction provides the rationale for the methodological choices and describes the overall research design. Section 3.2 Experiment and data collection details the laboratory setup, controlled conditions, and procedures used to obtain the different streams of data. Section 3.3 Data preparation and preprocessing explains the steps undertaken to clean, synchronise, and transform the multimodal datasets. Section 3.4 Statistical analysis methods introduces the non-parametric test for K related samples employed to examine the effects of thermal environment and task difficulty on thermal comfort, TCPD, and stress. Section 3.5 Model training and evaluation presents the predictive modelling pipeline, including the progressive integration of different feature groups and the assessment criteria for model performance. Section 3.6 Ethical considerations discuss the recruitment of subjects, informed consent, and data protection measures. Finally, Section 3.7 Summary synthesises the methodological decisions of this chapter and sets the stage for presenting the empirical findings in Chapter 4.

3.1.2 Background and rationale

The previous chapter established that human responses to the built environment are shaped not only by thermal conditions but also by the nature of the tasks being performed. Yet research in these areas has largely advanced in parallel. Thermal comfort studies often examine temperature effects independently of task-related factors, while workload studies typically control the environment without considering thermal influences. As a result, the combined impact of thermal environment and task demands remains insufficiently articulated.

Where concurrent effects have been explored, the role of individual variability, in age, sex, body mass, physiological regulation, or psychological disposition, has been either overlooked or treated as nuisance variance. This limits our ability to understand how different people experience and manage combined thermal and cognitive demands. Moreover, experimental traditions reveal an imbalance: laboratory studies provide strong control over conditions but often lack ecological validity, whereas field studies capture real-world settings but introduce confounds that hinder causal inference.

In terms of measurement, responses are still predominantly assessed through single-modal indicators such as questionnaires or isolated physiological signals. Few studies combine subjective, behavioural, physiological, and image-based markers in an integrated dataset. Similarly, analytical practice is dominated by classical statistical testing, which captures mean differences but cannot fully account for the interactive and nonlinear nature of combined environmental and task effects. Predictive modelling has begun to emerge, yet most approaches target only one outcome at a time, and the added value of multimodal inputs, including physiological and imaging data, has not been systematically quantified.

Against this backdrop, the present study is designed to fill these gaps. It focuses on thermal environment and task factors as the controlled experimental conditions, and investigates their influence on three sets of responses: thermal comfort, the task cognitive–performance domain (TCPD, comprising work performance and workload perception), and stress (with emphasis on anxiety). To achieve this, controlled laboratory experiments are conducted under three PMV-defined thermal environments and three levels of task difficulty. Multimodal datasets, including subjective reports, behavioural performance, physiological parameters, and thermal imaging, are collected. Data are then analysed using a Friedman test inferential framework to quantify specific condition effects, followed by predictive modelling strategies that integrate different feature groups. By placing thermal comfort, TCPD, and stress within the same analytical framework, the study aims to provide both empirical evidence and methodological innovation for understanding and anticipating human responses to combined environmental and task demands.

3.1.3 Research framework

This study is designed around a structured research framework that combines controlled experimentation, statistical analysis, and predictive modelling. The framework is illustrated in Figure 3.1. It comprises three main components: (i) experimental design and data collection, (ii) inferential statistical analysis, and (iii) predictive modelling.

Experimental design and data collection

The controlled conditions in this research are defined by two factors: thermal environment (three PMV-based settings: −1, 0, +1) and task difficulty (three levels: Stroop, arithmetic, and Same–Different tasks). These conditions are expected to influence a set of responses including thermal comfort, the task cognitive–performance domain (TCPD: work performance and workload perception), and stress (with focus on anxiety). To capture these outcomes, multimodal data are collected: subjective questionnaires, task performance indicators, physiological parameters (skin temperature, electrodermal activity, and heart rate), and thermal facial imaging.

Statistical analysis logic

The first layer of analysis employs a Friedman test inferential framework to systematically test the effect of thermal environment and task difficulty on each response. Specifically, two questions guide this layer:

- Within a given thermal environment, do different levels of task difficulty alter thermal sensation, satisfaction, and anxiety?

- Within a given task difficulty level, does thermal environment alter work performance, workload perception, and anxiety?

The Friedman test approach is chosen for its ability to quantify uncertainty and provide richer evidence than traditional frequentist methods, thereby addressing the gaps identified in the literature. This stage ensures that the experimental manipulations produce interpretable causal inferences before predictive models are developed.

Predictive modelling logic

The second layer of analysis extends beyond hypothesis testing to forecasting responses under varying conditions. The modelling framework is structured progressively:

Baseline models incorporating environmental variables, task difficulty, and personal characteristics to predict each response separately.

Extended models that add physiological data (ST, EDA, HR) to evaluate their contribution to predictive accuracy, generalisation, and computational efficiency.

Multimodal deep learning models that integrate thermal facial imaging with all other features, to assess whether image-based information further improves prediction.

Both single-output and multi-output models are compared, the latter capturing the interdependence of thermal comfort, TCPD, and stress.

Purpose and expected contribution

Through this dual strategy of Friedman test inference and predictive modelling, the research aims to achieve two objectives: (a) to provide empirical evidence of how thermal and task conditions jointly influence comfort, performance, and stress, and (b) to develop and evaluate computational models capable of anticipating human responses in multimodal, integrated ways. Together, these outcomes are expected to strengthen theoretical understanding while contributing practical tools for adaptive environmental and workload management.

Figure 3.1 presents an overview of this framework, showing the pathway from controlled thermal and task conditions, through multimodal data collection, to Friedman test inferential tests and predictive modelling.

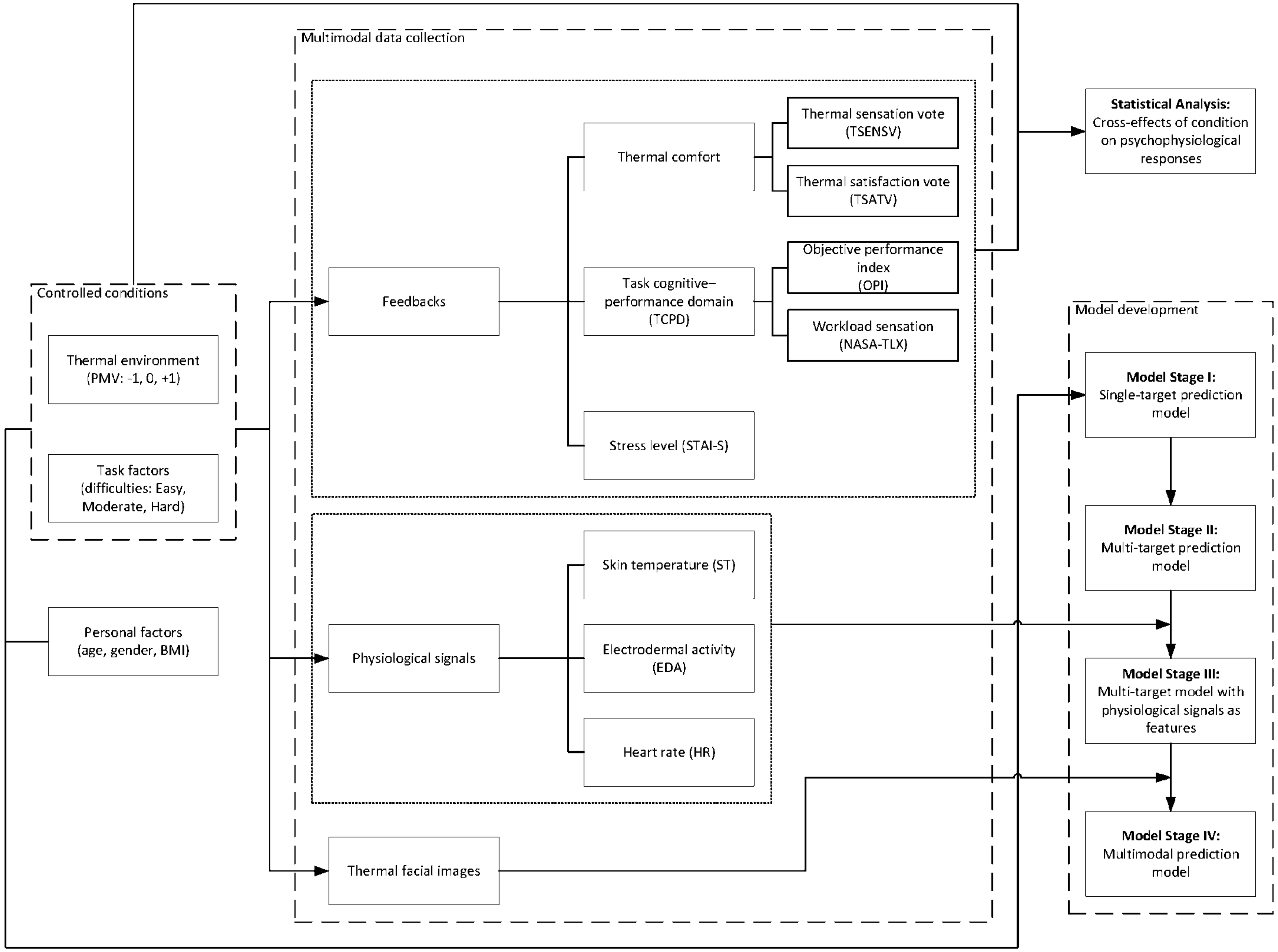


Figure 3.1 The framework of the study

3.2 Data collection

To address the research objectives, data were collected through a controlled laboratory experiment that enabled systematic manipulation of thermal conditions and task difficulty while recording subjects' responses across multiple modalities. The experiment was designed to produce a rich dataset combining subjective assessments, behavioural measures, physiological signals, and thermal imaging, thereby providing complementary perspectives on thermal comfort, TCPD, and stress. This section introduces the overall approach to data collection. Section 3.2.1 describes the experimental setup, including the laboratory environment and measurement instruments. Section 3.2.2 outlines the experimental design, specifying the independent and dependent variables as well as subject characteristics. Section 3.2.3 details the experimental procedure, presenting the chronological steps undertaken during each session.

3.2.1 Experiment setup

3.2.1.1 Laboratory setting

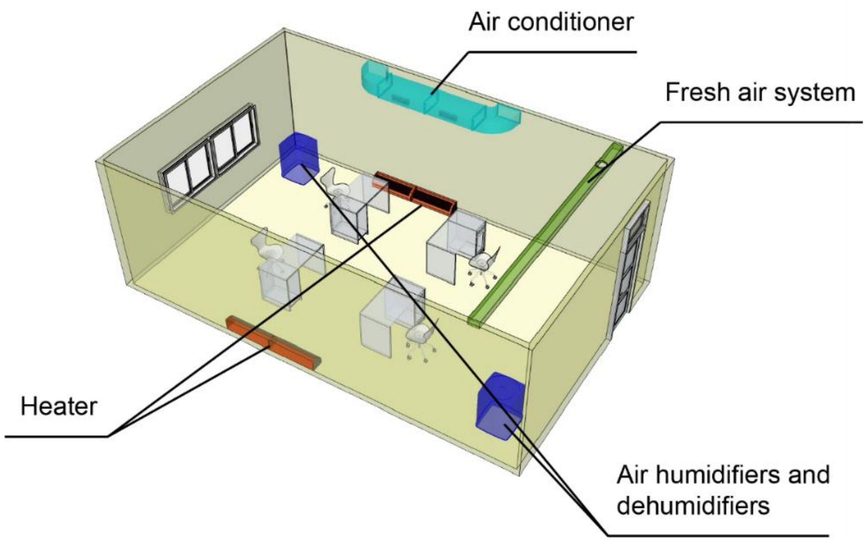
The data collection took place in October 2024 in Qinhuangdao, Hebei, using four identical office rooms located within the same building. Each room had the same floor area, ceiling height, and window orientation, ensuring a high level of comparability across experimental sessions. The walls, flooring, and ceiling materials were consistent, which helped to minimize differences in heat transfer and acoustic reflection.

A complete heating, ventilation, and air-conditioning (HVAC) system was installed in each office. The system integrated a fresh air ventilation unit, hot and cold air conditioners, electric heaters, humidifiers, and dehumidifiers, allowing fine-grained manipulation of thermal conditions and humidity within the room. All windows remained closed during the experiment to prevent outdoor influences, while the HVAC system maintained stable indoor conditions according to the experimental design.

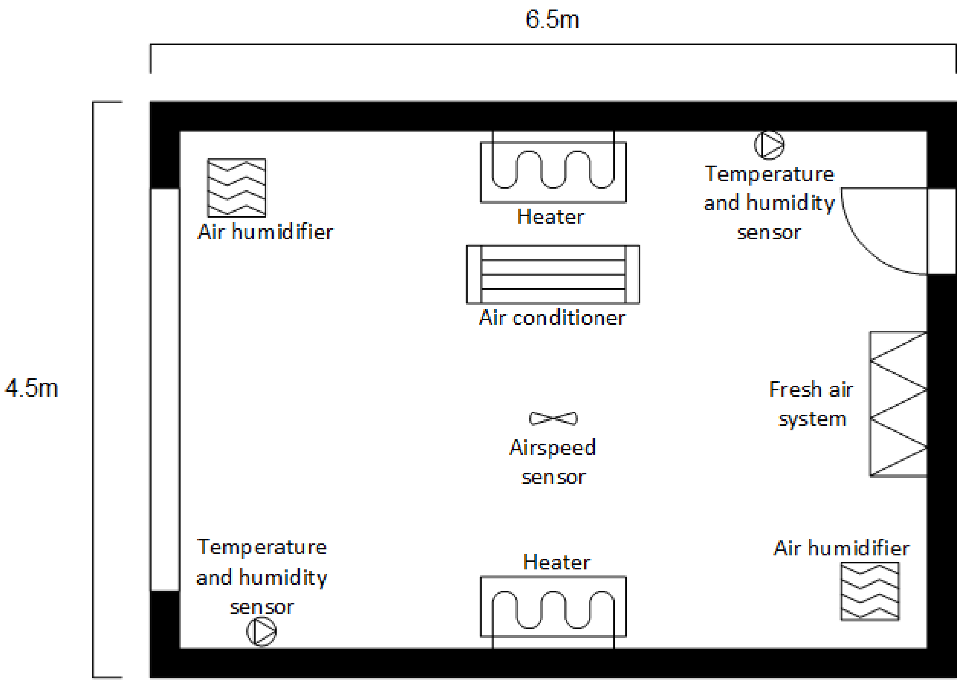
The office rooms were furnished in a standardised manner to minimise variability across experimental sessions. Four sets of desks and chairs were arranged in the centre of each room, positioned in pairs facing each other. On each workstation, a personal computer was placed together with an infrared camera, providing a consistent setup for task execution and physiological data collection. The standardisation of desk type, chair design, and monitor placement ensured that all subjects were exposed to near-identical physical conditions. This centralised layout also reduced obstructions to air circulation and provided fixed reference positions for the installation of the environmental sensors.

Noise and light conditions were carefully controlled. Artificial lighting was used exclusively, maintaining a constant illuminance level across sessions. External noise sources were minimized by the building's insulation, and no devices unrelated to the experiment were left operating during data collection.

Through this standardized laboratory configuration, the four offices provided equivalent environmental settings for all subjects, thereby ensuring reproducibility and comparability of the recorded data. A 3D model of the laboratory (Figure 3.2a) illustrates the overall appearance of the experimental setting, while a floor plan with equipment positions (Figure 3.2b) provides a detailed view of the spatial arrangement.



(a) 3D model of the experimental office



(b) Plan view of the equipments arrangement in the experimental office

Figure 3.2 Layout of the experimental office

3.2.1.2 Experiment equipment

(a) Environmental sensor

To monitor and record the indoor environmental conditions, two multifunctional sensors were installed in each laboratory room. These devices simultaneously measured and logged air temperature (°C), relative humidity (%RH), and carbon dioxide concentration (ppm), with both real-time display and internal storage.

In each room, the two sensors were fixed at head height relative to seated subjects. One sensor was mounted on the wall closer to the entrance door and the other on the opposite wall closer to the window. This arrangement allowed the detection of potential gradients in thermal or air quality parameters between the door and window zones, and ensured that the average conditions experienced by subjects were accurately captured. A schematic of the exact installation positions is provided in Figure 3.2b.

A photograph of the sensor model is shown in Figure 3.3, and its technical model together with main specifications (including measurement ranges, accuracy, and logging capabilities) are listed in Table 3.1. These features guaranteed that reliable reference environmental data were available throughout the experimental sessions.

Figure 3.3 Photograph of the environmental sensor used for monitoring indoor air temperature, relative humidity, and CO_2 concentration.

Table 3.1 Model and main specifications of the environmental sensor employed in the experiment

Model	**Parameters**	**Accuracy**	**Resolution**	**Sampling frequency**
QP NB-IoT	Temperature	±0.5°C (@ 0-50°C)	0.1°C	1 min
	Humidity	±5% RH (@ 10%-90%)	1%	1 min
	CO_2 concentration	±30ppm	1ppm	10 min

(b) Wearable device

For biometric data collection, smart wristband was be applied (shown in Figure 3.4). The device is lightweight, easy to wear, and enables real-time data transmission with access to raw signals, providing reliable monitoring of subjects in an office-like environment. Following physiological parameters can be collected by Empatica E4:

- Blood Volume Pulse (BVP): Photoplethysmography (PPG) technology is used to measure the volumetric variations of blood circulation by applying a photodetector at the skin surface [107]. The sampling frequency of BVP is 64Hz.

- Inter Beat Interval (IBI): By processing and remove the noise of BVP, the IBI signal can be obtained.

- Electrodermal Activity (EDA): The electrodermal activity (EDA) signal is an electrical manifestation of the sympathetic innervation of the sweat glands [108]. The sampling frequency of EDA is 4Hz.The resolution of EDA is from 1 digit to 900pSiemens

- Skin Temperature (ST): ST can be obtained through an infrared Thermople. The sampling frequency of ST is 4Hz. The resolution of ST is 0.02°C, and the accuracy is ±0.2°C.

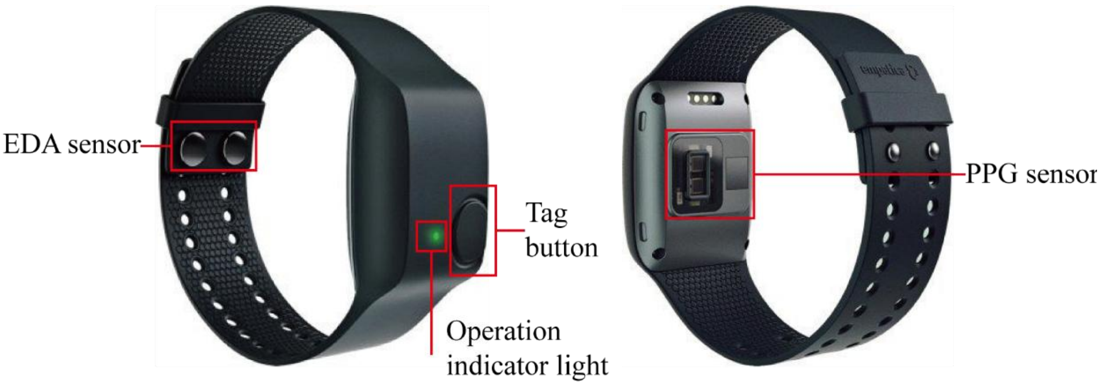


Figure 3.4 Figure 3.4 Photograph of the Empatica E4 wristband used for continuous collection of physiological signals, including blood volume pulse, inter-beat interval, electrodermal activity, and skin temperature.

(c) Imaging system

In addition to the environmental sensors and the wearable device, an imaging system was deployed to capture subjects' facial expressions and thermal emissions during the experimental sessions. BL-DT120 camera (see Figure 3.5) was mounted on each workstation at a fixed position relative to the computer monitor.

Figure 3.5 Photograph of the BL-DT120 imaging camera, which integrates visible light and thermal infrared modules for synchronised acquisition of facial and environmental data during the experiment.

The BL-DT120 integrates both RGB and infrared thermal imaging capabilities, enabling the simultaneous acquisition of visual and temperature-related features. This dual-channel configuration allowed for non-intrusive monitoring of facial responses, complementing the physiological data collected by the wearable wristband.

The detailed specifications of the BL-DT120, including resolution, spectral range, frame rate, and thermal sensitivity, are provided in Table 3.2. These parameters ensured that the imaging system could adequately support multimodal data analysis in subsequent stages of the study.

Table 3.2 Technical specifications of the BL-DT120 imaging camera used in the experiment

Parameter category	Parameter	Specification
Thermal imaging module	Sensor type	Uncooled focal plane array (FPA) microbolometer
	Resolution / frame rate	640×480@25fps
	Focal length	3.2mm
	Pixel size	12μm
	Spectral range	8 ~ 14μm
	Noise-equivalent temperature difference	≤50mK (@25°C, F#1.0)
	Temperature measurement range	-10°C ~ 550°C
	Temperature measurement accuracy	±2°C or ±2% of reading (Whichever is greater)
Visible light module	Sensor	1/2.7" CMOS sensor
	Resolution / frame rate	2592×1536@30fps
	Focal length	3.5mm
General specifications	Operating temperature	-20°C ~ 70°C
	Operating humidity	≤95%

3.2.2 Experiment design

After presenting the instrumentation in Section 3.2.1, this subsection introduces the overall experimental design. The purpose of the design was to create controlled laboratory conditions under which multimodal data, covering environmental parameters, physiological signals, and imaging information, could be systematically collected. Specifically, the design addressed four key aspects: the recruitment of human subjects, the control and monitoring of environmental conditions, the implementation of a task simulation system with different task types and difficulty levels, and the specification of dependent measures to be recorded. These components together ensured that the subsequent experimental procedure (Section 3.2.3) could be conducted in a standardized and reproducible manner.

3.2.2.1 Subjects

A total of 32 subjects (18 females, 14 males) were recruited from Qinhuangdao. The eligible age range for participation was 18–30 years, since substantial differences in thermal and cognitive responses across age groups may otherwise confound the experimental results. To ensure that subjects could meet the cognitive demands of the tasks, at least a secondary level of education was required for inclusion.

Statistical analysis in this study will be performed using the Friedman test, which is a non-parametric alternative to repeated-measures ANOVA and is suitable for within-subject designs. Because G*Power does not provide a direct option for the Friedman test, an a priori power analysis was conducted using its parametric counterpart (repeated-measures ANOVA within factors; Figure 3.6). Parameters were specified as follows: medium effect size ($f=0.25$), significance level ($\alpha=0.05$), statistical power ($1-\beta=0.80$), one group (all subjects experienced all conditions), and nine repeated measurements (3 thermal environments × 3 task difficulty levels). Assuming a correlation among repeated measures of 0.5 and a nonsphericity correction $\varepsilon=1$, the required sample size for the parametric test was 18 subjects. Considering the lower statistical efficiency of the Friedman test, approximately 15% more subjects were deemed necessary, resulting in an adjusted minimum of ~21 subjects. Our final sample size of 32 subjects therefore exceeds this requirement, ensuring sufficient power for the planned non-parametric analysis.

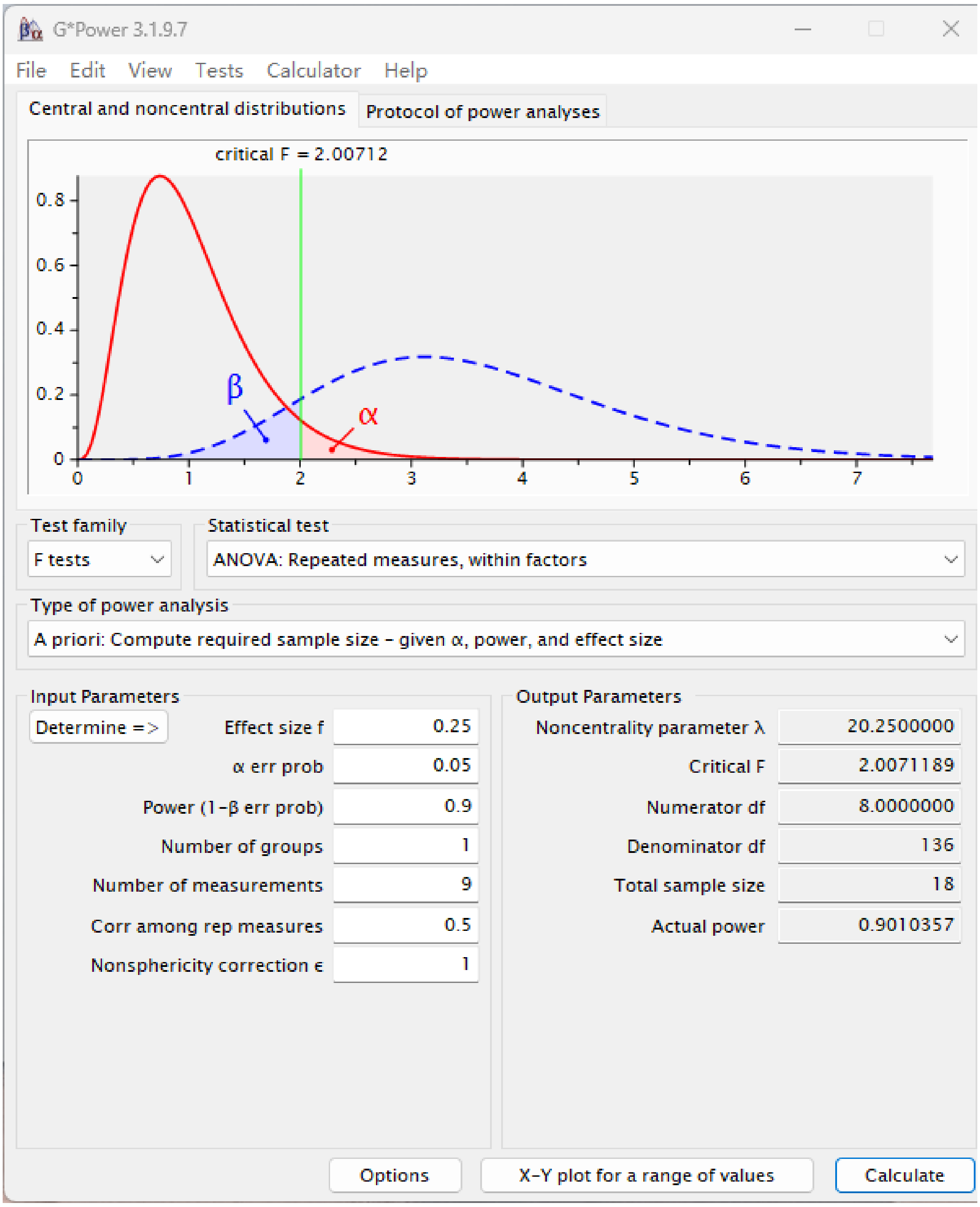


Figure 3.6 G*Power setup used for sample size estimation (repeated-measures ANOVA equivalent to Friedman test)

Exclusion criteria include subjects undergoing medical treatment (such as medication intake or physiotherapy), those who had been ill in the two weeks prior to the experiment, heavy drinkers (defined as consuming more than 500 ml of beer or the equivalent amount of alcohol daily), and heavy smokers (averaging more than five cigarettes per day). To safeguard the integrity of the measurements, all subjects were instructed to refrain from drinking, smoking, or taking any medication in the three days preceding the experiment, and to allow at least 1.5 hours after eating or exercising before participation.

During the experiment, clothing was standardised across all subjects. In contrast to allowing customised apparel for different temperature conditions, this experiment adopted a unified clothing requirement with an insulation value of 0.96 clo, corresponding to, for example, jack and long-sleeved T-shirt or shirt worn with long trousers [58]. The Garment insulation I_{clu} Table can be found at *Appendix A: Garment insulation I_{clu} Table*. This standardisation reduced variability due to clothing insulation between conditions and subjects.

All procedures complied with institutional ethical standards, and informed consent was obtained from each subject prior to participation.

3.2.2.2 Environmental control

The experimental sessions were performed in four climate-controlled rooms equipped with independent heating, ventilation, and air-conditioning (HVAC) systems. Relative humidity was maintained at approximately 75% and kept constant during all experiments. Air velocity was controlled to remain below 0.1 m/s, ensuring that subjects were not exposed to perceptible draughts. Carbon dioxide concentration was strictly maintained at levels lower than 550 ppm, thereby satisfying the requirements specified in ASHRAE Standard 55 [58].

Three of the rooms were used to create distinct thermal environments: Room I: 18 °C (slightly cool condition, PMV ≈ –1), Room II: 22 °C (neutral condition, PMV ≈ 0), and Room II: 26 °C (slightly warm condition, PMV ≈ +1). These conditions were selected to represent a range of typical indoor thermal states, covering slightly cool, thermally neutral, and slightly warm sensations according to the PMV scale [109], [110], [111], [112], [113]. Adopting three distinct temperatures allowed the experiment to examine how shifts in thermal comfort zones influence the physiological and behavioural responses of subjects when engaged in tasks of varying difficulty. The remaining room (Room IV) was also set to 22 °C and designated as a rest room where subjects could stay between trials to recover under neutral conditions. All experimental trials were carried out in the three temperature-controlled rooms.

To ensure accuracy and reproducibility, environmental parameters, including air temperature, relative humidity, air velocity, and CO_2 concentration, were continuously monitored throughout each session using calibrated sensors. The monitoring data were logged in real time, providing evidence that the target conditions were achieved and maintained during the experiments. These environmental controls minimised confounding factors and guaranteed that any physiological or behavioural responses measured in subjects could be attributed primarily to the experimental manipulations. A summary of the environmental conditions in the four experimental rooms is shown in Table 3.3.

Table 3.3. Ambient environment parameters maintained in the experimental and rest rooms

Parameters	**Room I**	**Room II**	**Room III**	**Room IV**
Air temperature (°C)	18.04 ± 0.13	21.95 ± 0.11	25.82 ± 0.21	22.01 ± 0.13
Relative air humidity (%)	74.18 ± 2.36	74.56 ± 3.12	75.63 ± 2.16	75.27 ± 4.37
Airspeed (m/s)	0.11 ± 0.04	0.09 ± 0.02	0.10 ± 0.03	0.12 ± 0.04
CO_2 concentration (ppm)	495 ± 9	498 ± 15	502 ± 13	541 ± 10

3.2.2.3 Task programme and difficulty control

To implement the experimental procedure, a dedicated computerised programme was developed that standardised the presentation of stimuli, instructions, timing, and data recording. Each experimental run lasted three minutes, during which subjects completed as many trials as possible under a global response-time limit that varied by difficulty level: 9 seconds per trial in the Easy condition, 7.2 seconds in the Moderate condition, and 6 seconds in the Hard condition. This timing constraint determined the approximate number of trials in each run, corresponding to about 20, 25, and 30 items respectively. If no response was registered within the allotted time window, the trial was automatically recorded as incorrect. Difficulty was thus operationalised through both the duration of the allowed response and the complexity of task parameters, enabling the systematic manipulation of cognitive workload.

The programme consisted of three main interface stages: the start screen, the task interface, and the end screen. All text and feedback were presented in English. The code of the programme can be found in *Appendix B: Code of the task simulation programme*.

Start screen (difficulty selection)

At the beginning of the experiment, subjects were presented with a start screen containing brief instructions and three buttons labelled *Easy*, *Moderate*, and *Hard* (as shown in Figure 3.7). By selecting one of these buttons, subjects initiated the task session at the corresponding difficulty level. Once a difficulty was chosen, the task interface automatically loaded, and the global countdown timer started.

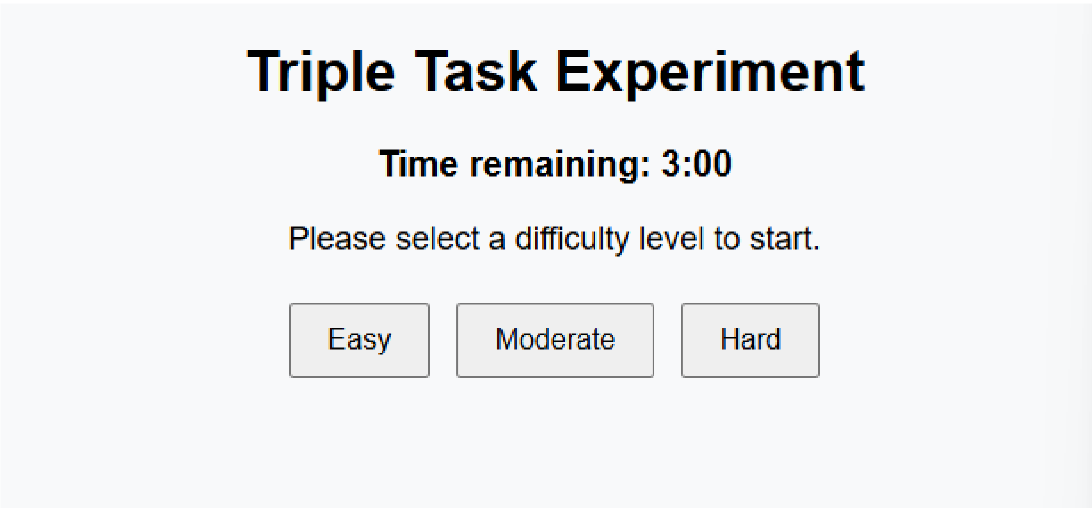


Figure 3.7 Start screen of the computerised task programme

Task interface

The main task window (shown as Figure 3.7) had the following components:

- Timer: displayed at the top of the screen, showing the remaining time of the 3-minute block.

- Stimulus display: located in the centre, presenting either a colour word (Stroop), two symbol sequences (pattern comparison), or an arithmetic equation.

- Task instructions: positioned directly below the stimulus, clarifying the required response (“Select the font colour”, “Are the patterns the same?”, or “Enter your answer”).

- Response area: contained clickable buttons (colour names or Same/Different) or an input field (for numerical answers), depending on the task.

- Feedback panel: appeared immediately after each response, displaying “Correct”, “Incorrect”, or “Timeout”.

- Statistics panel: shown on the left side, continuously reporting the number of completed tasks and correct responses.

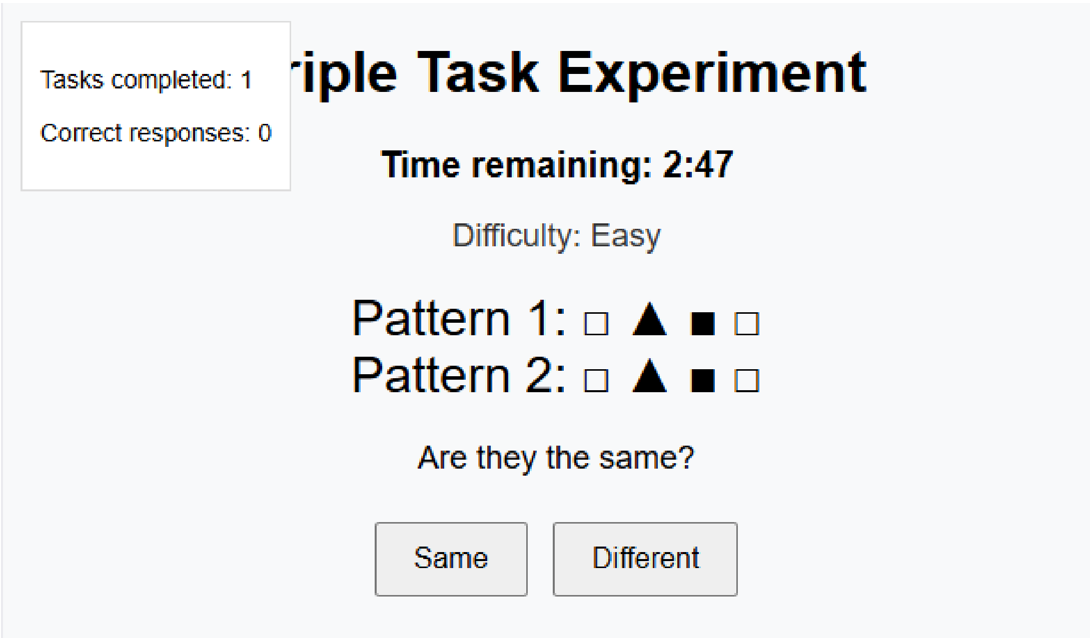
Task logic:

- Stroop task (Figure 3.8a): subjects identified the font colour of a colour word. Difficulty was manipulated by the ratio of congruent and incongruent stimuli.

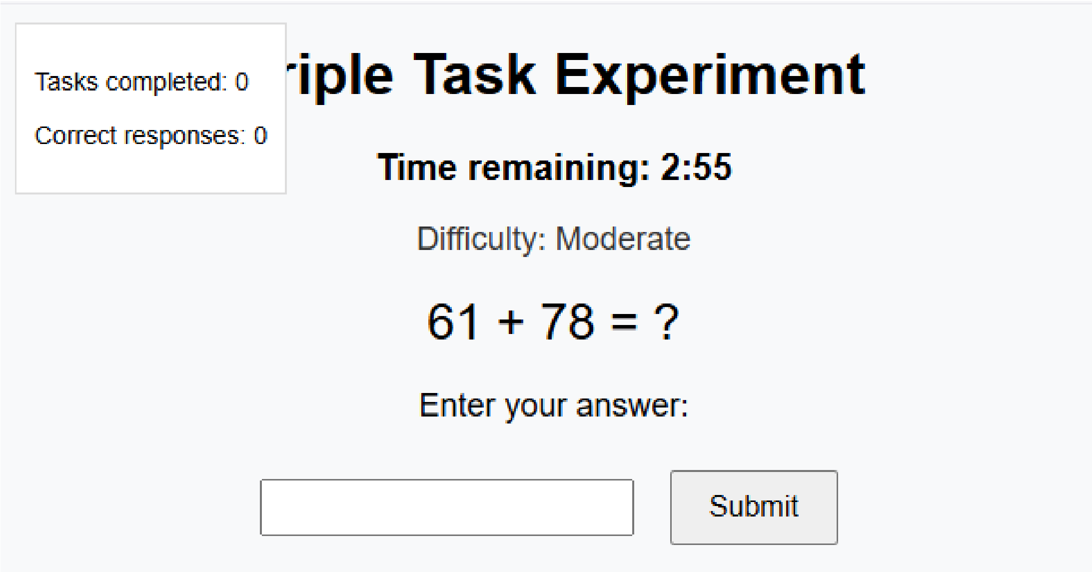
- Pattern comparison (Figure 3.8b): subjects compared two symbol sequences and judged whether they were identical. Difficulty increased by extending sequence length.

- Arithmetic task (Figure 3.8c): subjects solved addition or subtraction problems. Difficulty was determined by the numerical range of operands.

(a) Stroop colour–word task (selecting the font colour)



(b) Pattern comparison task (judging whether two symbol sequences are identical or different)



(c) Arithmetic task (solving a simple addition or subtraction problem)

Figure 3.8 Task interface of the computerised programme

End screen

At the conclusion of the 3-minute block, the programme automatically redirected to the end screen (shown as Figure 3.9). This summary view reported:

- Tasks completed
- Average response time (ms)
- Accuracy (%)

Below the results, a hyperlink to an online questionnaire (hosted on wjx.cn) was displayed, accompanied by the instruction to complete the survey and to press the wristband confirmation button after finishing. Two additional buttons were provided: Select difficulty (return to start screen) and Restart (repeat with the same difficulty).

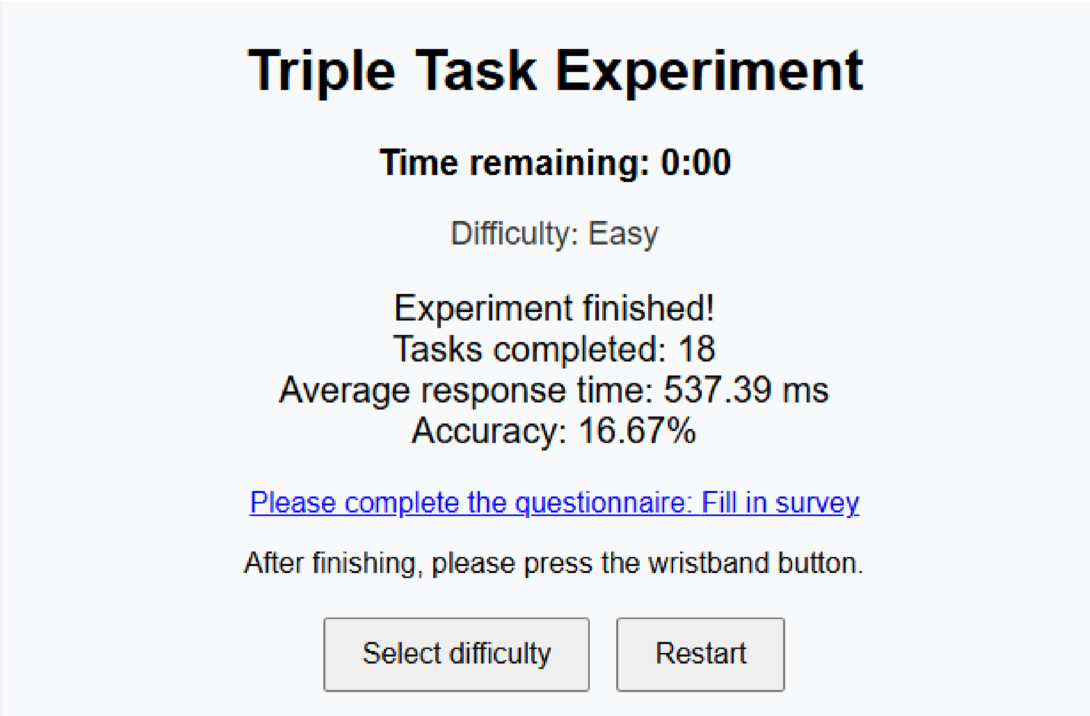


Figure 3.9 End screen of the programme, displaying overall performance statistics

The detailed configuration of task categories and their corresponding difficulty parameters is summarised in Table 3.4.

Table 3.4 Task categories and difficulty parameters

Task type	**Description**	**Easy**	**Moderate**	**Hard**
Global time	Time limitation of each question	9s	7.2s	6s
Stroop task	A colour word in English (Red, Blue, Green, Yellow) is displayed in a coloured font. Subject must select the font colour, not the word meaning.	Two options	Four options	Six options
Pattern comparison	Two symbol sequences (e.g., ○, ●, △, □) are	4-symbol sequence	6-symbol sequence	8-symbol sequence

	shown. Subject judges if they are identical; if different, exactly one symbol differs.			
Arithmetic	A simple addition or subtraction equation is shown. Subject submits the correct numerical answer.	Numbers 1–10	Numbers 1–99	Numbers up to 100

3.2.2.4 Dependent variables collection

To obtain a comprehensive dataset that captures both the internal states and the outward behavioural responses of subjects, data collection was designed to integrate subjective reports, task performance indices, physiological recordings, and multimodal imaging. The data served two main purposes: (i) to provide direct measures of subjects' perceived workload, affect, and thermal comfort, and (ii) to supply objective signals for quantifying physiological activation and behavioural expression under controlled environmental and task conditions.

Accordingly, the collection process was organised into three complementary domains. First, subjective scales and task performance were obtained during and after the experimental runs. All subjective feedback (thermal comfort, workload ratings, and state anxiety scales) as well as self-reported performance evaluations were administered via an electronic survey form provided by *WENJUANXING*. The unified electronic form ensured consistency of responses, reduced manual entry errors, and allowed immediate digital storage of all questionnaire data. A full copy of the questionnaire form is provided in *Appendix C Experiment Survey* for reference. Second, continuous physiological signals were acquired through the Empatica E4 wristband, ensuring high-resolution measurement of cardiovascular and electrodermal activity throughout the three-minute task blocks. Third, thermal video and image data were synchronously recorded using infrared and optical cameras, allowing subsequent analysis of both surface temperature distribution and subject behaviour.

Together, these three layers of data enable triangulation between perceived states, measured bodily reactions, and observable behaviours, providing a robust foundation for subsequent analysis of task and environmental effects.

I. Feedback

A total of 864 observations were obtained for each feedback variable (32 subjects × 3 thermal environments × 3 task difficulty levels × 3 repetitions). These data were generated through repeated sampling during the experimental sessions, which combined task performance blocks with self-report intervals. A detailed description of the experimental procedure, including session structure and timing, is provided in Section 3.2.3.

Thermal comfort assessment

Thermal comfort was measured using standardised single-item scales that have been widely adopted in environmental ergonomics. The rationale for selecting these instruments is threefold. First, the Thermal Sensation Vote (TSENV) and the Thermal Satisfaction Vote (TSATV) are part of the internationally recognised methodology for subjective thermal assessment, as defined in ISO 10551 [219]. Their use ensures comparability of the present findings with the established literature, such as Fanger's

PMV/PPD framework [172]. Second, they capture complementary dimensions of subjective thermal experience: perceived thermal state (descriptive) and subjective acceptability (evaluative). Finally, the single-item format minimises subject burden and remains sensitive enough to detect within-subject fluctuations during short experimental blocks, which has been demonstrated in laboratory thermal comfort studies [183]. After each three-minute simulation block, subjects were prompted to report their immediate thermal experience through two standardised single-item scales:

- Thermal Sensation Vote (TSENV): This item measures the perceived thermal state, anchored from −3 (“cold”) to +3 (“hot”), with 0 representing a thermally neutral sensation. The categories used in the study are summarised in Table 3.5.

Table 3.5 Thermal sensation vote (TSENV)

Value	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**
Label	Cold	Cool	Slightly cool	Neutral	Slightly warm	Warm	Hot

- Thermal Satisfaction Vote (TSATV): This item measures the degree of satisfaction with the thermal environment, anchored from −3 (“very dissatisfied”) to +3 (“very satisfied”), with 0 denoting a neutral judgment. The full scale is shown in Table 3.6.

Table 3.6 Thermal satisfaction vote (TSATV)

Value	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**
Label	Very dissatisfied	Dissatisfied	Slightly dissatisfied	Neutral	Slightly satisfied	Satisfied	Very satisfied

Administering these assessments after every run allowed the study to trace within-subject changes in both perceived thermal state and subjective acceptability across different environmental and cognitive conditions.

Task cognitive–performance domain (TCPD) assessment

The assessment of task cognitive–performance domain (TCPD) combined subjective workload rating and objective performance metrics in order to capture both perceived effort and actual operational outcomes. This two-layered approach was deliberately chosen. The NASA Task Load Index (NASA-TLX) is one of the most widely used subjective workload instruments, originally developed by Hart and Staveland (1988) [220], and later validated across a wide range of cognitive and operational domains (see Grier, 2015) [221]. Its multi-dimensional structure allows a nuanced breakdown of mental, physical, and temporal demands, as well as effort, frustration, and perceived performance. Coupling this with behavioural indicators, reaction time and response accuracy, which are established measures of task performance sensitivity to environmental or cognitive load [31], enables cross-validation of subjective perceptions against objective measures, thereby strengthening the interpretability and robustness of the results. The assessment of task cognitive–performance domain (TCPD) was divided into two complementary components: perceived workload and objective performance.

- Subjective workload: Workload was measured using the NASA Task Load Index (NASA-TLX), a widely adopted multi-dimensional tool for quantifying subjective workload across cognitive and operational tasks. Subjects completed the inventory after each experimental run, providing ratings on six sub-scales:

Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Each dimension was assessed on a 0–100 visual analogy scale, with higher scores reflecting greater perceived demand. A summary of the six evaluation dimensions is given in Table 3.7.

Table 3.7 NASA Task Load Index (NASA-TLX)

Mental Demand	How mentally demanding was the task?
Very Low	Very High
Physical Demand	How physically demanding was the task?
Very Low	Very High
Temporal Demand	How hurried or rushed was the pace of the task?
Very Low	Very High
Performance	How successful were you in accomplishing what you were asked to do?
Perfect	Failure
Effort	How hard did you have to work to accomplish your level of performance?
Very Low	Very High
Frustration	How insecure, discouraged, irritated, stressed and annoyed were you?
Very Low	Very High

- Performance: Objective task performance was derived from two simulation outputs, reaction time and accuracy rate, captured continuously during the three-minute blocks. These two indicators were later integrated into a composite performance index for analysis. The precise procedure for computing the composite index is described in Section 3.3 Data preparation.

Together, these two layers of TCPD assessment provide both subjective and behavioural measures, allowing cross-validation of workload perception against operational outcomes.

Subjective stress assessment

Stress level was operationalised as acute situational strain, reflecting short-term emotional responses to environmental and task demands. It was evaluated using the State–Trait Anxiety Inventory, State subscale (STAI-S). This instrument was selected because it is a validated tool for capturing momentary fluctuations in anxiety, originally developed and standardised by Spielberger and colleagues (1983) [222], and since widely applied in ergonomics, occupational health, and environmental stress research (e.g. Marteau & Bekker, 1992) [223]. Compared with more generic stress questionnaires, the STAI-S is highly sensitive to experimental manipulations and suitable for repeated administration within sessions [224], which makes it particularly appropriate for detecting changes induced by thermal exposure and cognitive workload. Stress in this study was operationalised as the acute psychological strain experienced during the experimental tasks. It was assessed at the beginning and the end of each experimental session to capture potential changes induced by prolonged exposure to thermal and cognitive challenges.

- State–Trait Anxiety Inventory (STAI-S): The state subscale of the STAI was administered to evaluate transient anxiety as a validated proxy of situational stress reactivity. This instrument captures momentary changes in affective state beyond general workload perception. The item prompts and response anchors are summarised in Table 3.8.

Table 3.8 State–Trait Anxiety Inventory (STAI-S)

Self-evaluation questionnaire	**1**	**2**	**3**	**4**
(1) I feel calm	□	□	□	□
(2) I feel secure	□	□	□	□
(3) I am tense	□	□	□	□
(4) I feel strained	□	□	□	□
(5) I feel at ease	□	□	□	□
(6) I feel upset	□	□	□	□
(7) I am presently worrying over possible misfortunes	□	□	□	□
(8) I feel satisfied	□	□	□	□
(9) I feel frightened	□	□	□	□
(10) I feel comfortable	□	□	□	□
(111) I feel self-confident	□	□	□	□
(12) I feel nervous	□	□	□	□
(13) I am jittery	□	□	□	□
(14) I feel indecisive	□	□	□	□
(15) I am relaxed	□	□	□	□
(16) I feel content	□	□	□	□
(17) I am worried	□	□	□	□
(18) I feel confused	□	□	□	□
(19) I feel steady	□	□	□	□
(20) I feel pleasant	□	□	□	□

1: Not at all, 2: Somewhat, 3: Moderately so, 4: Very much so

Administering the STAI-S allowed the study to quantify within-subject changes in situational stress and to relate these changes to thermal comfort, workload perceptions, and task performance across different environmental and cognitive conditions.

II. Physiological signals

Skin temperature (ST)

Skin temperature (ST) was chosen as a physiological index because it represents a direct manifestation of peripheral thermoregulation. Changes in skin vascular tone (vasodilation and vasoconstriction) are reflected rapidly in the skin surface temperature, making ST a sensitive marker of how the body responds to external thermal loads [96]. Unlike core temperature, which changes relatively slowly, ST reacts on a shorter timescale and is therefore particularly suitable for detecting acute fluctuations in thermal state under dynamic environmental conditions [96]. Furthermore, peripheral skin temperature has been shown to correlate with both subjective thermal sensation and stress-related vasomotor responses, allowing for integrative analysis alongside self-report data [44].

To record ST continuously, this study employed the Empatica E4 wristband, which measures skin temperature at the distal wrist. This wearable approach has several advantages: it is non-invasive and minimises disturbance during cognitive tasks, while still providing sufficient temporal resolution to capture dynamic physiological responses [44]. The E4 sampled skin temperature at 4 Hz, and data were subsequently down-sampled and averaged over each three-minute block, enabling alignment with subjective assessments and task performance metrics.

Electrodermal activity (EDA)

Electrodermal activity (EDA), also referred to as skin conductance, was included as a physiological measure because it provides a direct index of sympathetic nervous system activity. When stress or arousal increases, eccrine sweat gland activity is up-regulated, producing measurable changes in skin conductance [225]. As such, EDA has been widely validated as a sensitive marker of psychophysiological arousal and stress reactivity, complementing both subjective questionnaires (e.g. STAI-S) and performance-based measures [226]. Importantly, EDA responds on a short timescale (seconds), which makes it especially suitable for capturing transient fluctuations of stress levels during task performance under different environmental conditions [227].

EDA was recorded continuously using the Empatica E4 wristband, which measures skin conductance via two Ag/AgCl electrodes placed on the ventral wrist. Raw EDA signals were sampled at 4 Hz by the device and pre-processed by applying artifact removal and de-trending procedures. Compared with traditional laboratory systems, the E4 provides a wearable and minimally intrusive solution, allowing subjects to perform cognitive simulations without added constraint, while maintaining adequate signal quality for psychophysiological analysis [228].

Heart rate (HR)

Heart rate (HR) was incorporated as a physiological measure because it reflects the combined influences of the autonomic nervous system on cardiovascular regulation. Increases in cognitive demand or psychological stress are typically associated with higher sympathetic activation, resulting in elevated HR, while parasympathetic activity contributes to regulatory variability [130]. Thus, HR offers a sensitive index of both mental workload and stress reactivity, and has long been employed in ergonomics and psychophysiological research [163].

HR was continuously measured using the photoplethysmography (PPG) sensor of the Empatica E4 wristband. This sensor detects blood volume pulse at the wrist and derives inter-beat intervals (IBIs), from which HR is calculated. For the purposes of this study, HR was computed at a frequency of 1 Hz (one averaged value per second), providing sufficient temporal resolution to capture short-term fluctuations in cardiovascular activity while ensuring alignment with other physiological and subjective measures [44]. Compared with more obtrusive electrocardiogram systems, the E4 offers a non-invasive and minimally intrusive solution, allowing reliable monitoring during simulated task performance while preserving subject comfort [228].

III. Thermal video and images

Thermal imaging was included to provide an objective, contact-free measure of surface temperature distribution. Unlike single-point skin thermometry, infrared (IR) thermal video enables the capture of spatially resolved facial temperature patterns, which can reveal localised thermoregulatory responses such as vasodilation or cooling around the periorbital region [38]. Such measures have been increasingly employed in ergonomics and psychophysiology, as they permit the non-intrusive monitoring of emotional arousal, stress, and thermal comfort during task performance [44]. This method complements wearable biosensors by extending assessment to facial regions that cannot be covered by wrist-worn devices.

Thermal video and still images were recorded using the BL-DT120 infrared camera, which was installed at each workstation. The detailed specifications and installation configuration of the camera are provided in Section 3.2 Experiment setup. In this study, the imaging data were primarily used to extract regions of interest (ROIs) such as the forehead, nose, and cheeks, and to support multimodal analyses that aligned spatially resolved temperature signals with wearable physiological recordings (ST, EDA, HR).

This non-contact imaging strategy thus offered an additional perspective on facial thermoregulation, enhancing the robustness of physiological monitoring under combined thermal and cognitive loads.

3.2.3 Experiment procedure

The entire experiment was divided into four subsections, including one pre-experiment preparation subsection and three experimental subsections conducted in different rooms with distinct temperature conditions.

Pre-experiment preparation (Room IV)

At the beginning of each session, subjects entered the preparation room (Room 4) for a 10-minute pre-experiment preparation. They were assigned a unique identification number, their clothing was verified, and body height and weight were recorded. Subsequently, subjects completed a pre-experiment survey *(Appendix D: Pre-experiment Survey*) collecting background and physiological information. The experimental procedure was introduced in detail, and any questions were addressed. Finally, each pair of subjects drew lots to determine the sequence in which they would enter the three experimental rooms (Room I, Room II, Room III).

Experimental subsections (Rooms I–III)

Each experimental subsection lasted 105 minutes and followed the same procedure. After entering an experimental room, subjects were not allowed to leave until the subsection was completed. A 20-minute in-room preparation period was provided, allowing subjects' physiological state to stabilise under the given thermal condition. During this time, the Empatica E4 wristband was fitted on the left wrist for continuous skin temperature, electrodermal activity, and heart-rate measurements, and the thermal video system mounted at the workstation was activated for continuous facial imaging. At the same time, subjects drew lots to determine the order in which the three work tasks of different difficulty levels would be performed in that room.

Following acclimatisation, subjects then completed the work tasks in the predetermined order. Each task lasted 15 minutes, organised as three 3-minute performance blocks, each followed by a 2-minute reporting interval. At the end of each block, subjects completed the questionnaires listed in *Appendix C: Experiment survey*, covering thermal sensation, workload, and stress. Task performance, physiological signals, and thermal video were recorded continuously and synchronously during the full task period. Between successive tasks, a 20-minute in-room break was implemented to allow subjects' physiological state to return to baseline before commencing the next task.

Intermittent breaks (Room IV)

After completing the 105-minute subsection in one room, subjects returned to the preparation room (Room IV) for a break lasting at least one hour. The Empatica E4 devices were removed to upload physiological data, while thermal video recording ended upon leaving the experimental room. During this break, subjects could rest freely or briefly leave the laboratory. Subsequently, they proceeded to the next assigned room, where the same procedure was repeated.

This structured protocol ensured consistent recording of feedbacks, subjective reports, and thermal video data under all task sequences and environmental conditions, thereby enabling multimodal analyses of human responses to thermal environments and cognitive workload. The flows of the entire experiment, including pre-experiment preparation, in-room subsections, task cycles, and intermittent breaks, are illustrated in Figure 3.9.

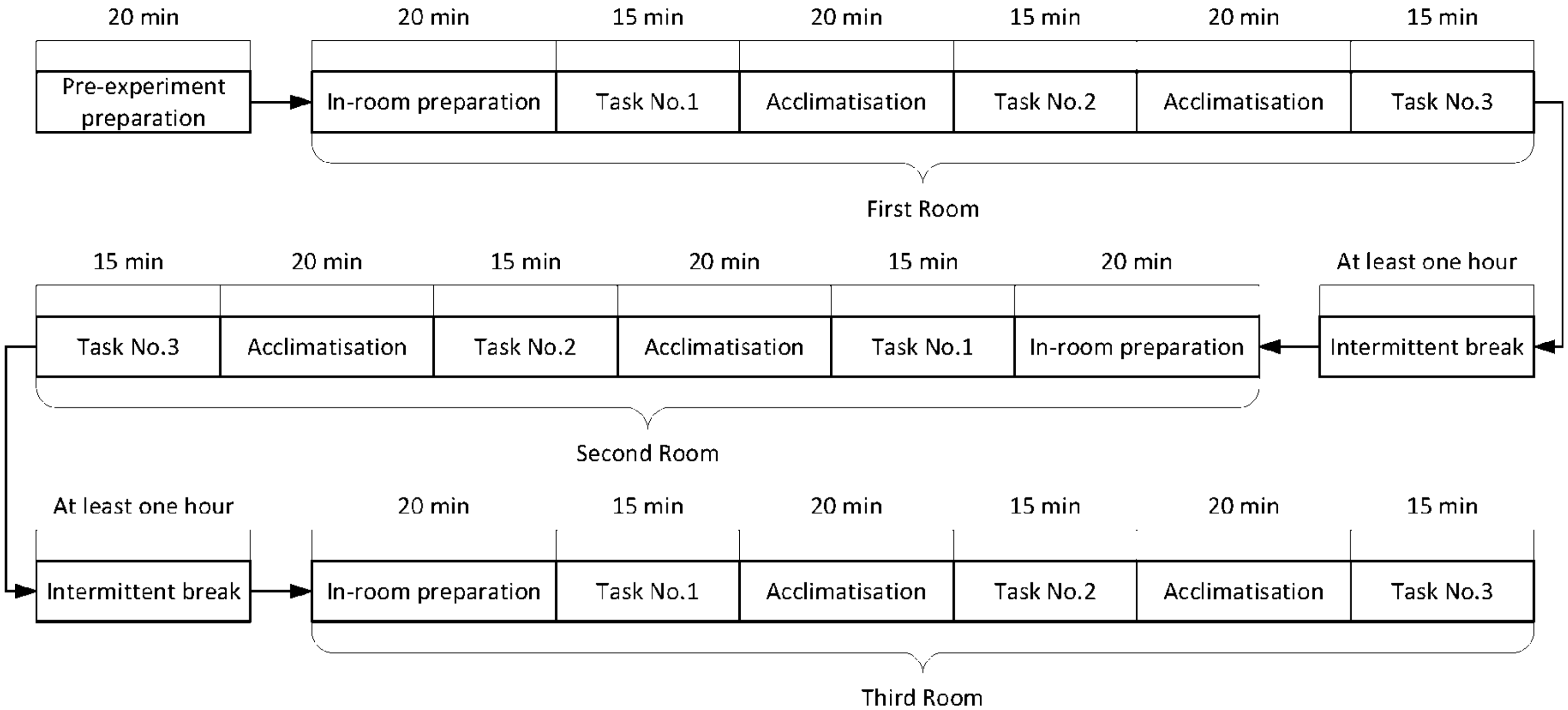


Figure 3.10 Experimental procedure timeline

3.3 Data preparation

The various data streams collected during the laboratory experiment required thorough cleaning, synchronisation, and transformation before subsequent analyses. Given the multimodal nature of the dataset, including survey-based personal factors, feedbacks, wearable sensor signals, and thermal imaging, data preparation was essential to ensure both data quality and comparability across conditions.

In line with the twofold aims of this study, the data were prepared in a manner that served both statistical validation and predictive modelling. For statistical validation, variables were retained in formats that respected their original scales to enable appropriate non-parametric tests. For predictive modelling, additional transformations such as feature derivation, numerical encoding, and normalisation were applied to allow integration of heterogeneous modalities within machine-learning pipelines.

To provide clarity, the following subsections describe data preparation by data type, starting from personal factors, together with the parameterisation of the two experimental factors (thermal environment and task difficulty), followed by feedback (thermal comfort, TCPD, and stress), then physiological signals (ST, EDA and HR), and finally thermal imaging features. Each subsection specifies the procedures applied and distinguishes between data forms intended for statistical analysis and those tailored for model training.

3.3.1 Personal factors

Personal information was collected from each subject through the pre-experiment questionnaire (*Appendix D: Pre-experiment Survey*), including gender, age, and body mass index (BMI). Basic demographic and anthropometric attributes were included to account for potential confounding effects and to enable exploration of inter-individual variability.

- Gender was reported as male or female.
- Age was self-reported in years.
- Body mass index (BMI) was calculated based on subjects' height and weight using the Eq. (3.1):

$$BMI = \frac{m_{weight}}{l_{height}^2}\left(\frac{kg}{m^2}\right) \qquad Eq. (3.1)$$

where m_{weight} represents the subject's weight (in kg), l_{height} represents the subject's height (in meter).

Processing for statistical analyses:

- Gender was retained as a binary categorical variable.
- Age was kept as a continuous variable (in years).
- BMI was retained as a continuous variable (kg/m^2).

Processing for modelling:

- Gender was numerically encoded (dummy variable: 0 = female, 1 = male).
- Age and BMI values were standardised using z-score normalisation across subjects to ensure comparability when included in the machine-learning models.
- These personal factors were primarily used as covariates or exploratory predictors to examine potential moderating effects of demographic and anthropometric characteristics.

3.3.2 Experimental factors (thermal environment & task difficulty)

Two controlled experimental factors were included in the study: thermal environment and task difficulty. Both were parameterised in ways that respected their theoretical order and allowed consistent use across statistical analysis and predictive modelling.

- Thermal environment: Three room conditions were designed to reflect equivalent shifts in thermal sensation according to the PMV scale: slightly cool (18 °C, PMV −1), neutral (22 °C, PMV 0), and slightly warm (26 °C, PMV +1). Given that these levels correspond to equal intervals on the thermal comfort scale, the variable was encoded directly using the values −1, 0, and +1.
- Task difficulty: The arithmetic tasks were presented at three difficulty levels: easy, moderate, and hard. These levels were conceptualised as ordinal with approximately equal steps in cognitive load. Accordingly, they were encoded numerically as 1, 2, and 3.

Processing for statistical analyses:

- Both variables were retained as categorical factors when applying non-parametric tests.
- For analyses involving trend effects, their ordinal encodings were used to capture potential ordered relationships.

Processing for modelling:

- Thermal environment was represented by its PMV-based ordinal values (−1, 0, +1).
- Task difficulty was represented by ordinal values (1, 2, 3).
- Both were further normalised if required by specific machine-learning algorithms.

This parameterisation preserved the theoretical meaning of the experimental conditions, reduced dimensionality compared to dummy encoding, and ensured their compatibility with both statistical evaluation and diverse machine-learning models.

3.3.3 Feedbacks

Three categories of feedbacks were considered: thermal comfort, the task cognitive–performance domain (TCPD), and stress. Each indicator was processed separately for statistical analyses and predictive modelling.

3.3.3.1 Thermal comfort (classification variable)

- Thermal sensation (TSENV): Reported on a 7-point ASHRAE-style scale ranging from −3 (cold), through 0 (neutral), to +3 (hot).

- Thermal satisfaction (TSATV): Reported on a bipolar 7-point scale ranging from −3 (very dissatisfied), through 0 (neutral), to +3 (very satisfied).

Processing for statistical analyses:

TSENV raw scores and TSATV values (after linear rescaling) were retained for hypothesis testing.

Processing for modelling:

TSENV was normalised as $(TSENV + 3)/6$, such that cold corresponds to 0, neutral corresponds to 0.5, and hot corresponds to 1.

TSATV was inverted and then normalised as $(TSATV + 3)/6$, resulting in values from 0 (very dissatisfied) to 1 (very satisfied).

3.3.3.2 Task cognitive–performance domain (TCPD) (regression variables)

- Objective performance index (OPI): Calculated from reaction time (RT, ms) and accuracy rate (AR, %) with equal weights (50%). To reduce the influence of extreme values, both RT and AR were normalised using the 10th and 90th percentiles of their distributions.

  For a given subject

$$RT_{norm,i} = \frac{P_{90}(RT) - RT_i}{P_{90}(RT) - P_{10}(RT)} \qquad Eq.(3.2)$$

$$AR_{norm,i} = \frac{AR_i - P_{10}(AR)}{P_{90}(AR) - P_{10}(AR)} \qquad Eq.(3.3)$$

where P_{10} and P_{90} denote the 10th and 90th percentiles across all subjects and trials.

Values below the 10th percentile were set to 0, and values above the 90th percentile to 100.

The composite index was then calculated as:

$$Performance_i = 0.5 \times RT_{norm,i} + 0.5 \times AR_{norm,i} \qquad Eq. (3.4)$$

Larger values represent better task performance.

- Subjective workload (NASA-TLX): Workload was evaluated using the NASA-TLX (0–120), with larger values indicating heavier perceived workload.

Processing for statistical analyses

- The raw OPI obtained from the task was directly used. Since the index had already been calculated to account for both reaction speed and accuracy, larger values indicated better performance.

- The NASA-TLX total score (0–120) was analysed in its raw form, with higher values reflecting heavier subjective workload.

- Thus, in statistical analyses, performance and workload were examined separately, and their original directional meaning was maintained (higher = better for performance; higher = worse for workload).

Processing for modelling:

- OPI was inverted as $(1 - Performance_i)$, so that 0 corresponded to the best performance and 1 to the worst.

- NASA-TLX was normalised as $\frac{NASA-TLX}{120}$, resulting in values from 0 (lowest workload) to 1 (highest workload).

3.3.3.3 Stress (regression variables)

- Stress level (STAI, state form): This questionnaire consists of 20 items, each scored from 1 to 4. Ten negative questions (question 3rd, 4th, 6th, 7th, 9th, 12th, 13th, 14th, 17th, 18th) were reverse-scored. The total scale ranges from 20 to 80, with larger values denoting lower levels of anxiety and stress.

Processing for statistical analyses:

- Recoded STAI totals were retained.

Processing for modelling:

- STAI scores were normalised as $(STAI - 20)/60$, rescaling the minimum of 20 to 0 and the maximum of 80 to 1.

3.3.3.4 Summary

In summary, subjects' feedbacks were assessed using subjective questionnaires and rating scales that capture different aspects of thermal and task experience. These included thermal sensation and comfort scales to assess perceived environmental conditions, as well as self-report measures of workload and arousal to characterise cognitive–emotional states.

Together, these measures complemented the objective physiological recordings and provided insights into subjective aspects of the human experience that cannot be inferred from physiological signals alone.

An overview of the feedback variables, their measurement instruments, and interpretation is provided in Table 3.8.

Table 3.8 Summary of feedback variables

Perspective	**Variable**	**For statistical analyses**	**For modelling**	**Value indicates**
Thermal comfort	Thermal sensation (TSENV)	Raw score, [-3, +3]	Normalised, [0, 1]	Larger values indicate feeling warmer
	Thermal satisfaction (TSATV)	Raw score, [-3, +3]	Normalised, [0, 1]	Larger values indicate greater thermal satisfaction
TCPD	Objective performance (OPI)	Raw composite index (Eq. 3.4), [0, 1]	Inverted composite, [0, 1]	Raw: larger values indicate better performance; Inverted: larger values indicate worse performance
	Subjective workload (NASA-TLX)	Raw score, [0, 120]	Normalised, [0, 1]	Larger values indicate heavier perceived workload
Stress	Stress level (STAI-S)	Raw score, [20, 80]	Normalised, [0, 1]	Larger values indicate lower situational anxiety

3.3.4 Physiological signals

Physiological signals were collected to obtain objective indicators of subjects' internal states that cannot be captured through self-report or performance measures alone. Skin temperature (ST) reflects peripheral vasomotor activity and thermoregulatory responses, providing information on thermal comfort and stress. Electrodermal activity (EDA), an established marker of sympathetic nervous system activation, is sensitive to changes in arousal and cognitive–emotional load. Heart rate (HR) reflects cardiovascular activation and stress responsiveness, integrating both thermal and cognitive influences.

These signals were selected because they are non-invasive, continuously recordable, and responsive to both environmental and task demands. After pre-processing and normalisation, they were used as continuous predictors to characterise short-term physiological states and to support the modelling of the overall human experience under experimental conditions.

After acquisition, all physiological signals were segmented into consecutive non-overlapping 30 seconds windows to ensure temporal alignment across modalities. Within each window, summary features were extracted after baseline signal correction and normalisation. Outlier detection was performed using the interquartile range (IQR) method, which was selected because it is non-parametric and does not assume normality of the underlying distribution, making it more robust for physiological signals that are often skewed, subject to slow drifts, or contaminated by sudden spike. The criterion was defined as

$$x \notin [Q_1 - 1.5 \times IQR, Q_3 + 1.5 \times IQR] \qquad Eq.(3.5)$$

where x represent an individual data point in the physiological signa, Q_1 and Q_3 represent the first (25th percentile) and third (75th percentile) quartiles of the distribution, respectively, and

$IQR = Q_3 - Q_1$ denotes the interquartile range. Values outside this interval were treated as outliers and replaced using linear interpolation to preserve temporal continuity of the signal.

3.3.4.1 Skin temperature (ST)

Skin temperature (ST) was included as it reflects peripheral vasomotor activity and is therefore a sensitive indicator of autonomic regulation under stress and fatigue. Prior studies have shown that decreases in ST are associated with sympathetic activation and less favourable physiological states, making it a relevant signal for assessing subjects' responses in this experiment.

Skin temperature was continuously collected with the Empatica E4 wristband at a sampling rate of 4 Hz. After the common preprocessing steps described above, the signal was further normalised to reduce inter-individual variability. Specifically, values were rescaled linearly to the $[0, 1]$ range using the subject-specific 10th and 90th percentiles:

$$ST_{norm,i} = \frac{ST_i - P_{10}(ST)}{P_{90}(ST) - P_{10}(ST)} \qquad Eq. (3.6)$$

where ST_i denotes the raw temperature at time i, $P_{10}(ST)$ and $P_{90}(ST)$ represent the 10th and 90th percentiles of the subject's distribution. In this way, higher values correspond directly to higher peripheral skin temperature.

From each 30-s window, four descriptive features were extracted: the mean skin temperature (ST_{mean}) as an indicator of the overall thermal level, the standard deviation (ST_{std}) to capture short-term fluctuations, the minimum (ST_{max}) and maximum (ST_{min}) values to represent extreme deviations, and the linear slope (ST_{slope}) obtained by least-squares regression to describe the local trend of change. Minimum and maximum values were included because sudden local drops or rises of skin temperature may reflect short-term vasoconstriction or vasodilation responses, which cannot be fully characterised by mean or variability alone. Together, these features constituted the representation of ST for subsequent modelling.

3.3.4.2 Electrodermal activity (EDA)

Electrodermal activity (EDA) was selected because it provides a direct and non-invasive read-out of sympathetic nervous system activity via eccrine sweat gland function. Unlike subjective ratings or behavioural measures, EDA is sensitive to both environmental and cognitive stressors, and has been extensively validated as a marker of arousal, mental workload, and emotional reactivity [82], [114]. This makes it appropriate for capturing changes in subjects' internal states under task demands that involve both thermal stress and cognitive load.

The EDA signal can be conceptually divided into two components: a phasic response (skin conductance responses, SCR), which reflects short-lived reactions to discrete stimuli, and a tonic component (skin conductance level, SCL), which reflects the slowly varying baseline of sympathetic arousal. Since the present experiment focused on continuous monitoring without discrete event markers, SCL was chosen as the more suitable index. SCL captures sustained changes in arousal driven by task difficulty and environmental conditions, whereas SCR analysis would require precise event alignment and stimulus-locked responses that were not the target of this design [229].

EDA was continuously recorded with the Empatica E4 wristband at a sampling rate of 4 Hz and decomposed into tonic and phasic components using the NeuroKit2 Python toolbox [230].

Only the tonic component, SCL, was retained for further processing. To reduce inter-individual variability, SCL was rescaled within each subject using percentile-based min–max scaling:

$$SCL_{norm,i} = \frac{SCL_i - P_{10}(SCL)}{P_{90}(SCL) - P_{10}(SCL)} \qquad Eq. (3.7)$$

where SCL_i denotes the raw temperature at time i, $P_{10}(SCL)$ and $P_{90}(SCL)$ represent the 10th and 90th percentiles of the subject's distribution.

From each 30-s analysis window, descriptive features were extracted from the normalised SCL, including the mean value (index of tonic arousal level, SCL_{mean}), the standard deviation (capturing variability in the tonic baseline, ST_{std}), and the linear slope (representing local temporal trends, ST_{slope}). Maximum and minimum values were not extracted because the tonic SCL component evolves very slowly, with limited range within 30 s windows, and these extrema are highly correlated with the mean level. Thus, including them would add little discriminative information beyond the chosen features.

3.3.4.3 Heart rate (HR)

Heart rate (HR) was selected because it provides a robust index of cardiovascular activation and is closely tied to both psychological and physiological demands. HR reflects the dynamic balance between sympathetic and parasympathetic nervous system activity, making it sensitive to stress, workload, and recovery processes [44]. In addition, HR is functionally related to metabolic rate and systemic oxygen demand [44, 157], which is particularly relevant in contexts where thermal and task-related challenges jointly influence subjects' physiological load. By capturing this integrated response, HR acts as a valuable proxy for overall physiological strain and energy expenditure.

HR was continuously recorded by the Empatica E4 wristband at a sampling rate of 1 Hz. This measurement represents an already processed signal derived from inter-beat intervals (IBIs) detected by the device's PPG sensor, ensuring reliable beat-to-beat monitoring without additional post-hoc detection algorithms.

After the general preprocessing steps, the HR signal was normalised for each subject using percentile-based min–max scaling to reduce inter-individual variability:

$$HR_{norm,i} = \frac{HR_i - P_{10}(HR)}{P_{90}(HR) - P_{10}(HR)} \qquad Eq. (3.8)$$

where HR_i denotes the raw temperature at time i, $P_{10}(HR)$ and $P_{90}(HR)$ represent the 10th and 90th percentiles of the subject's distribution.

From each 30-s analysis window, three descriptive features were extracted: the mean HR (overall cardiovascular activation, HR_{mean}), the standard deviation (short-term variability, HR_{std}), and the linear slope (temporal trend over the window, ST_{slope}). Maximum and minimum values were not included because the HR signal from the E4 is recorded at 1 Hz, yielding only 30 data points per window. In such short sequences, extrema are highly sensitive to artefacts or single noisy points and do not carry stable physiological meaning, unlike mean or variability-based metrics. These features provided compact and interpretable indicators of sustained cardiovascular responses under experimental conditions.

3.3.4.4 Summary

In summary, three physiological signals, skin temperature (ST), electrodermal activity in terms of skin conductance level (SCL), and heart rate (HR), were continuously recorded with the Empatica E4 wristband and processed into descriptive features on a 30-s basis. Each signal served as a complementary indicator of subjects' physiological states: ST reflected thermoregulatory and vasomotor activity, SCL provided a direct measure of tonic sympathetic arousal, and HR captured cardiovascular activation and metabolic demand.

Feature extraction was tailored to the properties of each signal. ST features included mean, variability, extrema, and slope, as rapid peripheral temperature deviations are informative of vasomotor changes. For SCL, only mean, variability, and slope were retained, as minima and maxima carry little additional information due to the tonic nature of the signal. For HR, mean, variability, and slope were extracted, while extrema were excluded because of the limited 1 Hz sampling rate and potential sensitivity to artefacts.

A consolidated overview of all extracted features, their data sources, and their physiological interpretation is provided in Table 3.9. This summary illustrates how the selected features jointly capture levels, variability, and temporal trends of the recorded physiological signals, supporting an integrated modelling of subjects' short-term internal states.

Table 3.9 Extracted physiological features

Signal	Rate	Feature	Value indicates
Skin temperature (ST)	4 Hz	Mean (ST_{mean})	Larger values indicate higher overall skin temperature

		Standard deviation (ST_{std})	Larger values indicate stronger short-term fluctuations in skin temperature
		Minimum (ST_{min})	Larger values indicate higher local minimum
		Maximum (ST_{max})	Larger values indicate higher peak skin temperature within the window
		Slope (ST_{slope})	Positive slope indicates upward warming trend; negative slope indicates cooling
Skin conductance level (SCL) of Electrodermal activity (EDA)	4 Hz	Mean (SCL_{mean})	Larger values indicate higher baseline sympathetic arousal
		Standard deviation (SCL_{std})	Larger values indicate more variability in tonic arousal within 30 s
		Slope (SCL_{slope})	Positive slope indicates increasing arousal; negative slope indicates decreasing arousal
Heart rate (HR)	1 Hz	Mean (HR_{mean})	Larger values indicate higher cardiovascular activation and metabolic demand
		Standard deviation (HR_{std})	Larger values indicate larger beat-to-beat variability within the 30 s window
		Slope (HR_{slope})	Positive slope indicates rising HR trend; negative slope indicates decreasing HR trend

3.3.5 Thermal images pre-processing

Thermal image recordings were obtained in parallel with physiological signals in the form of continuous video streams at 25 frames per second. To align the image data with the feedback windows while avoiding unnecessary redundancy, a temporal sampling procedure was implemented. Each 30-second window of the experiment contained approximately 750 frames; from these, a single representative image was extracted. Specifically, the frame corresponding to the mid-point of each 30-second interval was chosen as it minimised the influence of transient fluctuations that might occur at the beginning or end of the interval and provided a consistent rule across all subjects. The sampled images were subsequently stored as the unit of analysis and served as the input for further processing steps.

To focus the analysis on physiologically relevant regions, each sampled frame was further processed to extract a region of interest (ROI). Given the controlled experimental setup in which subjects were seated at a fixed distance and orientation relative to the thermal camera, the position of the face in the image plane remained stable across recordings. On this basis, a rectangular ROI was empirically defined to encompass the major facial areas, including the forehead, eyes, nose, and upper cheeks, which are known to provide reliable indicators of thermoregulatory responses. The same ROI parameters were consistently applied to all sampled images to ensure comparability between subjects and conditions. By removing background and peripheral elements, the cropped images reduced noise, lowered computational complexity, and retained the most informative thermal patterns for subsequent pre-processing and feature extraction.

Prior to model training, all sampled frames underwent a structured pre-processing pipeline designed to standardise the input format and enhance the robustness of subsequent feature extraction. To minimise the inclusion of background information, each image was first cropped to a consistent size of 100 × 100 pixels, thereby retaining only the most relevant facial region across subjects. The cropped images were then upscaled to 224 × 224 pixels, which corresponds to the fixed input size required by the convolutional neural network. This two-step operation both suppressed irrelevant information and ensured architectural compatibility. Following spatial standardisation, pixel intensity values, originally expressed as 8-bit greyscale levels in the range [0, 255], were normalised to [0, 1] using the ImageDataGenerator function implemented in Keras. The normalisation not only facilitated numerical stability during optimisation but also reduced the impact of absolute temperature differences between subjects or sessions. In addition, mild data augmentation procedures (e.g., small rotations, horizontal flips, and translations) were applied during training to increase sample variability while maintaining the integrity of the underlying thermal patterns. Through this pre-processing workflow, all images were transformed into a consistent, model-ready format that preserved essential thermal information and supported robust CNN feature learning.

After cropping and pre-processing, the thermal face images were standardised into a consistent, model-ready format. These processed images provided a reliable visual representation of subjects' dynamic thermal states and were temporally aligned with the feedback measurement windows. The resulting dataset was then used as the image input channel within the deep learning framework, the details of which are described in Section 3.5 Machine learning and deep learning model development. Enhance dataset diversity through randomidsed width offsets, height offsets, and hierarchical flipping.

3.4 Statistical analysis

The statistical analyses in this study were designed to evaluate how feedbacks varied under different thermal conditions. The focus was placed on the core response variables, thermal sensation (TSENV), thermal satisfaction (TSATV), the objective performance index (OPI), subjective workload (NASA-TLX), and state anxiety (STAI-S), which provided subjective and performance-related indicators of human thermal and cognitive states. The aim was twofold: first, to summarise the distributions of these measures using descriptive statistics; and second, to determine whether differences across experimental conditions were statistically significant. For each response variable, a total of 864 observations were obtained (96 for each of the nine environment–task combinations), providing adequate power for within-subject analyses. Because the study followed a within-subject design with repeated measurements, the assumptions of normality were first examined. Where these assumptions were not met, non-parametric procedures were applied, with the Friedman test serving as the principal method for detecting overall condition effects.

3.4.1 Descriptive statistics

Descriptive statistics were first calculated for subjects' personal factors, including age, gender, and body mass index (BMI). For continuous variables (age and BMI), means and standard deviations ($M \pm SD$) were reported to capture central tendency and variability, and medians and interquartile ranges (IQR) were additionally inspected to account for potential skewness. Gender distribution was summarised as frequency and percentage across the sample. These statistics provide an overview of the demographic and physiological background of the subjects and contextualise subsequent analyses.

Descriptive statistics were calculated separately for each feedback variable across the nine combinations of thermal environment (three levels) and task difficulty (three levels). For each condition, 32 subjects completed three repetitions, resulting in 96 observations per variable. The analysed variables included thermal sensation (TSENV) and thermal satisfaction (TSATV), which directly reflect subjective thermal comfort; the Objective Performance Index (OPI), derived from reaction time and accuracy, which quantifies task performance in an integrated manner; subjective workload as measured by the NASA-TLX, which captures the perceived cognitive and mental demands of the tasks; and stress level was measured by STAI-S, which serves as a psychological indicator of stress. For each of the nine condition groups, means and standard deviations ($M \pm SD$) were reported to capture central tendency and variability, while medians and interquartile ranges were additionally inspected to describe skewed or non-normal distributions. Presenting descriptive statistics in this way allowed us to profile how each response variable behaved across the experimental conditions and provided a foundation for subsequent comparisons using inferential statistics. All subsequent analyses were conducted on these groups of 96 observations, representing the basic analytical unit for each condition.

3.4.2 Normal distribution detection

The distributions of all feedback variables were examined to determine whether parametric analyses were appropriate. Checks were conducted for thermal sensation (TSENV), thermal satisfaction (TSATV), the objective performance index (OPI), subjective workload (NASA-TLX), and state anxiety (STAI-S) across each of the nine environment–task groups (three thermal environments × three task difficulties), with ninety-six observations per group.

Normality was primarily tested using the Shapiro–Wilk test ($\alpha = 0.05$), which is sensitive to deviations from Gaussian distributions in small to medium samples. For aggregated datasets the Kolmogorov–Smirnov test was also reviewed to confirm consistency. Distributional shape was further assessed with skewness and kurtosis coefficients, histograms, and Q–Q plots, because reliance on significance tests alone may inflate minor deviations in large repeated-measurement datasets.

When both test statistics and visual indicators pointed to clear departures from normality, data were treated as non-normal. These variables were analysed using the Friedman test for overall effects and Wilcoxon signed-rank tests with Bonferroni correction for pairwise comparisons. Where test results were borderline but descriptive indicators remained within conventional limits, the distributions were considered approximately normal; in such cases, repeated-measures ANOVA was inspected alongside non-parametric outcomes to verify robustness. Variables that satisfied all checks were analysed as normally distributed, though non-parametric results were also reported for completeness. This procedure ensured that inference was consistent and not overly dependent on the assumptions of any single method.

3.4.3 Friedman test

As each of the 32 subjects completed three repetitions per condition, the repeated measures within the same condition were first aggregated to obtain a single representative score for each subject under each experimental condition. This ensured that the unit of analysis corresponded to one observation per subject per condition. As the normality assumptions were not fulfilled for several of the feedback variables (the detailed results of the normality tests are presented in Chapter 4), subsequent analyses employed the Friedman test, a non-parametric alternative to repeated-measures ANOVA. The Friedman test is appropriate for within-subject designs with non-normal data and ordinal variables, such as subjective ratings of thermal comfort. Whereas ANOVA evaluates mean differences, the Friedman test assesses whether the ranks of repeated measurements differ systematically across conditions.

In the present study, the Friedman test was applied to compare each response variable across the nine experimental conditions defined by three thermal environments and three task difficulty levels. The hypotheses tested were formulated as follows:

- Null hypothesis (H_{0f}): there is no significant difference between different tasks.
- Alternative hypothesis (H_{af}): there is a significant difference between different tasks.

The Friedman statistic was calculated as a chi-square approximation based on rank sums of the repeated measures:

$$E(R) = \frac{N(k+1)}{2} \qquad Eq.(3.9)$$

$$Chi^2 = \frac{12}{N \cdot k \cdot (k+1)} \cdot \sum R^2 - 3 \cdot N \cdot (k+1) \qquad Eq.(3.10)$$

$$df = k - 1 \qquad Eq.(3.11)$$

where $E(R)$ is the expected rank sum value. N is the number of the subjects. k is the number of repeated conditions. The calculated Chi^2 values were evaluated against the Chi^2 value obtained from distribution tables according to the predefined significance level (p-value < 0.05) and the corresponding degrees of freedom. When the calculated Chi^2 exceeded the critical value, the null hypothesis (H_{0f}) was rejected, indicating that significant differences existed between conditions. Pairwise comparisons were then conducted using the Wilcoxon signed-rank test, with Bonferroni correction applied to adjust for multiple testing. All detailed statistical outcomes are reported in Chapter 4.

3.4.4 Post hoc test

When the Friedman test indicated statistically significant differences across conditions, pairwise comparisons were carried out to identify the specific sources of these effects. The Wilcoxon signed-rank test was employed as a non-parametric counterpart to the paired t-test, suitable for repeated-measures data without normality assumptions. All possible condition pairs were compared within each response variable.

All Wilcoxon signed-rank tests were conducted as two-tailed, because no strictly directional hypotheses were imposed; the analyses were therefore sensitive to differences in either direction. To account for inflated type I error due to multiple comparisons, the Bonferroni correction was applied by dividing the nominal α level (0.05) by the number of pairwise comparisons. Adjusted p-values below the corrected threshold were interpreted as evidence of significant differences between the paired conditions. Effect sizes were additionally reported using the rank-biserial correlation to complement significance testing.

The detailed outcomes of the post-hoc analyses, including adjusted p-values and effect sizes, are presented in Chapter 4, together with graphical summaries illustrating differences across environmental and task conditions.

3.5 Machine learning and deep-learning model development

The machine learning and deep learning analyses were designed as a progressive framework, gradually increasing in complexity from single-target classification tasks to multi-output and multimodal prediction. The overall data preparation and splitting strategy, which was consistently applied across all models, is first described in Section 3.5.1. After this, single-target classification models were trained to predict individual outcomes of TSENV, TSATV, OPI, NASA-TLX, STAI-S (Section 3.5.2). These tasks provided baseline performance and served as a benchmark for subsequent stages. The second stage (Sections 3.5.3–3.5.4) extended the models to handle multi-output learning, simultaneously predicting multiple response variables and incorporating physiological features such as skin temperature, electrodermal activity, and heart rate. Finally, a deep learning framework (Section 3.5.5) was implemented to combine structured numerical features with image features derived from facial thermal imaging, representing the most comprehensive multimodal approach. An overview of this modelling pathway, highlighting the progression from baseline classifiers to multimodal deep learning, is presented in Figure 3.11. To clarify the modelling sequence, Stage I comprised five independent classification models (Models I.1–I.5), each addressing a single target variable. For each model, the solid connections in Figure 3.11 represent the features used in training, while dashed connections indicate additional features that were tested to evaluate their contribution to predictive accuracy. The results of these baseline models provided reference points for subsequent extensions. These were then combined into a multi-output model (Stage II), before progressively incorporating physiological signals (Stage III) and facial thermal images (Stage IV), culminating in the most comprehensive multimodal framework.

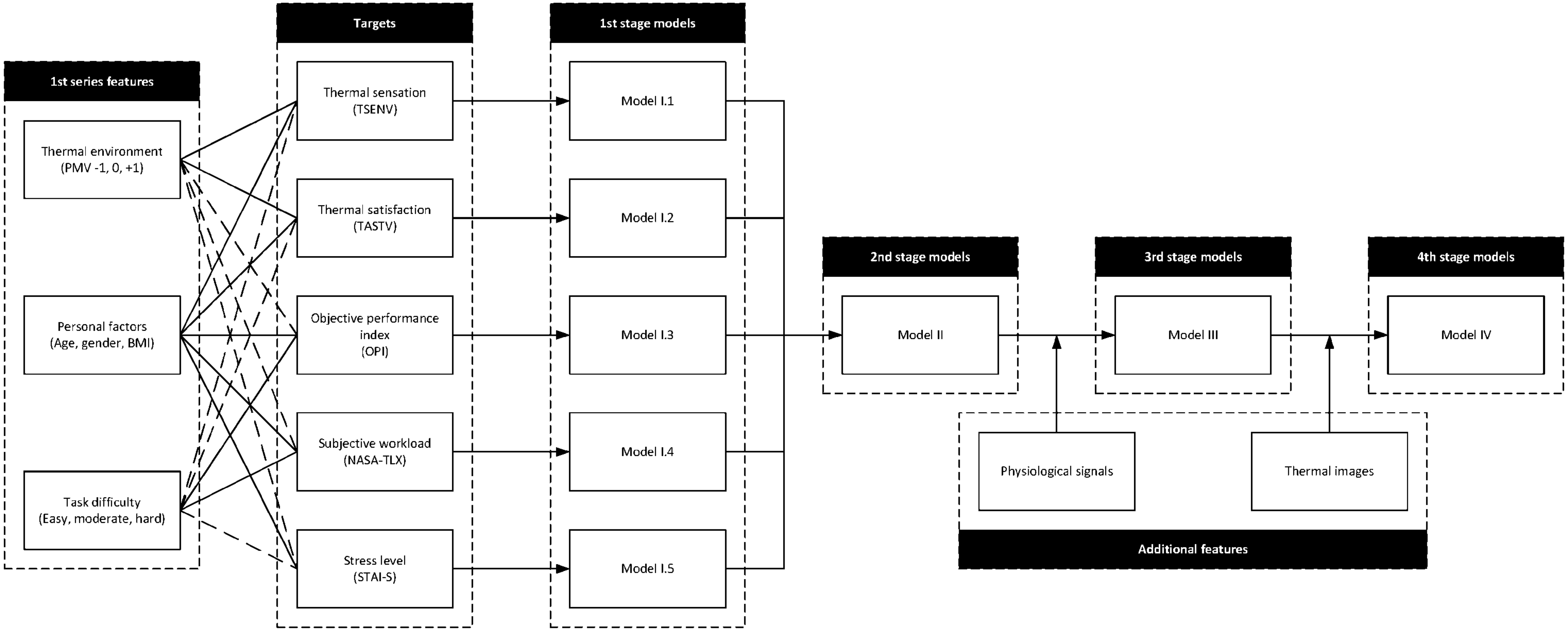


Figure 3.11. Illustration of the modelling pathway

Note: In Stage I (Models I.1–I.5), single-target classifiers were trained separately for thermal sensation, thermal satisfaction, work performance, workload sensation, and anxiety. Solid lines indicate feature sets used in training, while dashed lines indicate additional features tested for potential performance improvement. In Stage II (Model II), the five targets were combined into a multi-output framework to allow simultaneous prediction. Stage III (Model III) extended this framework by incorporating physiological signals (skin temperature, electrodermal activity, and heart rate). Finally, Stage IV (Model IV) further integrated thermal imaging with numerical features, resulting in a multimodal deep learning model.

3.5.1 Data preparation and splitting

All predictive analyses followed a consistent data partitioning strategy to ensure comparability of performance across single-target, multi-output, and multimodal frameworks. The dataset was divided into training, validation, and testing subsets. To reduce bias in model evaluation and avoid overfitting, a stratified K-fold cross-validation scheme was employed, maintaining balanced distributions of the target classes across folds. For each fold, approximately 80% of the data were used for model training, 10% for hyperparameter tuning, and 10% for final validation. Importantly, data from the same subject were kept within a single fold to prevent information leakage across training and testing sets. This splitting procedure was applied consistently across all stages of modelling, unless otherwise specified in later sections. The validation subsets within each fold were subsequently used for hyperparameter tuning, as detailed in later sections.

3.5.2 Single-target models (Model I.1 to Model I.5)

To establish baseline performance, four supervised learning algorithms were implemented for single-target prediction of thermal sensation, thermal satisfaction, work performance, workload sensation, and anxiety. The selected models were logistic regression (LR) [66], k-nearest neighbours (KNN) [66], random forest (RF) [66], and multilayer perceptron (MLP) [66]. These algorithms were chosen to cover complementary modelling paradigms, ranging from simple linear decision boundaries to flexible non-linear mappings:

- Logistic regression (LR)

  A parametric linear classifier that models outcome probabilities with interpretable coefficients [66]. It provides a straightforward benchmark and allows assessment of the relative contribution of input variables.

- k-nearest neighbours (KNN)

  A non-parametric, instance-based method that classifies samples based on their similarity to neighbouring training samples [66]. It is sensitive to local structure in the feature space and captures non-linear decision boundaries without explicit model parameters.

- Random forest (RF)

  An ensemble tree-based algorithm that aggregates predictions from multiple decision trees to improve robustness and reduce overfitting [66]. RF effectively handles heterogeneous features, evaluates feature importance, and generally performs well as a benchmark in classification tasks with moderate sample size.

- Multilayer perceptron (MLP)

  A feedforward neural network capable of modelling complex non-linear relationships between inputs and outputs [231]. MLP serves as a bridge between classical machine learning and deep learning approaches, demonstrating the potential benefit of increased model capacity while maintaining relatively simple architecture.

By applying these four algorithms, the analysis ensured coverage of diverse modelling approaches: interpretable linear models, memory-based non-linear models, ensemble learning, and neural networks. This diversity provided a balanced basis for benchmarking predictive performance across the five single-target outcomes.

As the prediction targets comprise both categorical outcomes (TSENV (Model I.1), TSATV (Model I.2), i.e., thermal comfort) and continuous outcomes (OPI (Model I.3), NASA-TLX (Model I.4), i.e. TCPD; and STAI-S (Model I.5), i.e., stress), algorithm variants for classification and regression were implemented accordingly. The same modelling families, LR, KNN, RF, MLP, were used, but their respective classification or regression formulations were applied depending on the outcome variable. This ensured methodological consistency across models while respecting the statistical nature of each target.

Hyperparameter optimisation was conducted through grid search within the cross-validation scheme described in Section 3.5.1. For each algorithm, candidate values were predefined (e.g., regularisation strength for LR; number of neighbours and distance metric for KNN; number and depth of trees for RF; hidden layer size, activation functions, and learning rate for MLP). The complete search ranges are summarised in Table 3.10, ensuring that all classifiers were tuned systematically under the same procedure. The evaluation of model performance followed a unified protocol, for classification models details of which are presented in Section 3.5.6.1 Single-target model validation: I. Classification model evaluation. For regression models, validation methods are provided in Section 3.5.6.1 Single-target model validation: II. Regression model evaluation

Table 3.10 Hyperparameter ranges explored for single-target models

Algorithm	**Main hyperparameters**	**Search range / options**
Logistic Regression (LR)	Regularisation strength (C), penalty type	C ∈ {0.01, 0.1, 1, 10, 100}, penalty ∈ {L1, L2}
K-Nearest Neighbours (KNN)	Number of neighbours (k), distance metric, weighting	k ∈ {3, 5, 7, 9, 11}, distance ∈ {Euclidean, Manhattan}, weights ∈ {uniform, distance}
Random Forest (RF)	Number of trees (n_estimators), Maximum depth, Minimum samples per split	n_estimators ∈ {100, 200, 500}, max_depth ∈ {None, 10, 20, 30}, min_samples_split ∈ {2, 5, 10}
Multilayer Perceptron (MLP)	Hidden layer sizes, activation function, learning rate, regularisation (α)	hidden_layer_sizes ∈ {(50,), (100,), (50,50)}, activation ∈ {ReLU, tanh}, leaning_rate_init ∈ {0.001, 0.01}, α ∈ {0.0001, 0.001, 0.01}

3.5.3 Multi-target model (Model II)

To jointly predict all behavioural and psychological outcomes within a unified framework, a multi-output model was constructed. Because the prediction targets were heterogeneous, two categorical variables (TSENV, TSATV) and three continuous variables (OPI, NASA-TLX, STAI-S), only the multilayer perceptron (MLP) was employed, as neural networks natively allow heterogeneous outputs to be combined in a single architecture. The network consisted of shared hidden layers for representation learning and a multi-branch output layer, with two softmax heads for classification and three linear heads for regression. This design ensured that input features were encoded into a common latent representation while still producing task-specific outputs.

A joint objective function was defined as Eq. (3.12).

$$\mathcal{L} = w_{cls1} \cdot BCE_1 + w_{cls2} \cdot BCE_2 + w_{reg1} \cdot MSE_1 + w_{reg2} \cdot MSE_1 + w_{reg3} \cdot MSE_3 Eq. (3.12)$$

where binary/categorical cross-entropy was applied to the classification tasks and mean squared error (MSE), or alternatively MAE or Huber loss, was applied to the regression tasks. Since manually fixing loss weights (w) risks favouring certain objectives, dynamic adjustment was introduced using GradNorm. This strategy balances the training process by adapting the weights according to gradient norms, thereby harmonising learning speeds across tasks and mitigating domination by a single loss.

Several implementation choices were explicitly specified to stabilise optimisation. For regression heads, each target was standardised using its training-set mean and variance prior to model fitting. Predictions were inverse-transformed during inference to recover the original physical scale. For classification heads, the network produced probability distributions over classes through softmax activations. Instead of adopting a universal threshold of 0.5 for decision boundaries, task-specific thresholds were optimised on the validation set to best align probabilistic outputs with categorical labels. In instances where training samples lacked annotations for one or more targets, a masking mechanism was incorporated into the loss function: the corresponding component was excluded from the joint objective for that sample, ensuring that only available labels contributed to gradient updates.

Through this formulation, the MLP was able to integrate heterogeneous classification and regression outcomes within one framework. The architecture leveraged representation sharing to reduce computational redundancy and promote cross-task generalisation, while GradNorm dynamically maintained balance among objectives. These design choices were not defaults of the modelling library but were purposefully implemented to suit the characteristics of the dataset, ensuring that the resulting multi-output model provided a principled and efficient foundation for subsequent benchmarking.

3.5.4 Integration of physiological features (Model III)

To investigate whether physiological signals could enhance predictive accuracy, the multi-output models were extended with additional features extracted from thermal imaging and wearable sensors. A major challenge was the mismatch in temporal resolution between the subjective questionnaires (administered every 3 min) and the physiological recordings (sampled every 30 s). To harmonise both datasets, each questionnaire response was replicated six times and aligned with the six corresponding 30-second physiological windows within the same 3-minute interval [31]. This ensured that each physiological measurement was associated with the relevant set of subjective outcomes.

The physiological features incorporated into the models included peripheral skin temperature, electrodermal activity, heart rate, and descriptors derived from thermal images [44]. All variables were normalised using the same preprocessing steps as in the baseline multi-output models to maintain comparability.

Given that the prediction task involved heterogeneous outcomes (two classification and three regression targets), only multilayer perceptrons (MLPs) were employed, as neural networks natively accommodate multiple output types within a single architecture. Following temporal alignment and feature integration, the extended datasets were used to train the multi-target MLPs described in Section 3.5.3. Since the inclusion of physiological signals increased both feature dimensionality and predictive complexity, hyperparameter optimisation was repeated for the extended models [66]. The tuning procedure adhered to the same cross-validation scheme as in Section 3.5.2, with adjustments to parameter ranges where appropriate, for example, larger hidden layers and a slightly broader learning-rate range were explored to accommodate the richer input representation.

To evaluate the benefit of physiological feature integration, the performance of the extended MLPs was benchmarked directly against the multi-target model (Model II) from Section 3.5.3. Both sets of models were trained under the same validation protocol, and their predictive performance was compared with respect to the five outcomes: thermal sensation (TSENV), thermal satisfaction (TSATV), work performance (OPI), workload (NASA-TLX), and anxiety (STAI-S). Detailed definitions of evaluation metrics are provided in Section 3.5.6.2 Multi-target model validation.

3.5.5 Integration of thermal images through MLP-CNN (Model IV)

To further enhance predictive capability, thermal images were incorporated into the modelling framework using a hybrid architecture that combines a multilayer perceptron (MLP) branch with a convolutional neural network (CNN) branch. CNNs are well-suited for automatically extracting spatial and thermal patterns from infrared image data [39], while MLPs have demonstrated effectiveness in modelling thermophysiological and categorical features [232]. Although such hybrid approaches have been applied in fields such as medicine and material science [186], their use in thermal comfort research is still limited.

The structure of a single residual block is illustrated in Figure 3.12. Each block contains two or three convolutional layers followed by batch normalisation and a non-linear activation function (ReLU). The identity mapping is then added element-wise to the block output. In cases where the input and output dimensions differ, a 1×1 convolution in the shortcut path is applied to project the input to the appropriate dimensionality. These residual operations enable deep feature representations to be learned effectively without compromising stability.

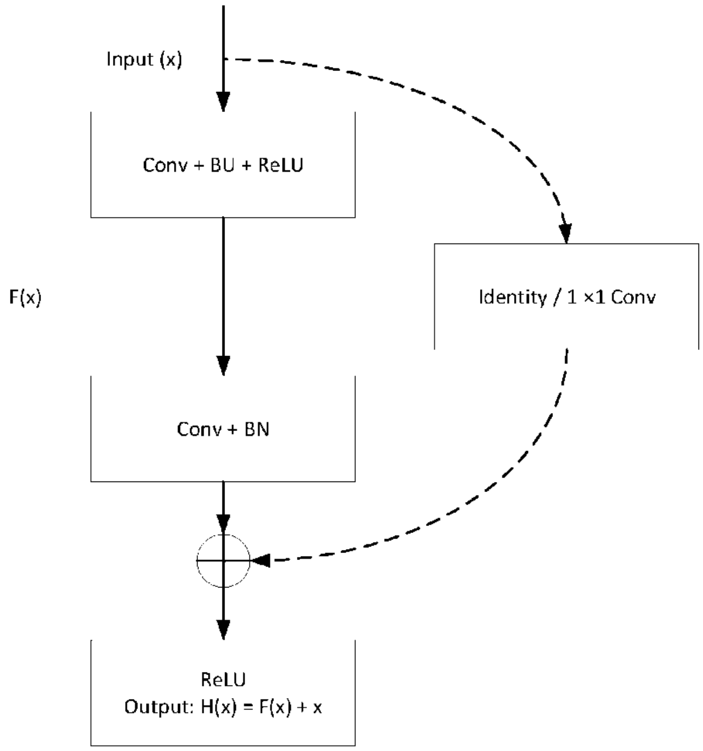


Figure 3.12 Residual block structure in ResNet50

The overall ResNet50 branch used in this study is summarised in Figure 3.13. The network was divided into six stages, corresponding to the canonical implementation of ResNet50:

- Stage 1 (Input convolution and pooling). A 7×7 convolution with 64 filters and a stride of 2 processes the input thermal images, followed by batch normalisation, ReLU activation, and a 3×3 max-pooling operation with stride 2.

- Stages 2–5 (Residual module groups). Four sequential groups of residual blocks progressively increase the feature depth:

- Stage 2: 3 residual blocks, each with three convolutional layers, expanding to 256 channels.
- Stage 3: 4 residual blocks, expanding to 512 channels.
- Stage 4: 6 residual blocks, expanding to 1024 channels.
- Stage 5: 3 residual blocks, expanding to 2048 channels.

  Each block applies 1×1, 3×3, and 1×1 convolutions, with shortcut connections maintained throughout.

- Stage 6 (Global average pooling). A global average pooling layer reduces each feature map to a single scalar value, producing compact, translation-invariant embeddings of the thermal image.

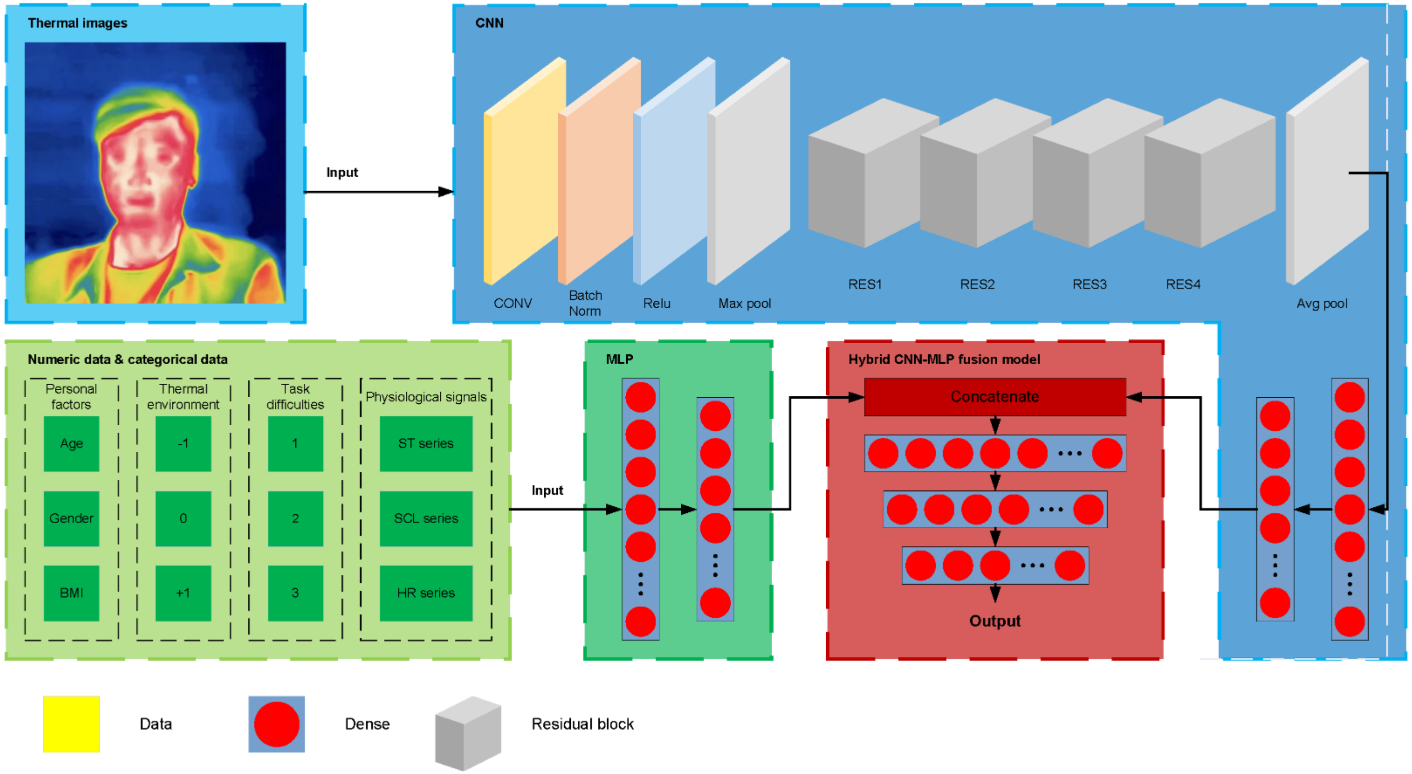


Figure 3.13 Structural overview of the ResNet50 branch

In parallel, the MLP branch processed numeric and categorical data, including personal factors (gender, age, BMI), physiological signals (skin temperature, electrodermal activity, heart rate), and task category indicators. The outputs of the ResNet50 and MLP branches were then concatenated, passed through three fully connected layers, and finally mapped to the prediction layer.

The modelling followed a progressive approach. Unidimensional models (one for each outcome) were first tested to validate the stability of the architecture and ensure successful feature integration. These implementations were not ends in themselves, but methodological steps to confirm feasibility and optimise the joint representation learning process. Building on this foundation, a multidimensional model was derived by expanding the final layer to five neurons, allowing simultaneous prediction of TSENV, TSATV, OPI, NASA-TLX, STAI-S. This multi-output design enabled the network to exploit correlations across outcomes, while still retaining target-specific information. The conceptual development process is illustrated in Figure 3.14.

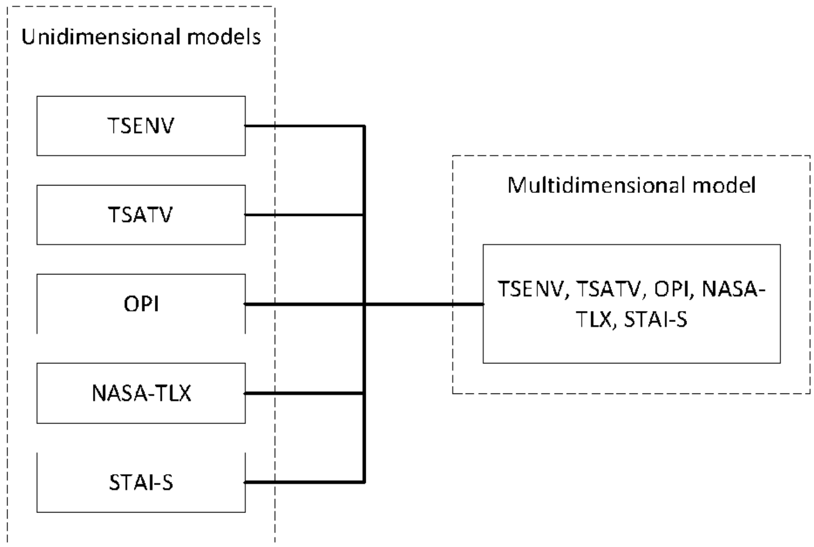


Figure 3.14 Conceptual development of the progressive modelling framework: from unidimensional model validation to the final multidimensional architecture

Hyperparameter optimisation was conducted through grid search following the cross-validation scheme described in Section 3.5.2. Candidate parameters included convolutional filter configurations, pooling strategies, dense layer sizes, dropout rates, and learning rates [233]. Optimal configurations were selected according to validation results. Details of the evaluation metrics are provided in Section 3.5.6.2.

3.5.6 Model validation

3.5.6.1 Single-target model validation

I. Classification model evaluation

To validate the performance of the developed thermal comfort models (i.e., Model I.1, Model I.2), model predictions were assessed using confusion matrices, classification metrics, and threshold-independent evaluation curves. The confusion matrix provides a tabular comparison between actual and predicted classes in terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). From these values, the following metrics were derived:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad Eq.(3.13)$$

$$Precision = \frac{TP}{TP + FP} \qquad Eq.(3.14)$$

$$Recall\ (True\ Positive\ Rate) = \frac{TP}{TP + FN} \qquad Eq.(3.15)$$

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad Eq.(3.16)$$

Accuracy explains the overall correctness of predictions, precision measures the reliability of positive predictions, recall reflects the sensitivity of the model in detecting a given class, and the F1 score provides a balanced combination of precision and recall.

In addition, two threshold-independent metrics were adopted to evaluate model robustness under class imbalance. The Area Under the Receiver Operating Characteristic curve (AUROC) measures the ability of the classifier to distinguish between positive and negative samples by plotting the True Positive Rate (TPR/Recall) against the False Positive Rate (FPR) across all thresholds:

$$False\ Positive\ Rate\ (FPR) = \frac{FP}{FP + TN} \qquad Eq.(3.17)$$

$$AUROC = \int_0^1 TPR(FPR)d(FPR) \qquad Eq.(3.18)$$

The Area Under the Precision-Recall Curve (AUPRC) was also computed. This metric plots Precision against Recall for different thresholds and is particularly informative in imbalanced datasets, as it prioritises the model performance on minority classes. The AUPRC equals 1.0 for a perfect classifier and approaches 0.0 for a random or entirely incorrect classifier:

$$AUPRC = \int_0^1 Precision(Recall)d(Recall) \qquad Eq.(3.19)$$

Together, Accuracy, Precision, Recall, F1 score, AUROC, and AUPRC provide a comprehensive and balanced validation framework to compare the predictive ability of the four unidimensional thermal comfort models.

II. Regression model validation

To validate the predictive accuracy of the regression-based TCPD and stress models (i.e., Model I.3–Model I.5), model predictions were assessed using error-based metrics and diagnostic plots. From the prediction–observation pairs, the following indices were derived:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad Eq.(3.20)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \qquad Eq.(3.21)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad Eq.(3.22)$$

where y_i denotes the observed (ground-truth) value of the response variable for sample i, $\widehat{y_i}$ is the corresponding model prediction, and $\bar{y}$ represents the mean of the observed values across the dataset. RMSE explains the overall magnitude of prediction errors and penalises larger deviations more strongly. MAE reflects the average absolute deviation and thus provides an intuitive measure of typical discrepancy between predicted and observed values. The coefficient of determination R^2 quantifies the proportion of variance in the observed outcomes that is explained by the model predictions.

In addition to numerical indices, two graphical diagnostics were employed. Residual histograms were plotted to examine the error distribution and detect skewness or heteroscedasticity, with a symmetric distribution around zero indicating unbiased predictions. Scatter plots of predicted versus actual values were further generated to assess alignment with the identity line $(y = \hat{y})$ and to reveal potential systematic biases.

Together, $RMSE$, MAE, R^2, and graphical diagnostics provide a consistent and balanced validation framework for regression-based thermal comfort models.

3.5.6.2 Multi-target model validation

To validate the performance of the multi-target models (Model II, Model III and Model IV), which integrates image features with tabular data to jointly predict two classification and three regression targets, a combined evaluation framework was applied. Each target was assessed individually according to its output type while following the same procedure adopted in single-task models.

For the classification outputs, confusion matrices were first constructed to characterise the prediction distribution in terms of TP, TN, FP, and FN. From these, accuracy, precision, recall, and F1 were derived to capture overall correctness, reliability of positive predictions, sensitivity, and the balance between precision and recall. To further ensure robustness under class imbalance, AUROC and AUPRC were computed as threshold-independent metrics, providing insight into the discriminative power of the classifiers across varying thresholds.

For the regression outputs, performance was quantified using RMSE, MAE and R^2. These indices capture the magnitude of large errors, the average absolute deviation, and the proportion of variance explained, respectively. To complement numerical indices, residual histograms were examined to assess error distributions, nd predicted-versus-actual scatter plots were generated to evaluate agreement with the 1:1 line and identify systematic deviations.

By combining classification-based and regression-based validation for each output dimension, the adopted framework provides a rigorous and balanced assessment of the multi-target MLP models, ensuring that predictive performance is consistently evaluated across heterogeneous outputs.

For the multi-target CNN-MLP model (Model IV), which integrates image features (processed by CNN) with tabular features (processed by MLP), model outputs were validated per target using the same evaluation strategy described above: classification-type targets were assessed with confusion-matrix-based indices and threshold-independent curves, while regression-type targets were evaluated with error-based measures and diagnostic plots.

Chapter 4 Results and Findings

4.1 Introduction

This chapter presents the findings and results obtained from the analyses described in Chapter 3. The aim is to provide a clear and systematic account of the statistical properties of the collected data and the performance of the developed modelling approaches. First, descriptive statistics are reported to characterise the dataset in terms of central tendency and variability. Subsequently, tests of normality are conducted to examine distributional assumptions, followed by non-parametric inferential analyses. In particular, the Friedman test is applied to identify overall differences across conditions, and post-hoc analyses are performed where appropriate to determine pairwise contrasts. Building on these statistical examinations, the chapter then reports the predictive performance of both unidimensional and multidimensional models. For classification outputs, model performance is presented using accuracy, precision, recall, F1 score, AUROC, and AUPRC, complemented by confusion matrices to enable detailed class-wise evaluation. For regression outputs, performance is quantified using RMSE, MAE, and R^2, together with residual histograms and predicted-versus-observed scatter plots for diagnostic assessment. This sequential presentation, from descriptive statistics to advanced modelling results, ensures that the empirical findings are transparent, reproducible, and closely aligned with the methodological framework established in Chapter 3.

4.2 Statistical results

4.2.1 Descriptive statistics

Descriptive statistics of the subjects' personal factors are summarised in Table 4.1. The analysed factors included age, gender, and body mass index (BMI). For age and BMI, means and standard deviations ($M \pm SD$) were reported together with medians, maximum value and minimum value to capture both central tendency and potential skewness, while gender distribution was reported as frequency and percentage.

For each of the nine experimental conditions (3 thermal environments × 3 task difficulty levels), means and standard deviations ($M \pm SD$) were reported to capture central tendency and variability, while medians (Md) and IQR were additionally inspected to describe skewed or non-normal distributions. The descriptive results for TSENV, TSATV, OPI, NASA-TLX, and STAI-S provide an overview of the dataset, highlight potential distributional characteristics, and serve as a preparatory step for subsequent normality testing and inferential analyses. The summary statistics are presented in Table 4.2, while Figure 4.1 illustrates the corresponding violin plots (for TSENV and TSATV) and boxplots (for OPI, NASA-TLX and STAI-S), enabling a visual comparison of distributional patterns and variability across the nine experimental conditions

Table 4.1 Personal factors of subjects

Gender	n (%)	Age			BMI (kg/m^2)		
		M ± SD	Median	[min, max]	M ± SD	Median	[min, max]
Female	18 (56.25%)	20.00 ± 1.33	20	[18, 24]	21.50 ± 2.09	21.05	[18.07, 25.54]
Male	14 (43.75%)	19.19 ± 2.52	19	[18, 28]	23.21 ± 4.08	22.22	[19.03, 32.66]
Total	100 (100%)	19.90 ± 1.92	19.5	[18, 28]	22.25 ± 3.20	21.59	[18.07, 32.66]

Table 4.2 Descriptive statistics (Mean ± SD and Median [IQR]) of outcome variables across nine conditions

(a) Descriptive statistics for thermal comfort perspective

Thermal environment	Task difficulties	N	TSENV		TSATV	
			M ± SD	Md [IQR]	M ± SD	Md [IQR]
Slightly warm	Easy	96	1.72 ± 0.83	2.0 [1.00, 2.00]	0.16 ± 1.33	0.0 [-1.00, 1.00]
	Moderate	96	1.11 ± 0.56	1.0 [1.00, 1.00]	0.76 ± 1.24	1.0 [0.00, 2.00]
	Hard	96	0.73 ± 0.55	1.0 [0.00, 1.00]	1.18 ± 1.01	1.0 [0.75, 2.00]
Neutral	Easy	96	0.11 ± 0.56	0.0 [0.00, 0.00]	1.93 ± 0.94	2.0 [1.00, 3.00]
	Moderate	96	-0.07 ± 0.53	0.0 [0.00, 0.00]	1.78 ± 0.93	2.0 [1.00, 3.00]
	Hard	96	3.91 ± 0.60	0.0 [0.00, 0.00]	1.55 ± 1.00	1.0 [1.00, 2.00]
Slightly cool	Easy	96	-1.60 ± 0.91	-1.0 [-2.00, -1.00]	-0.14 ± 1.14	0.0 [-1.00, 0.00]
	Moderate	96	-1.03 ± 0.47	-1.0 [-1.00, -1.00]	0.52 ± 1.08	1.0 [0.00, 1.00]
	Hard	96	-0.59 ± 0.64	-1.0 [-1.00, 0.00]	1.02 ± 0.93	1.0 [1.00, 1.00]

(b) Descriptive statistics for task cognitive-performance domain (TCPD) perspective

Task difficulties	Thermal environment	N	OPI		NASA-TLX	
			M ± SD	Md [IQR]	M ± SD	Md [IQR]

	Slightly warm	96	55.56 ± 7.72	56.0 [51.75, 61.00]	24.36 ± 14.73	22.5 [12.00, 32.00]
Easy	Neutral	96	58.35 ± 10.62	59.0 [50.75, 67.25]	21.05 ± 14.24	19.50 [12.00, 28.00]
	Slightly cool	96	55.58 ± 7.36	56.0 [51.00, 60.25]	23.97 ± 18.83	21.0 [9.00, 32.50]
	Slightly warm	96	50.41 ± 8.00	51.0 [45.75, 56.00]	33.76 ± 18.26	29.0 [20.00, 41.00]
Moderate	Neutral	96	54.24 ± 8.25	54.5 [50.75, 59.00]	30.39 ± 14.26	27.0 [23.00, 40.25]
	Slightly cool	96	50.55 ± 7.36	52.0 [47.00, 56.00]	33.46 ± 19.54	30.5 [18.75, 45.25]
	Slightly warm	96	40.27 ± 7.37	41.0 [35.75, 45.00]	45.58 ± 25.77	37.0 [27.00, 69.50]
Hard	Neutral	96	43.76 ± 6.97	44.0 [40.75, 48.00]	38.97 ± 19.10	37.0 [25.00, 49.50]
	Slightly cool	96	40.54 ± 8.26	39.0 [34.00, 47.00]	44.96 ± 21.71	40.5 [28.00, 62.00]

(c) Descriptive statistics for stress (STAI-S) perspective

Thermal environment	Task difficulties	N	STAI	
			M ± SD	Md [IQR]
	Easy	96	70.96 ± 11.22	68.5 [62.00, 78.00]
Slightly warm	Moderate	96	65.77 ± 10.04	62.5 [57.00, 72.00]
	Hard	96	63.21 ± 8.89	57.0 [57.00, 71.00]
	Easy	96	73.22 ± 10.21	73.0 [64.00, 78.00]
Neutral	Moderate	96	73.26 ± 9.80	71.5 [64.00, 79.00]
	Hard	96	70.93 ± 11.03	68.0 [60.00, 78.00]
	Easy	96	69.14 ± 12.56	65.0 [59.00, 75.25]
Slightly cool	Moderate	96	65.33 ± 11.84	57.0 [57.00, 70.00]
	Hard	96	62.33 ± 9.00	57.0 [57.00, 64.25]

Table 4.2 summarises the descriptive statistics of outcome variables across nine experimental conditions, stratified by thermal environment and task difficulty. The results are presented for three perspectives: thermal comfort (TSENV, TSATV), task cognitive-performance domain (OPI, NASA-TLX), and stress (STAI).

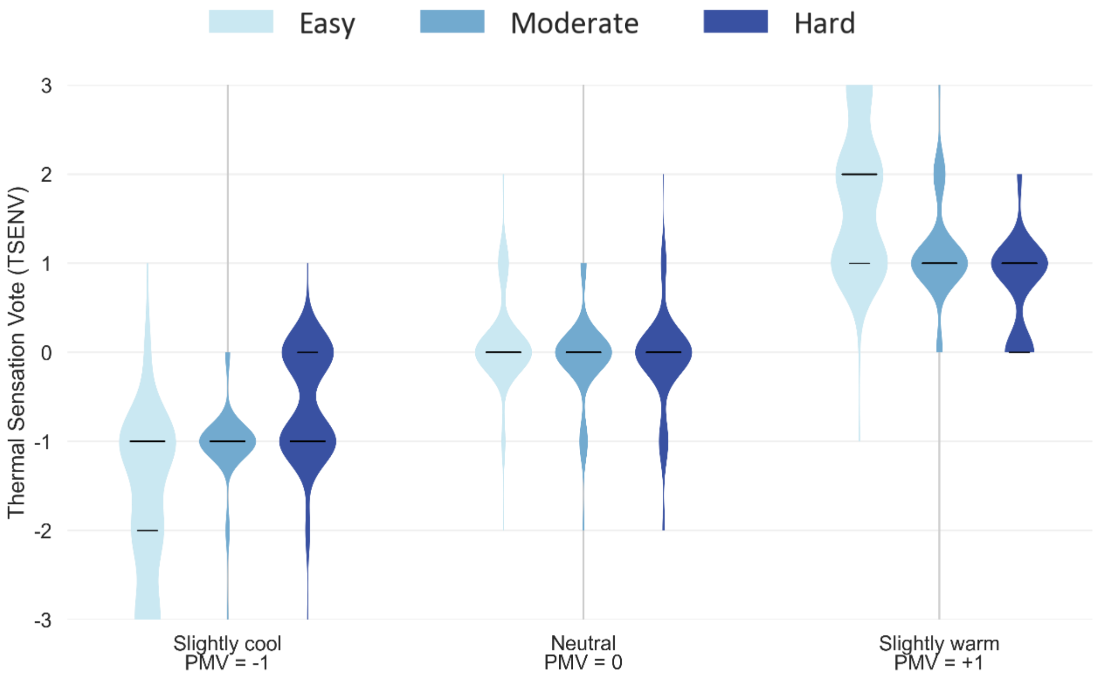
From the thermal comfort perspective, TSENV scores showed a clear trend towards neutrality (0) as task difficulty increased. For instance, under the slightly warm condition, TSENV was 1.72±0.83 at easy level, but decreased to 0.73±0.55 at hard level. A similar pattern was observed under slightly cool, with mean TSENV rising from −1.60 (easy) to −0.59 (hard). In parallel, TSATV values suggested that subjects experienced higher satisfaction at greater task difficulty: in the slightly warm condition, TSATV increased from 0.16±1.33 (easy) to 1.18±1.01 (hard). These patterns indicate that increased cognitive load may have moderated perceived deviations from thermal neutrality, leading to a relative increase in comfort.

For the task cognitive-performance domain, OPI was consistently highest under neutral thermal conditions, regardless of difficulty. At the moderate difficulty level, OPI reached 54.24±8.25, compared to 50.41±8.00 (slightly warm) and 50.55±7.36 (slightly cool). Slightly cool environments produced marginally better OPI scores than slightly warm, e.g., under hard conditions, OPI was 40.54±8.26 versus 40.27±7.37. In terms of perceived workload (NASA-TLX), values were lowest in the neutral condition, for example, 21.05±14.24 under easy difficulty compared with 24.36±14.73 (slightly warm) and 23.97±18.83 (slightly cool). At moderate and hard levels, slightly cool conditions were accompanied by slightly lower workload scores than slightly warm (e.g., 33.46±19.54 vs. 33.76±18.26 at the moderate level).

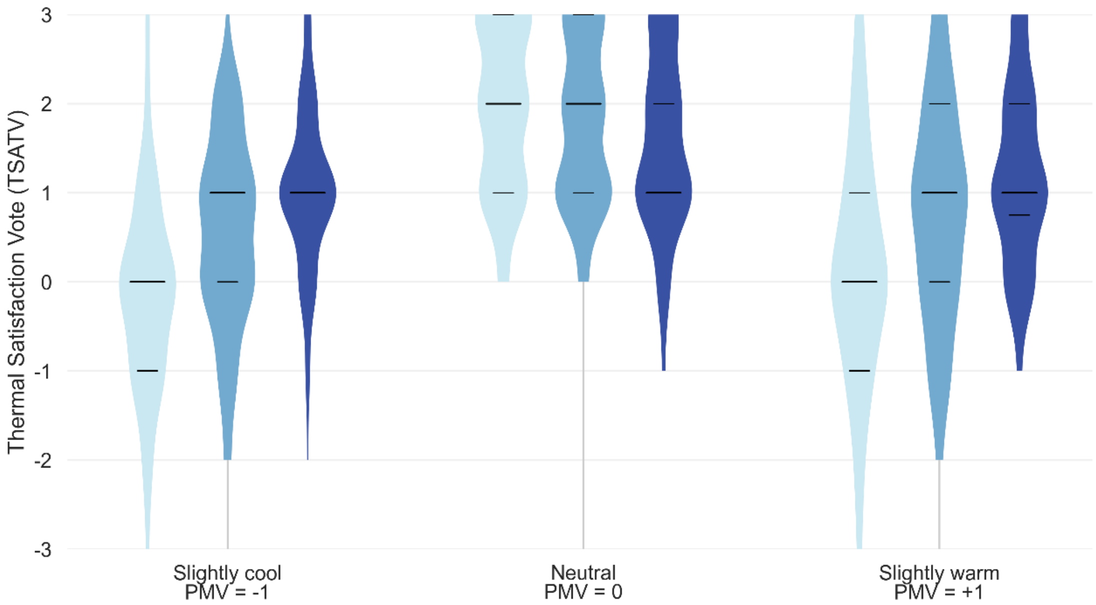
For the stress perspective, STAI scores were lowest in neutral environments and at the easy task level. Under neutral/easy conditions, STAI averaged 73.22±10.21, compared with 70.96±11.22 at slightly warm/easy and 69.14±12.56 at slightly cool/easy. Across difficulty levels, stress tended to rise from hard to easy, with neutral conditions consistently producing lower values at moderate and hard difficulties as well (e.g., 73.26±9.80 vs. 65.77±10.04 and 65.33±11.84 at moderate).

Together, these descriptive trends reveal that higher task difficulty was associated with thermal neutrality and greater comfort, neutral environments supported better performance and lower workload, and stress levels were lowest under neutral environments and easy tasks (see Figure 4.1).

(a) Violin plots for thermal comfort perspective (TSENV, TSATV)

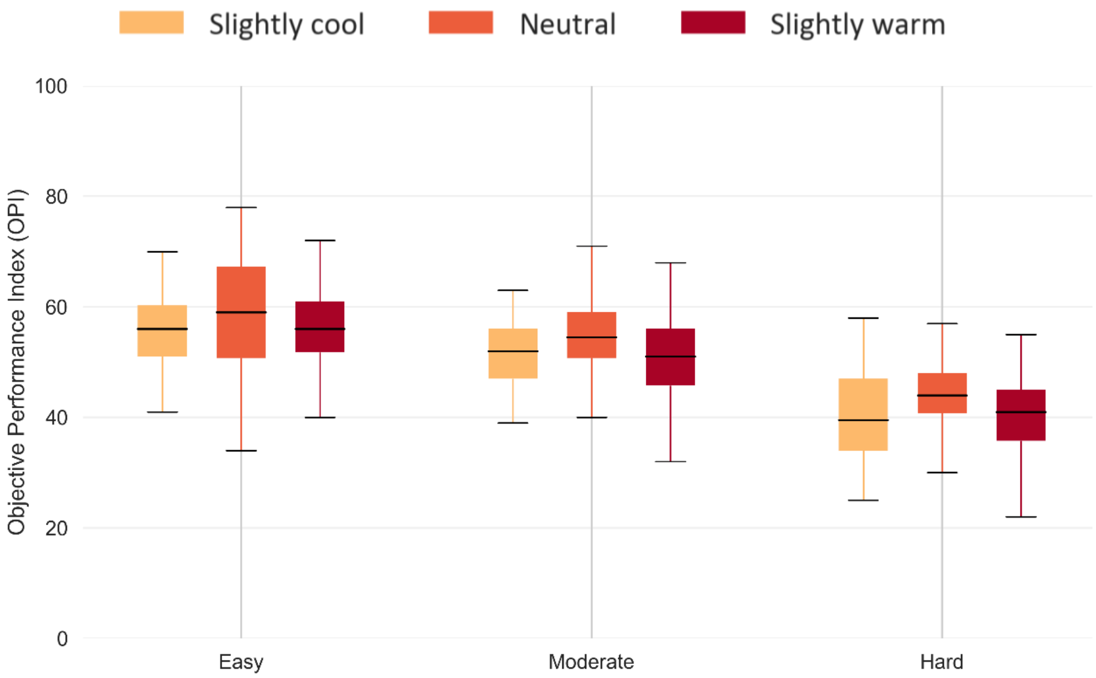


(a.i) Thermal Sensation Vote (TSENV)

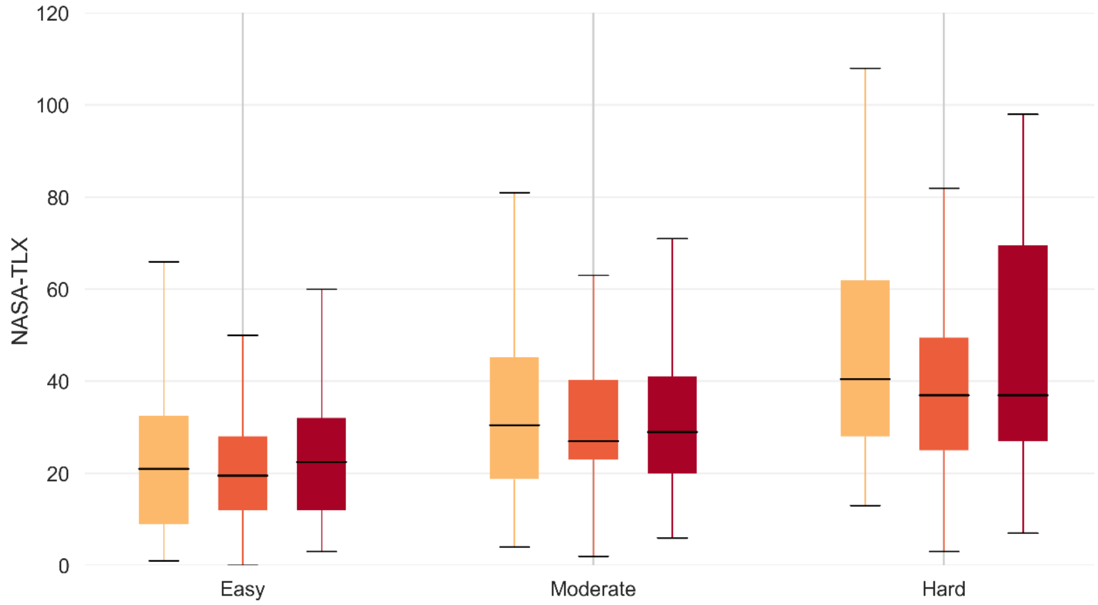


(a.ii) Thermal Satisfaction Vote (TSATV)

(b) Boxplots for TCPD perspective (OPI, NASA-TLX)

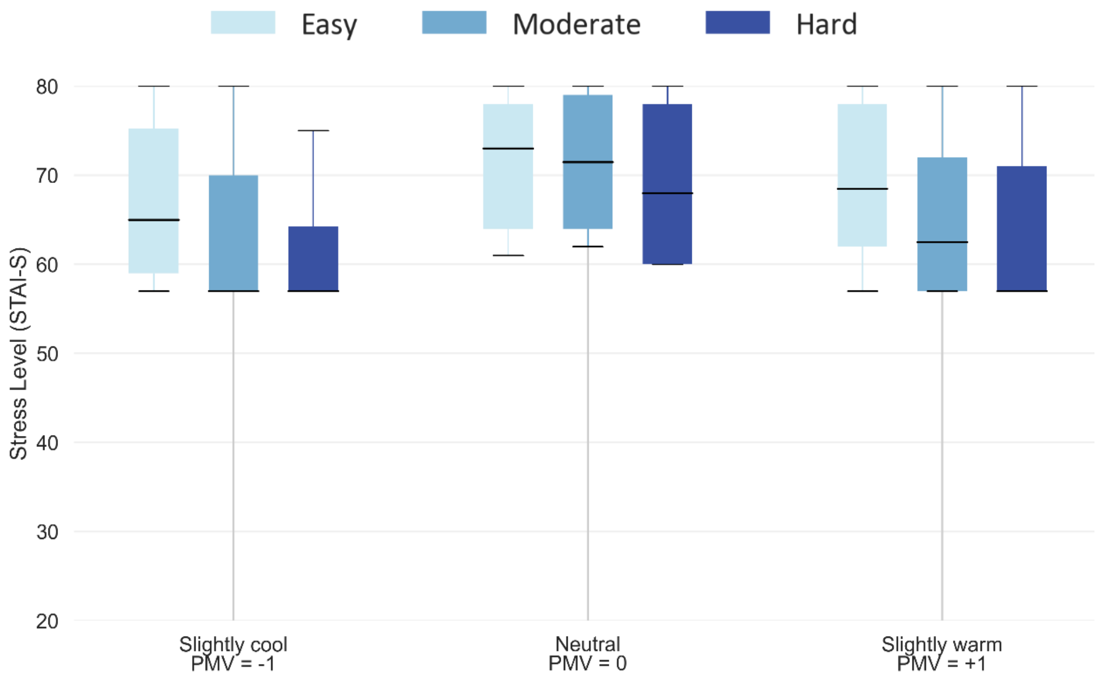


(b.i) Objective Performance Index (OPI)

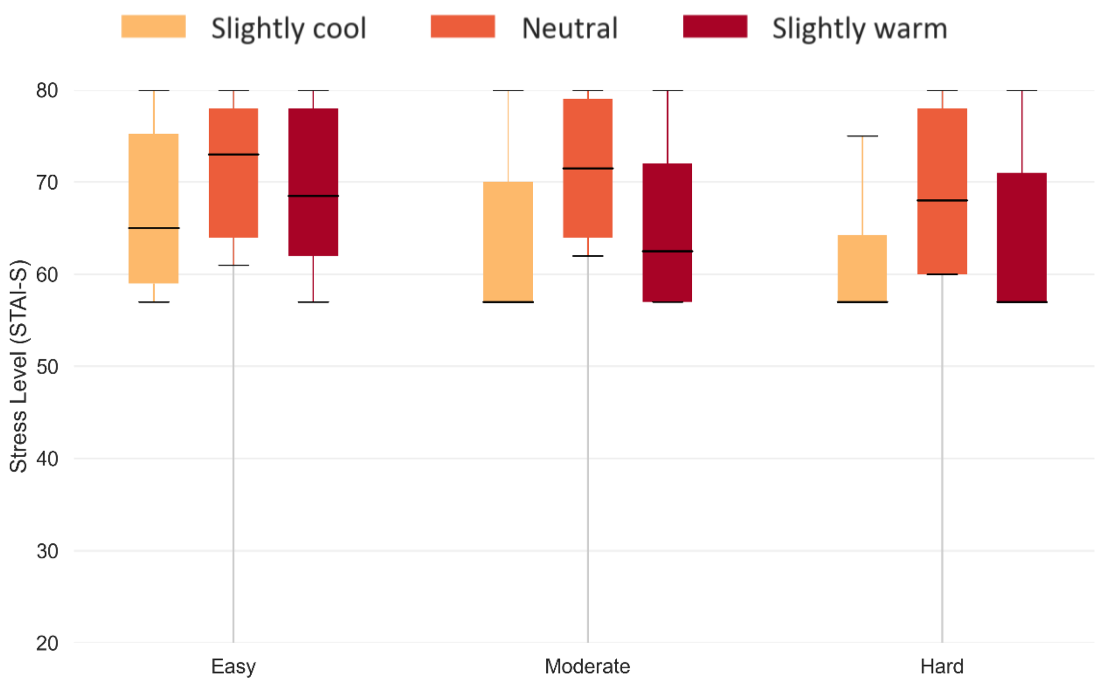


(b.ii) NASA-TLX

(c) Boxplots for stress level (STAI-S)



(c.i) STAI-S, categorised by thermal environment



(c.ii) STAI-S categoried by task difficulty

Figure 4.1 Descriptive figures for feedbacks

4.2.2 Test of normality

The feedback outcomes of the Shapiro–Wilk tests are summarised in Table 4.3. Overall, the thermal comfort indices (TSENV and TSATV) significantly departed from normal distribution under all experimental conditions (all $p < 0.001$). The same pattern was observed for the stress measure (STAI), with no subset satisfying the assumption of normality. For workload (NASA-TLX), most distributions also violated normality; only a few subsets approached borderline significance, particularly under neutral environments at medium or hard difficulty ($p \approx 0.03$–0.05). In contrast, task performance (OPI) showed more heterogeneous results: several conditions were approximately normal (e.g., Easy –1: $p = 0.222$; Hard +1: $p = 0.367$), whereas at medium difficulty, all subsets exhibited significant deviations ($p < 0.01$).

Taken together, although isolated variables such as OPI and NASA-TLX were occasionally close to normality under specific conditions, the majority of outcome distributions failed to meet the assumption. Consequently, subsequent inferential analyses were conducted using non-parametric methods, namely the Friedman test followed by pairwise post-hoc comparisons where appropriate.

Table 4.3 Shapiro–Wilk test summary of normality across outcome variables

Outcomes	**Variables**	**Normality results summary**
Thermal comfort	TSENV	All conditions rejected normality (all $p < 0.001$)
	TSATV	All conditions rejected normality (all $p < 0.001$)
TCPD	OPI	Hardly easy and hard conditions approx. normal (e.g. Easy –1: $p = 0.22$; Hard +1: $p = 0.37$); Majority of conditions rejected (all $p < 0.01$)
	NASA-TLX	Majority of conditions rejected; borderline normality under neutral PMV at moderate/hard ($p \approx 0.03–0.05$)
Stress level	STAI-S	All conditions violated normality (all $p < 0.001$)

4.2.3 Friedman test and post-hoc test

Because each subject completed three repetitions per condition, the values were aggregated by computing the mean, resulting in one representative score for each variable under each experimental condition. Based on the non-parametric requirements identified in Section 4.3, Friedman tests were then performed to examine within-subject differences. Specifically, thermal comfort variables (TSENV, TSATV) were analysed across thermal environments (slightly warm, neutral, slightly cool, hereinafter refer as +1, 0, -1, respectively), while subjective workload (NASA-TLX) and objective performance index (OPI) were examined across task difficulty levels (easy, moderate, hard, hereinafter refer as E, M, H, respectively). Stress level (STAI-S) was additionally tested across both difficulty levels and thermal environments, allowing assessment of separate and combined effects. Post-hoc Wilcoxon signed-rank comparisons with Holm correction were applied following significant omnibus effects (two-tailed). The Friedman test and post-hoc test results were presented in Table 4.4.

Table 4.4 Friedman test results and post-hoc test results

(a) TSENV (within thermal environment)					
Slightly cool	Main effect	n	Chi^2	df	p
		32	42.396	2	0.0000***
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.6916	0.0001	0.0001***
	E vs H	32	0.8124	0.0000	0.0000***
	M vs H	32	0.7363	0.0000	0.0001***
Neutral	Main effect	n	Chi^2	df	p
		32	6.758	2	0.0341*
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.2853	0.1066	0.2131
	E vs H	32	0.4405	0.0127	0.0381*
	M vs H	32	0.0321	0.8559	0.8559
Slightly warm	Main effect	n	Chi^2	df	p
		32	44.839	2	0.0000***
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.7543	0.0000	0.0000***
	E vs H	32	0.8355	0.0000	0.0000***
	M vs H	32	0.6162	0.0005	0.0005***
(b) TSATV (within thermal environment)					
Slightly cool	Main effect	n	Chi^2	df	p
		32	35.412	2	0.0000***
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.6802	0.0001***	0.0001***
	E vs H	32	0.7662	0.0000***	0.0000***
	M vs H	32	0.5698	0.0013**	0.0013**
Neutral	Main effect	n	Chi^2	df	p
		32	4.380	2	0.1119
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.1786	0.3125	0.3125
	E vs H	32	0.4058	0.0217	0.0651
	M vs H	32	0.3072	0.0822	0.1645
Slightly warm	Main effect	n	Chi^2	df	p
		32	28.852	2	0.0000***

	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.5565	0.0016**	0.0033**
	E vs H	32	0.7800	0.0000***	0.0000***
	M vs H	32	0.4664	0.0083**	0.0083**

(c) OPI (within task difficulty)

Easy	Main effect	n	Chi^2	df	p
		32	0.492	2	0.7819
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.2628	0.1371	0.4113
	-1 vs +1	32	0.0936	0.5966	0.9128
	0 vs +1	32	0.1317	0.4564	0.9128
Moderate	Main effect	n	Chi^2	df	p
		32	9.349	2	0.0093**
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.5267	0.0029	0.0087**
	-1 vs +1	32	0.0104	0.9531	0.9531
	0 vs +1	32	0.3074	0.0820	0.1641
Hard	Main effect	n	Chi^2	df	p
		32	9.571	2	0.0083**
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.4083	0.0209	0.0418*
	-1 vs +1	32	0.0468	0.7913	0.7913
	0 vs +1	32	0.4816	0.0064	0.0193*

(d) NASA-TLX (within task difficulty)

Easy	Main effect	n	Chi^2	df	p
		32	6.438	2	0.0400*
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.1983	0.2619	0.5237
	-1 vs +1	32	0.1521	0.3896	0.5237
	0 vs +1	32	0.3273	0.0641	0.1924
Moderate	Main effect	n	Chi^2	df	p
		32	4.333	2	0.1146
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.2720	0.1239	0.2814
	-1 vs +1	32	0.1521	0.3897	0.3897
	0 vs +1	32	0.2962	0.0938	0.2814
Hard	Main effect	n	Chi^2	df	p
		32	5.056	2	0.0798
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.3482	0.0489	0.1246
	-1 vs +1	32	0.0139	0.9375	0.9375
	0 vs +1	32	0.3603	0.0415	0.1246

(e) STAI-S (within thermal environment)

Slightly cool	Main effect	n	Chi^2	df	p
		32	24.692	2	0.0000***
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.4874	0.0058	0.0058**
	E vs H	32	0.7030	0.0001	0.0001***
	M vs H	32	0.5485	0.0019	0.0019**
Neutral	Main effect	n	Chi^2	df	p
		32	9.921	2	0.0070**
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.1611	0.3620	0.3620
	E vs H	32	0.3725	0.0351	0.0351*
	M vs H	32	0.3091	0.0404	0.0404*
Slightly warm	Main effect	n	Chi^2	df	p
		32	24.436	2	0.0000***
	Comparison	n_pairs	r	p_raw	p_holm
	E vs M	32	0.6710	0.0001	0.0003***
	E vs H	32	0.6976	0.0001	0.0002***
	M vs H	32	0.4519	0.0106	0.0106*
(f) STAI-S (within task difficulty)					
Easy	Main effect	n	Chi^2	df	p
		32	5.048	2	0.0801
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.4799	0.0066	0.0199*
	-1 vs +1	32	0.1988	0.2607	0.2607
	0 vs +1	32	0.2678	0.1298	0.2595
Moderate	Main effect	n	Chi^2	df	p
		32	28.711	2	0.0000***
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.5868	0.0009	0.0018**
	-1 vs +1	32	0.1308	0.4593	0.4593
	0 vs +1	32	0.7919	0.0000	0.0000***
Hard	Main effect	n	Chi^2	df	p
		32	42.780	2	0.0000***
	Comparison	n_pairs	r	p_raw	p_holm
	-1 vs 0	32	0.8003	0.0000	0.0000***
	-1 vs +1	32	0.0947	0.5922	0.5922
	0 vs +1	32	0.8069	0.0000	0.0000***

I. TSENV (thermal sensation)

Across all three thermal environments, task difficulty significantly influenced thermal sensation. In the slightly cool condition, a clear main effect was found ($Chi^2 = 42.40$, $p < 0.001$). Post-hoc results showed that sensation ratings differed significantly between every pair of difficulty levels (E vs M: $r = 0.69$, $p < 0.001$; E vs H: $r = 0.81$, $p < 0.001$; M vs H: $r = 0.74$, $p < 0.001$). In the neutral condition, the omnibus test was weaker but still significant ($Chi^2 = 6.76$, $p = 0.034$); however, only the easy–hard contrast reached significance ($r = 0.44$, $p = 0.038$), while easy–moderate and moderate–hard did not differ. In the slightly warm condition, a robust main effect emerged ($Chi^2 = 44.84$, $p < 0.001$), and all pairwise comparisons were significant (effect sizes $r = 0.62$–0.84, all $p < 0.001$).

II. TSATV (thermal satisfaction)

Task difficulty also affected satisfaction, but the pattern depended on the environment. Under the slightly cool condition, the Friedman test was strongly significant ($Chi^2 = 35.41$, $p < 0.001$). Post-hoc tests indicated decreasing satisfaction as difficulty rose (E vs M: $r = 0.68$, $p < 0.001$; E vs H: $r = 0.77$, $p < 0.001$; M vs H: $r = 0.57$, $p = 0.001$). In contrast, the neutral condition did not show a reliable omnibus effect ($Chi^2 = 4.38$, $p = 0.112$), and none of the pairwise comparisons survived Holm adjustment. In the slightly warm condition, the main effect was again significant ($Chi^2 = 28.85$, $p < 0.001$), with all pairwise differences significant ($r = 0.47$–0.78, $p \leq 0.01$).

III. OPI (objective performance index)

For easy tasks, no significant differences in performance were detected across thermal environments ($Chi^2 = 0.49$, $p = 0.782$). At the moderate difficulty level, however, performance varied significantly (Chi^2) = 9.35, $p = 0.009$). Post-hoc tests revealed worse performance at PMV = –1 compared to neutral ($r = 0.53$, $p = 0.009$), whereas other contrasts (–1 vs +1, 0 vs +1) did not reach significance. Similarly, for hard tasks, the omnibus test was significant ($Chi^2 = 9.57$, $p = 0.008$). Performance at neutral was higher compared to both cool ($r = 0.41$, $p = 0.042$) and warm ($r = 0.48$, $p = 0.019$), while the comparison between PMV = –1 and +1 was not significant ($p = 0.791$).

IV. NASA-TLX (subjective workload)

Perceived workload showed more limited sensitivity. In the easy difficulty condition, an overall effect was found ($Chi^2 = 6.44$, $p = 0.040$), although none of the pairwise tests remained significant after adjustment (all $p > 0.19$). For moderate ($Chi^2 = 4.33$, $p = 0.115$) and hard tasks ($Chi^2 = 5.06$, $p = 0.080$), no significant omnibus effects were detected, indicating that thermal environment had little consistent impact on workload ratings.

V. STAI-S (stress level)

Stress was clearly modulated by both task difficulty and thermal condition. In the slightly cool environment, a strong main effect emerged ($Chi^2 = 24.69$, $p < 0.001$). Stress increased across difficulty, with all pairwise comparisons significant ($r = 0.49$–0.70, all $p \leq 0.006$). Similar results appeared in the slightly warm environment ($Chi^2 = 24.44$, $p < 0.001$), again with all pairwise contrasts significant ($r = 0.45$–0.67, all $p \leq 0.011$). In the neutral environment, stress also varied significantly ($Chi^2 = 9.92$, $p = 0.007$), but only easy–hard ($r = 0.37$, $p = 0.035$) and moderate–hard ($r = 0.31$, $p = 0.040$) contrasts reached significance, not easy–moderate. When analysed within difficulty levels, no omnibus effect was present for easy tasks (Chi^2= 5.05, $p = 0.080$);

however, in the moderate difficulty condition, stress differed markedly across environments ($Chi^2 = 28.71$, $p < 0.001$), with lower stress at PMV = 0 than both PMV = –1 and +1 (neutral vs ±1 contrasts: $p < 0.002$), while –1 vs +1 was non-significant. For hard tasks, the effect was even stronger ($Chi^2 = 42.78$, $p < 0.001$), showing higher stress at PMV = 0 compared to both cooler and warmer conditions ($r \approx 0.80$, $p < 0.001$), whereas –1 vs +1 remained non-significant.

4.2.4 Summary

Across the tested conditions, clear interactions were observed between task difficulty and thermal environment. With increasing task difficulty, TSENV tended to shift toward neutrality, indicating that subjects perceived less deviation from the neutral state as cognitive load increased. In parallel, TSATV improved significantly with higher task demands, suggesting that subjects became more tolerant of the thermal environment under increased cognitive engagement. Regarding performance, OPI was generally highest under neutral thermal conditions, slightly lower in the cool condition, and lowest under warm conditions; however, not all pairwise contrasts reached statistical significance. Similarly, NASA-TLX tended to be lowest at neutral temperatures compared with cool and warm environments, with a significant difference detected only in the easy task condition. Finally, STAI-S showed pronounced and consistent effects across both factors: significant differences were found among all combinations of task difficulty and thermal environment, except under the easy task condition, where stress levels did not differ significantly between temperatures.

4.3 Machine-learning and deep learning results

To comprehensively examine the complex interaction between thermal environment and task difficulty, a range of machine-learning and deep-learning models were developed. The objective of these models was to identify how environmental conditions and cognitive load jointly influence human responses in terms of thermal sensation, thermal satisfaction, task performance, subjective workload, and stress. In addition to the subjective and behavioural data, thermal image features and physiological indicators (e.g., heart rate, skin temperature, and other biosignals) were incorporated into the modelling framework to enhance predictive power and to capture latent physiological–psychological relationships. By leveraging these multimodal inputs, the models aimed to uncover nonlinear and interactive mechanisms that underlie human thermal-cognitive responses.

The analytic framework consisted of two predictive paradigms distinguished by their output configuration. The single-target models (Model I.1 – I.5) were designed to predict each response variable independently, establishing baseline relationships between environmental and physiological inputs and individual outcomes. The multi-target models (Model II, Model III, and Model IV) simultaneously predicted multiple dependent variables to capture interdependence across comfort, performance, and stress dimensions. This integrated approach provided a holistic understanding of how environmental parameters, task factors, thermal images, and physiological indicators interact to shape human thermal experience and cognitive performance.

4.3.1 Single target models (Model I.1 to Model I.5)

The single-target modelling framework aimed to predict each target variable (TSENV, TSATV, OPI, NASA-TLX, and STAI-S) independently, based on the features introduced in Chapter 3, Section 3.5 Model Development. TSENV and TSATV were treated as classification tasks, while OPI, NASA-TLX, and STAI-S were handled as regression tasks.

Each model employed four representative algorithms, Logistic Regression (LR), k-Nearest Neighbours (KNN), Random Forest (RF), and Multilayer Perceptron (MLP), to ensure comparability across different learning paradigms.For models predicting thermal-related outcomes, Stage 1 utilized the original thermal and physiological features, whereas Stage 2 further integrated task-difficulty–related factors. Conversely, for models targeting task-related outcomes, Stage 1 included task features, and Stage 2 incorporated additional thermal-environment descriptors to examine cross-domain contributions. Regarding the stress model (STAI-S), Stage 1 included personal and thermal-environmental factors only, Stage 2 used personal and task-related factors, and Stage 3 combined all feature domains to evaluate their joint predictive value.

4.3.1.1 Model I.1 (TSENV)

Model I.1 was developed to predict the thermal sensation vote (TSENV) and to evaluate whether incorporating task difficulty as an additional feature could enhance model performance. This model specifically examined the interaction between individual characteristics, thermal conditions, and cognitive load in shaping perceived thermal sensation. Stage 1 established a baseline prediction using only personal and environmental variables, while Stage 2 introduced task difficulty levels to assess their incremental predictive impact. The comparison between these two stages allowed the evaluation of whether variations in task demand contribute to improved estimation of thermal sensation.

Target:

- Thermal sensation vote (TSENV)

Features:

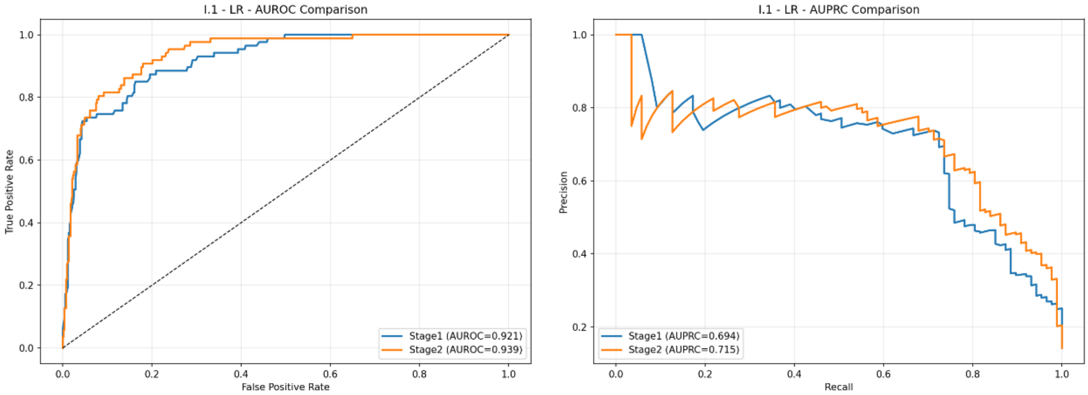
- Stage 1: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1)
- Stage 2: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1), task difficulties (easy, moderate and hard)

Table 4.5 summarizes the classification performance of Model I.1 across four algorithms (LR, KNN, RF, and MLP) in Stage 1 and Stage 2, evaluated using Accuracy, Precision, Recall, F1, AUROC, and AUPRC. Overall, the results show mixed effects after introducing task-difficulty features. For LR, accuracy slightly decreased (from 0.724 to 0.713), whereas AUROC and AUPRC increased from 0.921 to 0.939 and 0.694 to 0.715, respectively, suggesting a minor improvement in discriminative capability despite the small drop in overall accuracy. KNN exhibited a nearly unchanged accuracy (from 0.563 to 0.575) but achieved substantial increases in precision, recall, and F1, indicating that the additional feature improved sensitivity and balance among classes even though global accuracy remained modest. Random Forest (RF) showed consistent improvement across most metrics, particularly in accuracy (0.621 to 0.644) and F1 (0.349 to 0.384), reflecting enhanced stability and generalization after cognitive-load information was included. MLP maintained the highest overall performance, with accuracy increasing from 0.724 to 0.736 and F1 from 0.337 to 0.395, accompanied by consistently high AUROC and AUPRC values in both stages. Collectively, these findings suggest that adding task-difficulty features contributed to slight but meaningful gains in discriminative performance for most algorithms, while some fluctuations in accuracy may reflect trade-offs between overall classification balance and sensitivity. These trends are illustrated in Figure 4.2, which depicts the AUROC and AUPRC comparison across the four algorithms between Stage 1 and Stage 2, highlighting the nuanced performance changes following the inclusion of task-difficulty features.

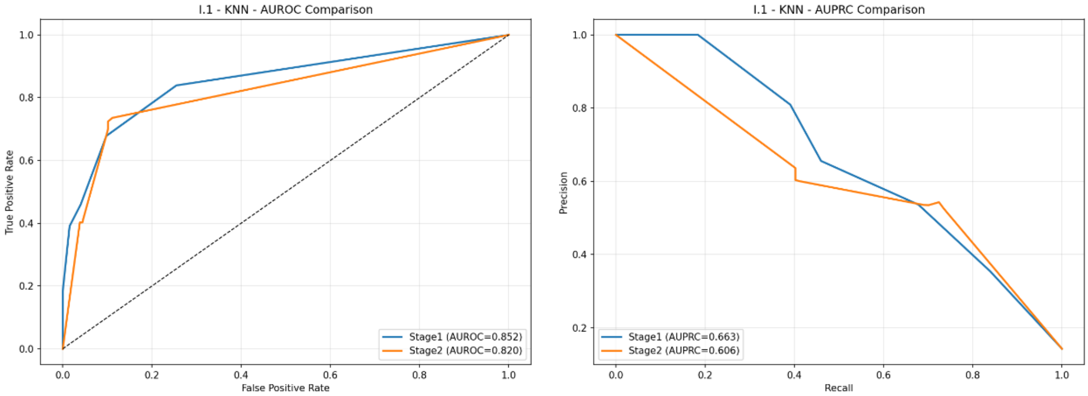
Table 4.5 Performance of Stage 1 and Stage 2 of Model I.1 (TSENV)

Algorithms	**Stage**	**Accuracy**	**Precision**	**Recall**	**F1**	**AUROC**	**AUPRC**
LR	1	0.724	0.312	0.372	0.338	0.921	0.694
	2	0.713	0.310	0.365	0.335	0.939	0.715
KNN	1	0.563	0.306	0.358	0.317	0.852	0.663
	2	0.575	0.386	0.466	0.407	0.820	0.606
RF	1	0.621	0.346	0.360	0.349	0.900	0.716
	2	0.644	0.375	0.402	0.384	0.918	0.675
MLP	1	0.724	0.309	0.372	0.337	0.921	0.720
	2	0.736	0.389	0.415	0.395	0.930	0.708

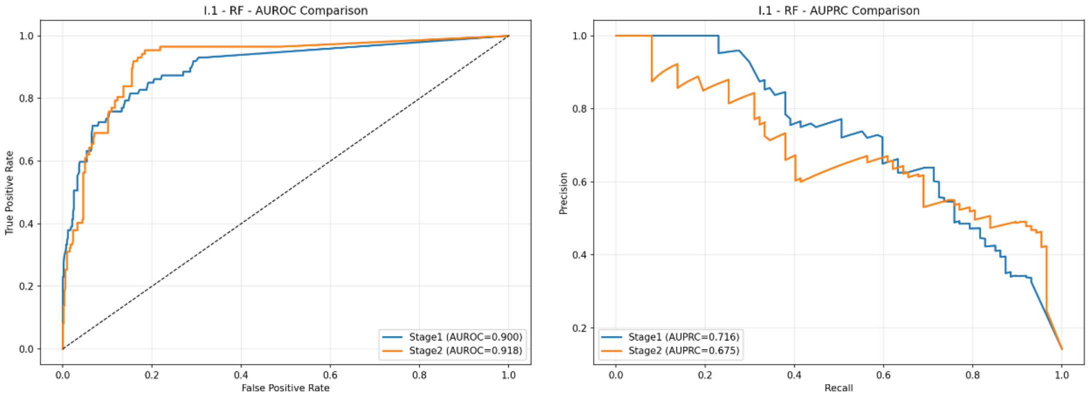
(a) Logistic Regression (LR)



(b) k-Nearest Neighbors (KNN)



(c) Random Forest (RF)



(d) Multilayer Perceptron (MLP)

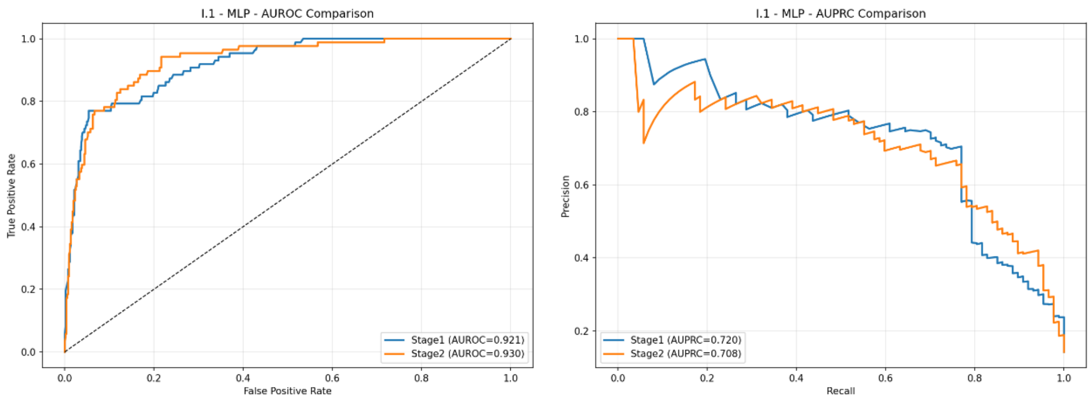


Figure 4.2 AUROC/AUPRC comparison across algorithms between Stage 1 and Stage 2 (Model I.1: TSENV)

4.3.1.2 Model I.2 (TSATV)

Model I.2 was developed to predict the thermal satisfaction vote (TSATV) and to evaluate whether incorporating task difficulty as an additional feature could enhance model performance. This model focused on exploring how personal characteristics, thermal environmental conditions, and cognitive workload together influence individuals' perceived thermal satisfaction. Stage 1 established a baseline prediction using only personal and environmental variables, whereas Stage 2 introduced task difficulty levels to assess their incremental predictive effect. The comparison between the two stages allowed the investigation of whether variations in task demand contribute to a more accurate estimation of thermal satisfaction responses.

Target:

- Thermal satisfaction vote (TSATV)

Features:

- Stage 1: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1)
- Stage 2: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1), task difficulties (easy, moderate and hard)

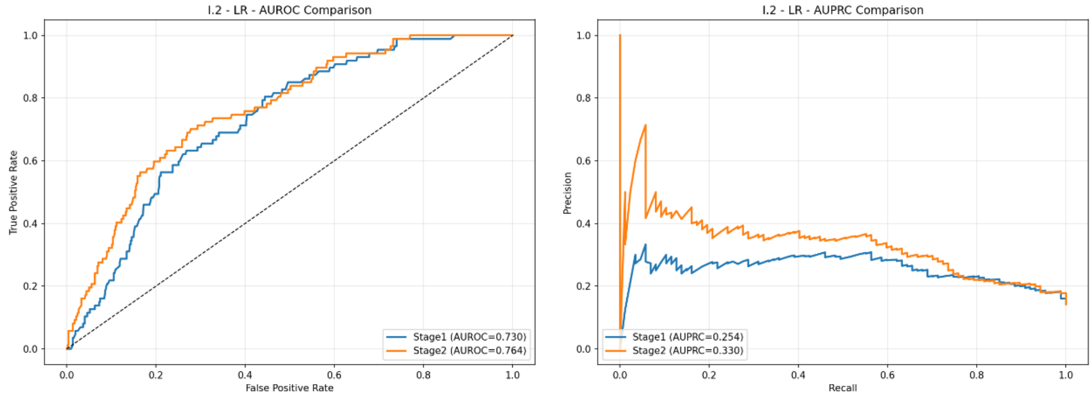
Table 4.6 presents the classification performance of Model I.2, developed to predict the thermal satisfaction vote (TSATV) using four algorithms (LR, KNN, RF, and MLP) in Stage 1 and Stage 2. Overall, model performance in Stage 1 was relatively weak, as reflected by the low Accuracy and F1 scores across all algorithms. This outcome can be attributed to the complexity of the TSATV variable, which involves seven satisfaction levels and is strongly influenced by individual differences and subjective perceptions. As a result, the use of only personal and environmental factors in Stage 1 was insufficient to capture these nuanced variations in thermal satisfaction.

Table 4.6 Performance of Stage 1 and Stage 2 of Model I.2 (TSATV)

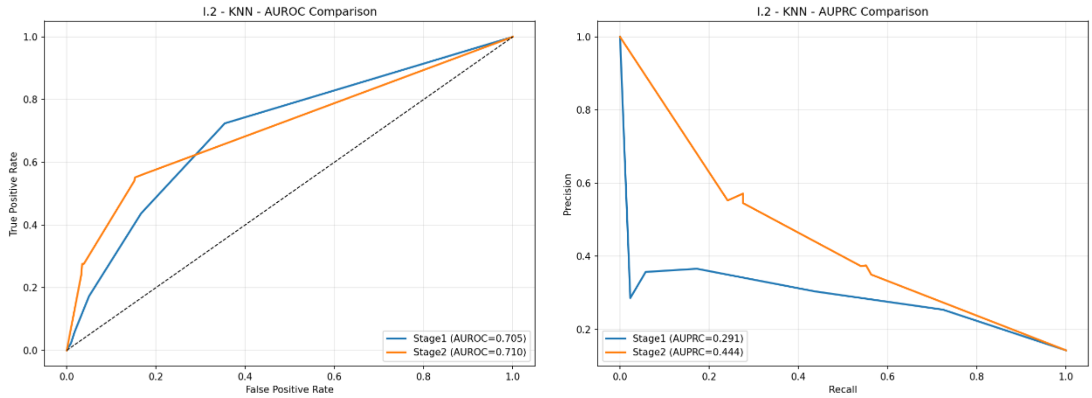
Algorithms	Stage	Accuracy	Precision	Recall	F1	AUROC	AUPRC
LR	1	0.276	0.041	0.137	0.063	0.730	0.254
	2	0.367	0.110	0.194	0.136	0.764	0.330
KNN	1	0.287	0.255	0.220	0.196	0.705	0.291
	2	0.391	0.267	0.289	0.261	0.710	0.444
RF	1	0.299	0.224	0.196	0.194	0.750	0.322
	2	0.425	0.293	0.316	0.301	0.793	0.439
MLP	1	0.310	0.211	0.198	0.197	0.770	0.343
	2	0.460	0.314	0.325	0.307	0.823	0.472

After incorporating task difficulty as an additional feature in Stage 2, all four algorithms exhibited substantial performance improvements. Logistic Regression (LR) showed increases in Accuracy from 0.276 to 0.367 and in F1 from 0.063 to 0.136, along with higher AUROC (0.730 to 0.764) and AUPRC (0.254 to 0.330), indicating enhanced discriminative ability. KNN displayed notable gains in precision, recall, and F1, suggesting that task-difficulty information improved sensitivity and class balance despite only a modest rise in accuracy. Random Forest (RF) demonstrated robust enhancement across all metrics, with Accuracy improving from 0.299 to 0.425 and F1 from 0.194 to 0.301, implying more stable prediction after the inclusion of cognitive-load information. MLP achieved the highest performance overall, with Accuracy rising from 0.310 to 0.460, F1 from 0.197 to 0.307, and noticeable gains in AUROC (0.770 to 0.823) and AUPRC (0.343 to 0.472). Collectively, these findings indicate that while Stage 1 suffered from low predictive accuracy due to the complexity and subjectivity inherent in thermal satisfaction assessment, the integration of task-difficulty features in Stage 2 significantly improved the model's predictive capability and reliability. The improvement trends are illustrated in Figure 4.3, which compares AUROC and AUPRC values between Stage 1 and Stage 2 across all algorithms.

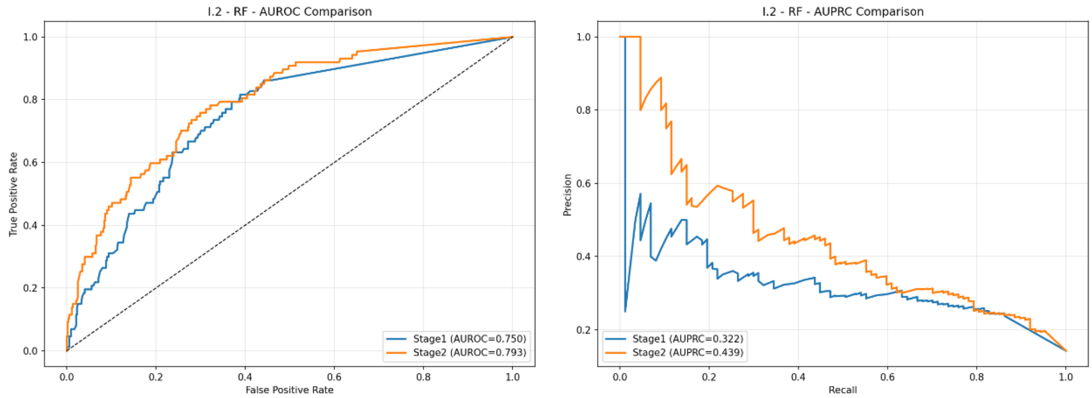
(a) Logistic Regression (LR)



(b) k-Nearest Neighbors (KNN)



(c) Random Forest (RF)



(d) Multilayer Perceptron (MLP)

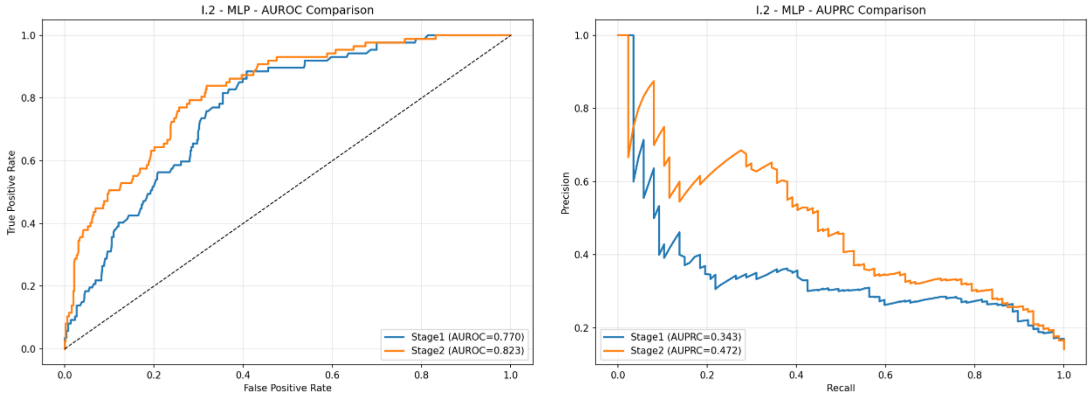


Figure 4.3 AUROC/AUPRC comparison across algorithms between Stage 1 and Stage 2 (Model I.2: TSATV)

4.3.1.3 Model I.3 (OPI)

Model I.3 was developed to predict the objective performance index (OPI) and to examine whether incorporating thermal environmental conditions as additional features could improve the model's predictive accuracy. Unlike Models I.1 and I.2, which addressed categorical thermal perception outcomes, Model I.3 is a regression model designed to estimate a continuous measure of task performance under varying thermal and cognitive conditions. This model aimed to investigate how individual characteristics, cognitive workload, and thermal environment jointly influence objective performance outcomes. In Stage 1, the model utilized only personal variables and task difficulty levels to establish a baseline, reflecting performance variation primarily driven by individual capability and cognitive demand. In Stage 2, thermal environmental factors (PMV = –1, 0, +1) were incorporated to assess their additional explanatory power. The comparison between the two stages allowed the evaluation of whether the inclusion of thermal parameters could enhance the estimation of OPI and reveal potential interactions between thermal conditions and cognitive load.Target:

- Objective performance index (OPI)

Features:

- Stage 1: Personal factors (gender, age and BMI), task difficulties (easy, moderate and hard)

- Stage 2: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1), task difficulties (easy, moderate and hard)

Table 4.7 summarises the regression performance of Model I.3, developed to predict the objective performance index (OPI) using four algorithms LR, KNN, RF and MLP,across Stage 1 and Stage 2.

Table 4.7 Performance of Stage 1 and Stage 2 of Model I.3 (OPI)

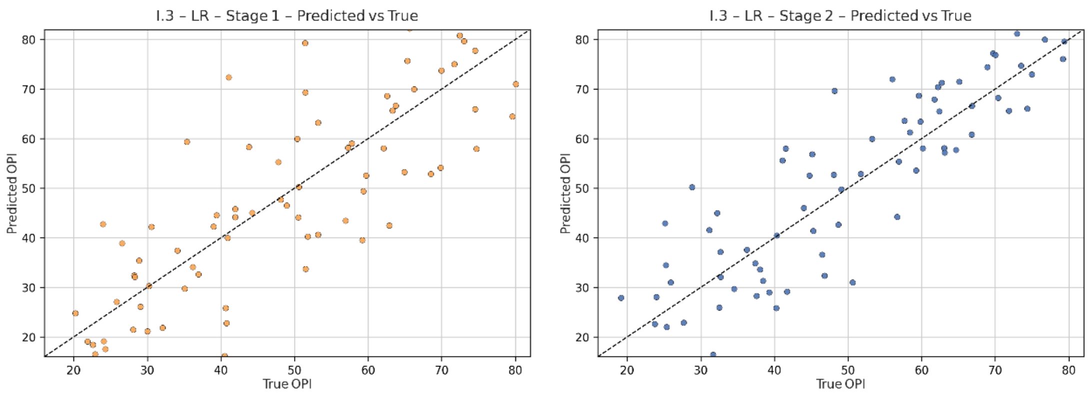
Algorithms	**Stage**	**R^2**	**RMSE**	**MAE**
LR	1	0.213	11.583	9.564
	2	0.268	9.653	7.959
KNN	1	0.336	10.476	7.824
	2	0.410	8.670	6.574
RF	1	0.347	10.163	7.547
	2	0.431	8.510	6.394
MLP	1	0.208	11.583	9.564
	2	0.259	9.717	7.999

In Stage 1, when only personal variables and task-difficulty levels were included, all models achieved modest explanatory capability, with R^2 values between 0.208 and 0.347 and relatively high RMSE and MAE errors. This indicates that individual traits and task demand alone could not fully explain the variability in OPI, likely due to the categorical nature of task difficulty versus the continuous nature of the OPI scale.

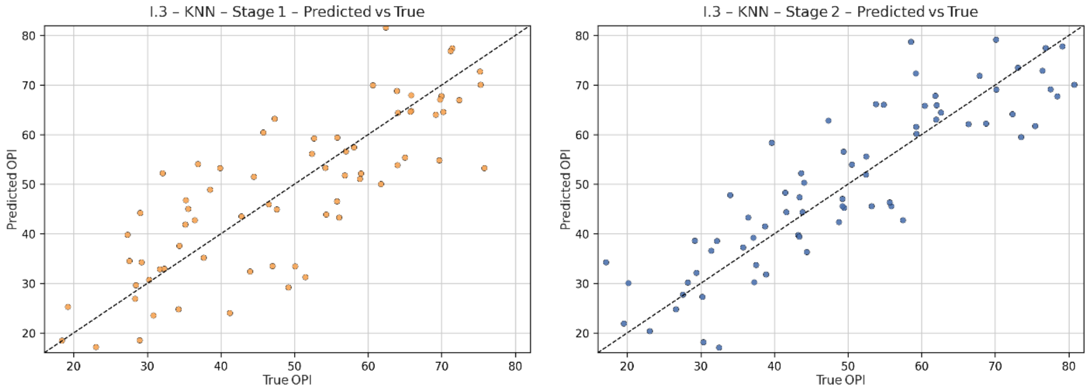
After incorporating thermal environmental features (PMV = –1, 0, +1) in Stage 2, every algorithm exhibited measurable improvement. The R^2 values increased consistently, for example, LR rose from 0.213 to 0.268, KNN from 0.336 to 0.410, and RF from 0.347 to 0.431, accompanied by notable reductions in RMSE and MAE. Among the four methods, RF achieved the highest R^2 (0.431) and the lowest RMSE (8.510) and MAE (6.394), followed closely by KNN, indicating that nonlinear and ensemble-based approaches captured the data structure more effectively than either linear regression or the neural-network model under this configuration.

Overall, the inclusion of thermal parameters in Stage 2 enhanced model stability and predictive accuracy, confirming that environmental conditions contribute additional explanatory power to task-performance modelling. The improvement trends and model fits are visualised in Figure 4.4, which presents the Predicted vs. Actual distributions for all four algorithms.

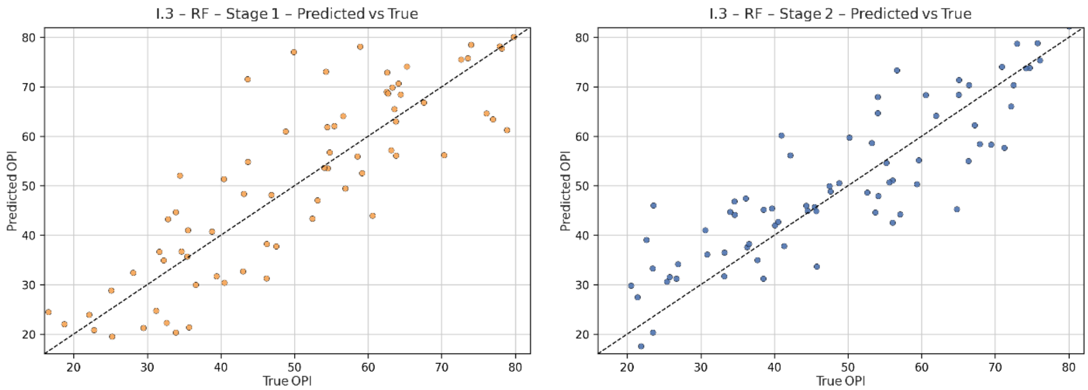
(a) Logistic Regression (LR)



(b) k-Nearest Neighbors (KNN)



(c) Random Forest (RF)



(d) Multilayer Perceptron (MLP)

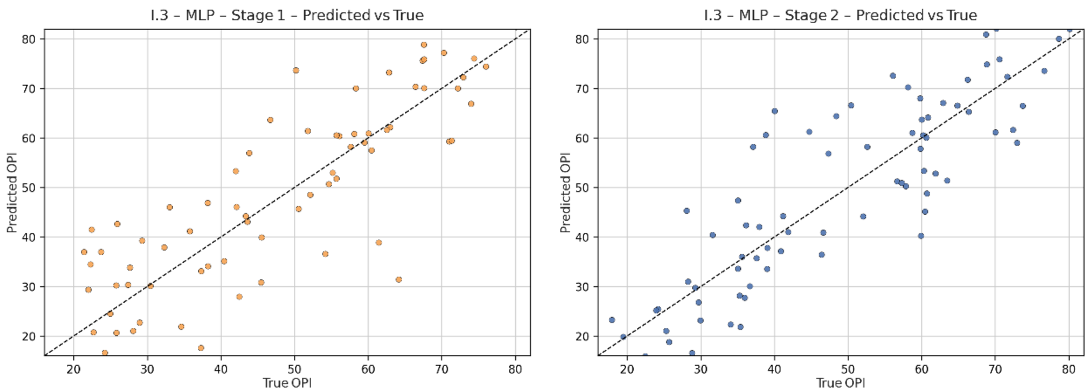


Figure 4.4 Predicted vs True comparison across algorithms between Stage 1 and Stage 2 (Model I.3: OPI)

4.3.1.4 Model I.4 (NASA-TLX)

Model I.4 was developed to predict the subjective workload level, as assessed by the NASA-TLX, and to determine whether incorporating thermal environmental parameters could enhance the model's explanatory and predictive capability. In contrast to Model I.3, which focused on objective task performance, Model I.4 addressed subjects' perceived workload under varying task difficulties and thermal conditions. This regression model aimed to examine how individual differences, task difficulty, and thermal environment jointly shape subjective workload responses. In Stage 1, the model utilized personal factors (gender, age, and BMI) together with task difficulty levels (easy, moderate, and hard) to establish a baseline that primarily reflected cognitive demand and individual characteristics. In Stage 2, thermal environmental factors (PMV = –1, 0, +1) were added alongside the personal and task features to evaluate the additional influence of thermal conditions on perceived workload. The comparison between the two stages enabled an assessment of whether the inclusion of thermal parameters improved the estimation of NASA-TLX scores and revealed potential interactions between thermal and task-related demands.

Target:

- Subjective workload (NASA-TLX)

Features:

- Stage 1: Personal factors (gender, age and BMI), task difficulties (easy, moderate and hard)

- Stage 2: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1), task difficulties (easy, moderate and hard)

Table 4.8 summarises the regression performance of Model I.4, which was developed to predict the subjective workload level (NASA-TLX) using four machine-learning algorithms, LR, KNN, RF and MLP, across Stage 1 and Stage 2.

Table 4.8 Performance of Stage 1 and Stage 2 of Model I.4 (NASA-TLX)

Algorithms	**Stage**	**R^2**	**RMSE**	**MAE**
LR	1	0.332	19.672	115.875
	2	0.428	15.779	12.590
KNN	1	0.651	13.144	8.979
	2	0.744	10.551	6.969
RF	1	0.669	12.883	8.682
	2	0.754	10.337	6.885
MLP	1	0.166	23.156	19.053
	2	0.221	18.409	15.183

In Stage 1, when only personal attributes (gender, age, BMI) and task-difficulty levels were included, the models exhibited moderate predictive capability, with R^2 values ranging from 0.166 to 0.669 and relatively large RMSE and MAE errors. These results suggest that without considering environmental conditions, the models could explain only limited variance in self-reported workload. The restricted feature set likely constrained the detection of interactions between individual characteristics and perceived task demand.

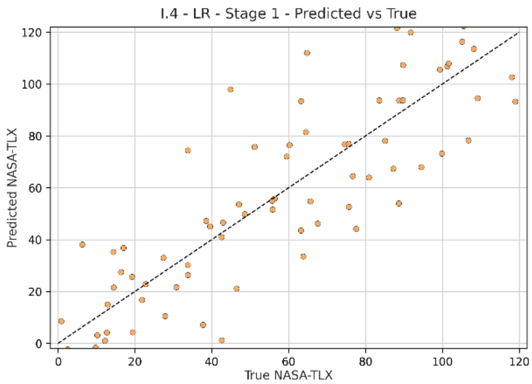
After thermal environmental factors (PMV = –1, 0, +1) were incorporated in Stage 2, all algorithms demonstrated clear and consistent performance improvements.

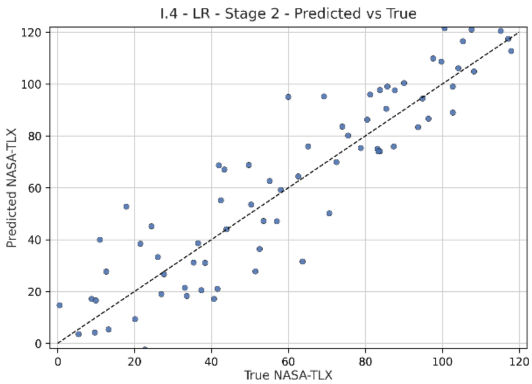
The R^2 values increased across all models, e.g., LR from 0.332 to 0.428, KNN from 0.651 to 0.744, RF from 0.669 to 0.754, and MLP from 0.166 to 0.221, accompanied by reductions in both RMSE and MAE. These gains indicate that the inclusion of thermal features enhanced the models' capacity to represent variations in subjective workload by capturing the influence of environmental comfort on cognitive and physical effort.

Among all algorithms, RF and KNN achieved the best overall performance in Stage 2, yielding the highest R^2 values (0.754 and 0.744, respectively) and the lowest error metrics. Their superior results highlight the advantage of nonlinear and ensemble-based approaches in capturing the complex, multidimensional mechanisms underlying workload perception, compared with the comparatively rigid linear regression and neural-network models in this setting.

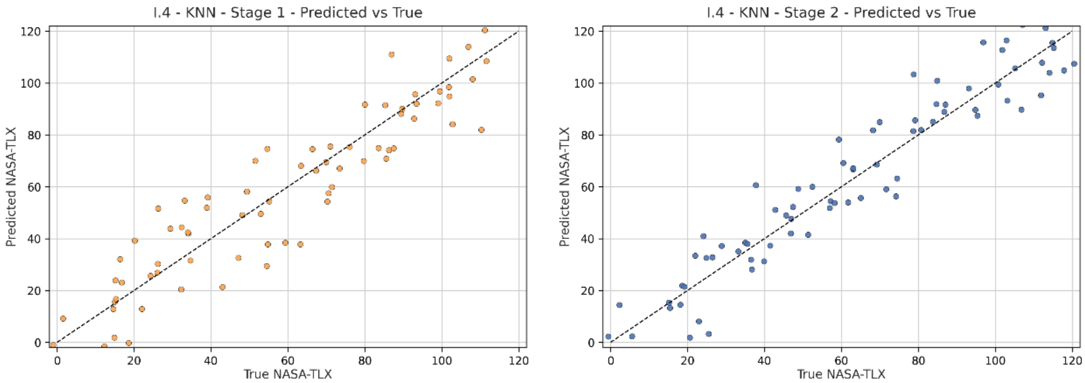
Overall, these findings confirm that Model I.4 achieved markedly stronger explanatory and predictive power in Stage 2, demonstrating the essential role of environmental parameters in improving workload prediction accuracy. The improvement trends and model fit are illustrated in Figure 4.5, which presents the Predicted vs. Actual plots for all four algorithms across both stages.

(a) Logistic Regression (LR)

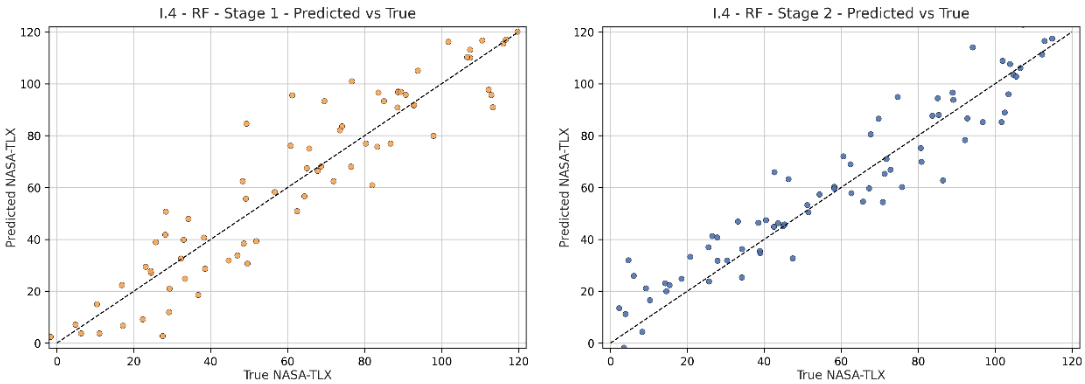




(b) k-Nearest Neighbors (KNN)



(c) Random Forest (RF)



(d) Multilayer Perceptron (MLP)

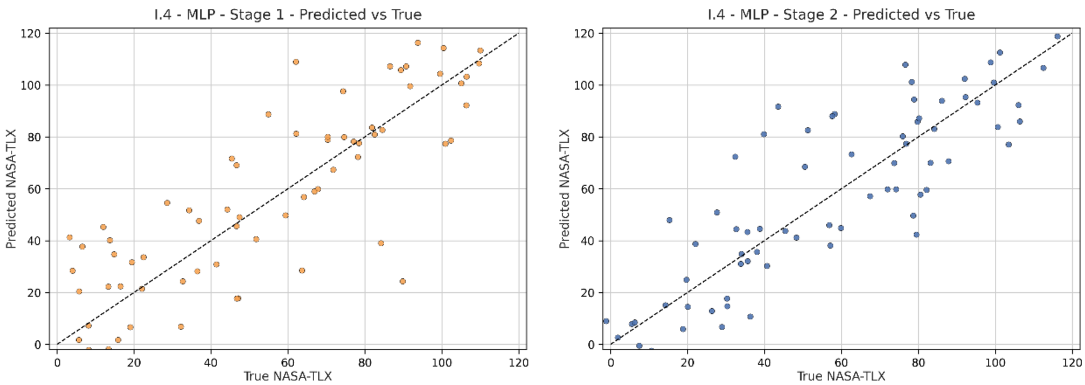


Figure 4.5 Predicted vs True comparison across algorithms between Stage 1 and Stage 2 (Model I.4: NASA-TLX)

4.3.1.5 Model I.5 (STAI-S)

Model I.5 was developed to predict the subjective stress level, as assessed by the State–Trait Anxiety Inventory – State subscale (STAI-S), and to evaluate how different combinations of personal, environmental, and task-related factors contribute to stress perception. Model I.5 extends the investigation to psychological stress responses under varying environmental and task difficulties conditions. This regression framework aimed to determine how individual characteristics, thermal environment, and task difficulty interact to influence self-reported stress.

In Stage 1, the model used personal factors (gender, age, and BMI) together with thermal environmental conditions (PMV = –1, 0, +1) to establish a baseline representing the combined effects of individual traits and thermal comfort on stress. In Stage 2, the feature set was reconfigured to include personal factors and task difficulty levels (easy, moderate, hard) to examine the impact of cognitive demand on stress in the absence of environmental variation. In Stage 3, all predictive factors, personal, thermal, and task-related, were integrated, enabling the model to capture the joint effects and possible interactions between environmental stressors and task demand.

The stepwise comparison across the three stages provides insight into the relative contributions of thermal and task factors to psychological stress and facilitates an evaluation of whether incorporating multi-dimensional contextual features leads to improved explanatory power and overall predictive accuracy of STAI-S scores.

Target:

- Stress level (STAI-S)

Features:

- Stage 1: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1)
- Stage 2: Personal factors (gender, age and BMI), task difficulties (easy, moderate and hard)
- Stage 3: Personal factors (gender, age and BMI), thermal environment (PMV -1, 0, +1), task difficulties (easy, moderate and hard)

Table 4.9 summarises the regression performance of Model I.5, which was developed to predict the subjective stress level (STAI-S) using four algorithms, LR, KNN, RF and MLP, across three stages.

Table 4.9 Performance of Stage 1 and Stage 2 of Model I.5 (STAI-S)

Algorithms	**Stage**	**R^2**	**RMSE**	**MAE**
LR	1	0.107	9.847	7.738
	2	0.040	10.213	7.900
	3	0.266	7.825	6.188
KNN	1	0.462	7.644	5.520
	2	0.637	6.276	3.846
	3	0.710	5.039	3.041
RF	1	0.442	7.787	5.643
	2	0.636	6.284	3.971
	3	0.707	5.027	3.166
MLP	1	0.295	8.750	6.999
	2	0.253	9.007	7.023
	3	0.436	6.916	5.474

In Stage 1, the models incorporated personal factors (gender, age, BMI) together with thermal environmental conditions (PMV = −1, 0, +1). The results showed moderate predictive capability, with R^2 values ranging from 0.107 to 0.462, indicating that while these combined features captured part of the variance in subjects' stress ratings, substantial unexplained variability remained.

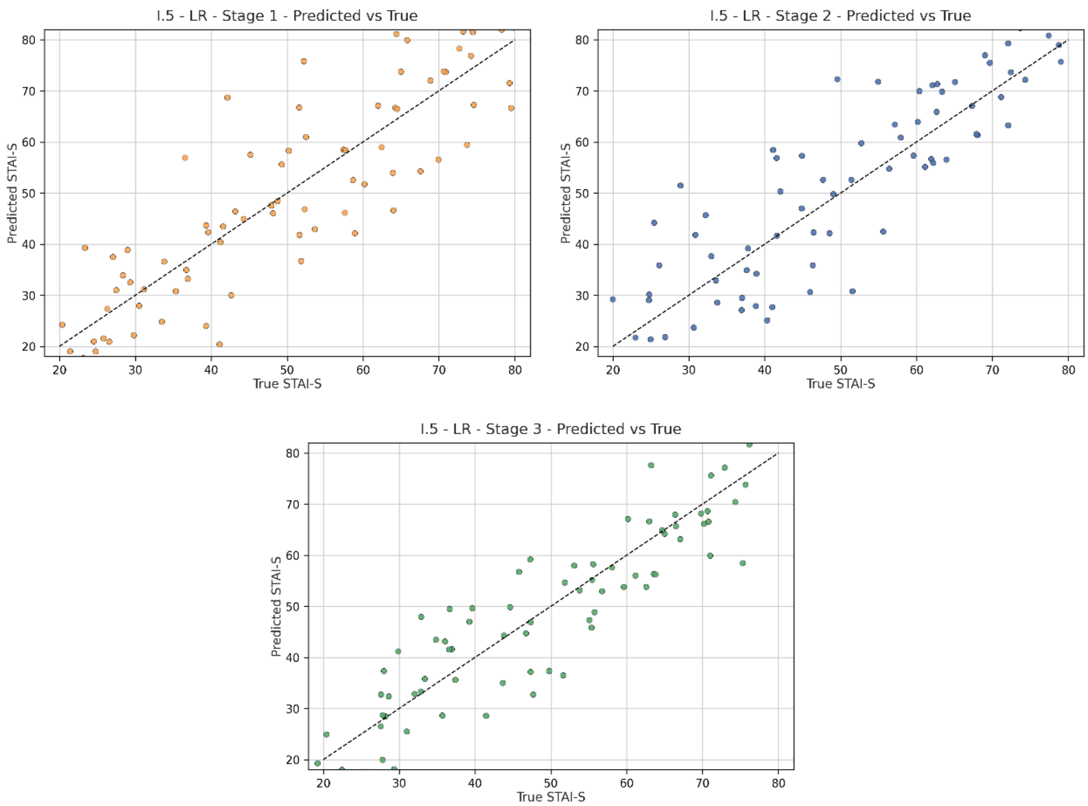
In Stage 2, thermal parameters were replaced by task-difficulty levels (easy, moderate, hard) to examine stress responses under varying cognitive demands. The results revealed mixed changes in performance across algorithms. For instance, KNN and RF improved markedly ($R^2 = 0.637$ and 0.636, respectively), reflecting stronger associations between perceived stress and workload intensity, whereas LR and MLP showed slight decreases in explanatory power ($R^2 = 0.040$ and 0.253, respectively). These divergent patterns suggest that the predictive contribution of task difficulty was meaningful but not uniformly superior to that of thermal influences, implying that individual stress responses were differentially sensitive to the two contextual domains.

In Stage 3, all three feature categories, personal, thermal, and task-related, were integrated simultaneously. This comprehensive configuration produced substantial and consistent improvements across all models, yielding the highest R^2 values and the lowest RMSE and MAE errors. For example, KNN and RF increased from approximately 0.64 to 0.710 and 0.707, respectively, while MLP rose from 0.253 to 0.436. Corresponding reductions in RMSE and MAE further confirmed the strengthened predictive performance. These results demonstrate that incorporating both environmental and cognitive factors enabled the models to more accurately capture the multifaceted mechanisms underlying stress perception.

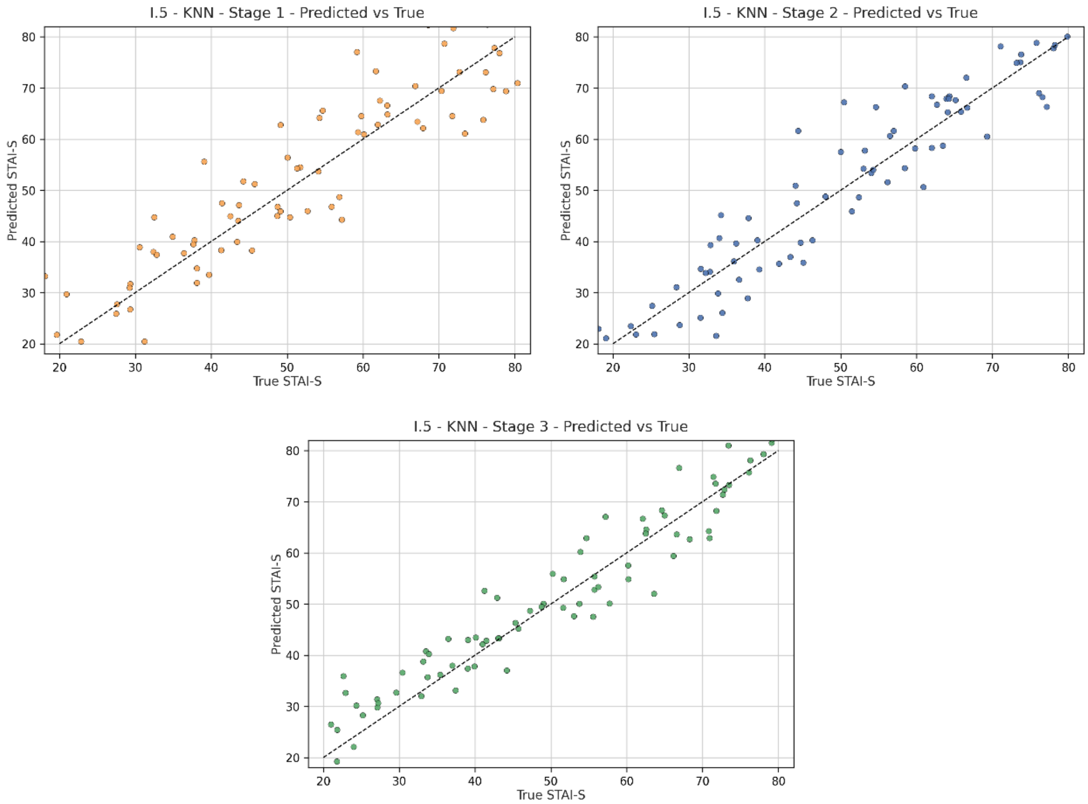
Among the algorithms, KNN and RF achieved the best overall results in Stage 3, delivering the highest accuracy and stability. Their superior performance again highlights the advantage of nonlinear and ensemble-based learning techniques for modelling complex psychological responses influenced by multiple simultaneous stressors.

Overall, Model I.5 reached its strongest explanatory power when all contextual influences were included, underscoring the complementary roles of thermal and task factors in shaping subjective stress levels. The performance improvement trends and model fits across the three stages are illustrated in Figure 4.6, which presents the Predicted vs. Actual distributions for all four algorithms.

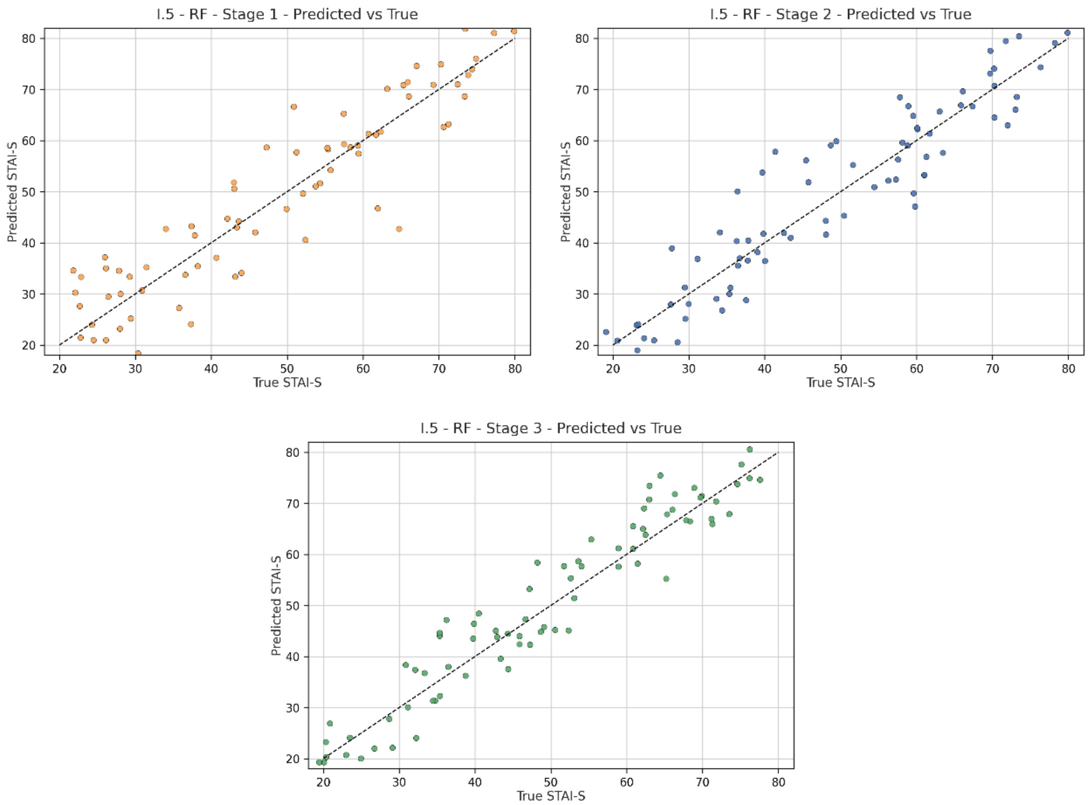
(a) Logistic Regression (LR)



(b) k-Nearest Neighbors (KNN)



(c) Random Forest (RF)



(d) Multilayer Perceptron (MLP)

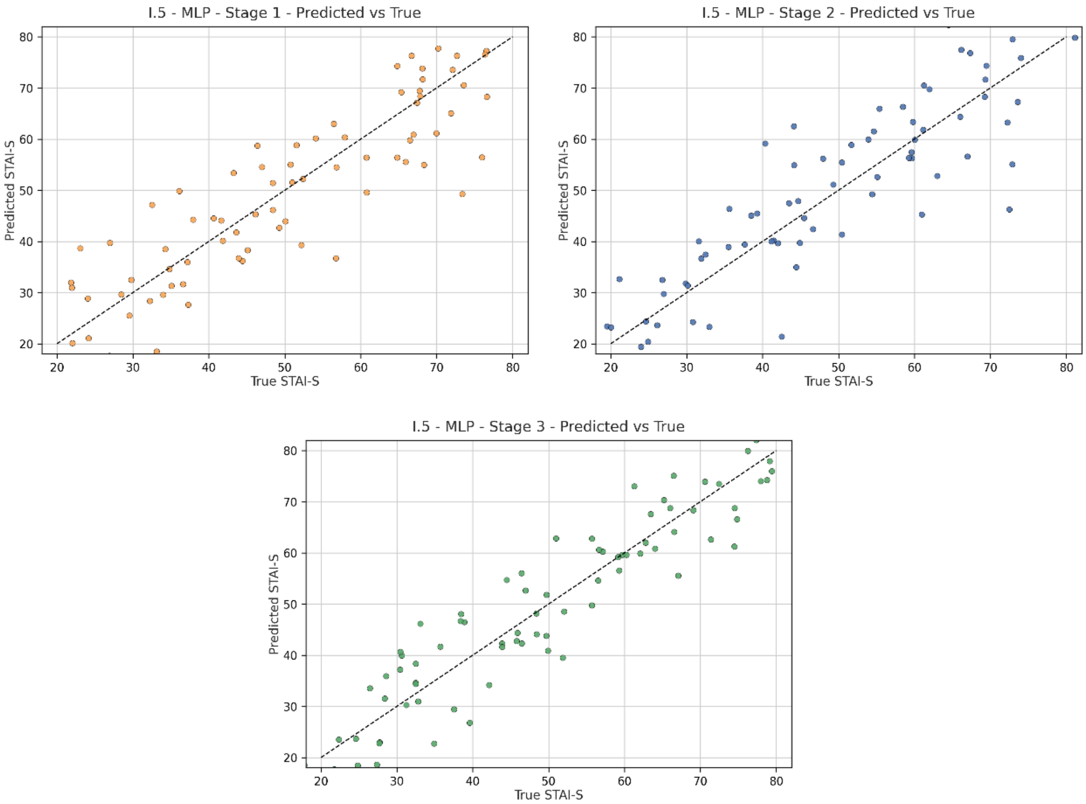


Figure 4.6 Predicted vs True comparison across algorithms between Stage 1 and Stage 2 (Model I.5: STAI-S)

4.3.1.6 Summary

Models I.1 to Model I.5 were systematically constructed to examine how personal characteristics, thermal environmental conditions, and task-related factors jointly influence human thermal perception, task performance, workload, and psychological stress. Collectively, these models form an integrated analytical framework linking physiological comfort, cognitive demand, and affective state under varying thermal and task conditions.

Models I.1 and I.2 focused on discrete thermal perception outcomes, the thermal sensation vote (TSENV) and the thermal satisfaction vote (TSATV). In both models, Stage 1 utilized personal and thermal factors, whereas Stage 2 introduced task-difficulty information. The addition of task features led to slight-to-moderate improvements in classification metrics, particularly in precision, recall, F1, and area-under-curve scores, suggesting that cognitive workload exerts measurable influence on thermal perception and satisfaction.

Model I.3 extended this approach to a continuous performance measure (OPI), revealing that models based only on personal and task factors explained limited variance in objective outcomes. The inclusion of thermal parameters in Stage 2 markedly improved model fit and reduced prediction error, confirming that environmental conditions significantly modulate task performance. Shifting from objective to subjective evaluation, Model I.4 predicted the perceived workload level (NASA-TLX). The results demonstrated clear increases in R^2 and reductions in RMSE and MAE after adding thermal features, indicating that thermal discomfort amplifies perceived workload. Nonlinear and ensemble algorithms (RF and MLP) consistently achieved higher predictive accuracy than linear methods, underscoring the complexity of these interactions.

Finally, Model I.5 targeted psychological stress levels (STAI-S) across three progressive stages. Stage 1 incorporated personal and thermal factors; Stage 2 substituted thermal features with task difficulty; and Stage 3 integrated all predictors simultaneously. The comprehensive Stage 3 configuration produced the strongest results across all algorithms, confirming that stress perception emerges from combined thermal and cognitive influences rather than any single factor alone.

Taken together, the progressive development from Model I.1 to Model I.5 demonstrates a coherent trajectory, from physical comfort to cognitive and psychological states, highlighting how incremental feature integration enhances explanatory power and predictive accuracy. The findings collectively emphasize that thermal environment, task demand, and personal characteristics interactively shape human comfort, performance, workload, and stress, providing an important empirical foundation for the design of adaptive environments that optimize both well-being and productivity.

4.3.2 Multi-target models

Building on the outcomes of the single-target models (Models I.1 – I.5), this section introduces a new generation of multi-target learning frameworks designed to simultaneously predict multiple behavioural and psychological outcomes within unified architectures. The motivation for transitioning from independent single-output estimators to integrated multi-output models lies in the high interdependence among thermal perception, cognitive workload, task performance, and stress responses. By allowing shared representations across related outputs, multi-target models are expected to improve overall predictive coherence, reduce redundancy, and enhance generalisation across heterogeneous human-state variables.

Among the developed frameworks, Model II and Model III were implemented using multilayer perceptron (MLP) architectures. The choice of MLP was driven by its flexibility and capacity to handle heterogeneous target types, specifically, both classification (e.g., thermal sensation, thermal satisfaction) and regression tasks (e.g., performance, workload, stress), within a single network. The MLP structure enables efficient joint representation learning across behavioural and psychological domains, facilitating the capture of nonlinear relationships among personal, environmental, task-related, and, in the case of Model III, physiological features.

Building upon these developments, Model IV extends the multi-target paradigm by integrating a convolutional neural network (CNN) with an MLP, forming a CNN-MLP hybrid architecture. This configuration was introduced to further enhance spatiotemporal feature extraction from physiological signal patterns, capturing subtle fluctuations in biosignals that may precede perceptual or cognitive changes. The convolutional layers effectively learn local temporal patterns and correlations within physiological inputs, while the MLP component simultaneously manages contextual and behavioural variables, merging both modalities into a unified predictive model.

Together, Models II–IV progressively evolve from contextual feature-based learning toward multimodal, representation-driven human-state modelling, advancing the framework's capability to estimate thermal comfort, performance, workload, and stress through integrated, data-driven neural architectures.

4.3.2.1 Model II

Model II marks the first application of the multi-target learning framework, developed to estimate multiple behavioural and psychological outcomes simultaneously within a unified neural model. Structured as a multi-output multilayer perceptron (MLP), the model jointly predicts thermal sensation (TSENV), thermal satisfaction (TSATV), objective performance (OPI), subjective workload (NASA-TLX), and stress level (STAI-S). This joint structure enables the exploration of shared representations among related psychological and behavioural responses rather than treating them as independent outputs.

The input feature set was designed to comprehensively capture the key contextual determinants of human responses, including personal factors (gender, age, BMI), thermal environmental parameters (PMV = –1, 0, +1), and task-related variables (task-difficulty levels). Integrating these domains allows the model to reflect the complex interactions between individual differences, thermal conditions, and cognitive demands that influence both perception and performance.

The selection of an MLP architecture was motivated by its capacity to learn nonlinear mappings across heterogeneous input spaces and to handle mixed prediction objectives that span both classification and regression tasks. Through shared representational layers, the model is able to draw connections between related outcomes, facilitating more consistent and generalisable predictions.

Overall, Model II provides a foundation for examining how unified representation learning can improve the modelling of human behavioural and psychological states within thermal-cognitive contexts. The subsequent results section evaluates its predictive performance across all target variables and discusses the interpretability and robustness of this multi-target configuration.

Table 4.10 summarises the predictive outcomes of Model II, the first multi-output MLP developed to estimate behavioural and psychological variables within a unified learning framework. While the model successfully implemented multitask representation learning, its overall predictive performance remained relatively limited across both classification and regression targets.

Table 4.10 Predictive performance of Model II

Target type	Target	Metric	
Classification		Accuracy	F1
	TSENV	0.641	0.305
	TSATV	0.369	0.136
Regression		R2	RMSE
	OPI	0.363	8.253
	NASA-TLX	0.278	17.472
	STAI-S	0.228	9.920

For the classification tasks, Model II achieved an accuracy of 0.641 and F1 of 0.305 for thermal sensation (TSENV), and a notably lower accuracy of 0.369 with F1 of 0.136 for thermal satisfaction (TSATV). These results indicate that the network was able to capture general tendencies in thermal perception but struggled to differentiate subtle categorical patterns, possibly due to overlapping subjective responses and insufficiently distinct thermal–task conditions.

In the regression tasks, the model demonstrated moderate predictive capacity, yielding $R^2 = 0.363$ (RMSE = 8.253) for objective performance (OPI), $R^2 = 0.278$ (RMSE = 17.472) for subjective workload (NASA-TLX), and $R^2 = 0.228$ (RMSE = 9.920) for stress level (STAI-S). These outcomes suggest that while the unified model could extract some meaningful associations among behavioural and psychological indicators, its explanatory power remained modest.

Overall, the performance of Model II can be characterised as relatively weak, demonstrating only a preliminary level of predictive capability. The multitask configuration appeared to impose compromises in domain-specific optimisation, leading to underfitting across heterogeneous targets. Nevertheless, it provided a valuable baseline for subsequent models by establishing the feasibility of joint learning and highlighting the need for richer, more informative features, such as physiological or thermal imaging inputs, to improve representational depth and predictive accuracy.

4.3.2.2 Model III

Model III represents an advancement of the multi-output learning framework introduced in Model II through the integration of physiological indicators into the predictive feature set. While Model II captured contextual influences derived from personal characteristics, thermal environment, and task difficulty, Model III enhances this representation by incorporating dynamic physiological variables that reflect the subjects' internal responses to changing thermal and cognitive demands. As detailed in Section 3.5.4 (Integration of Physiological Features), the added features include heart-rate-based measures and peripheral temperature variations, which serve as observable correlates of autonomic and thermal regulation processes.

By integrating these physiological attributes with contextual inputs, Model III provides a richer and more comprehensive description of the factors underlying human behavioural and psychological states. This multimodal design enables the model to capture complex bio-psychophysiological interdependencies that may not be fully represented by environmental or task variables alone. Retaining the multi-output MLP architecture ensures consistency with the previous framework while allowing the network to learn shared representations across heterogeneous targets, TSENV, TSATV, OPI, NASA-TLX, and STAI-S, augmented by physiological signal inputs.

The inclusion of physiological information is expected to enhance the sensitivity of the predictive model to subtle individual differences and transient states of thermal discomfort, workload fluctuation, and stress variation. The following section presents the results of Model III, illustrating how physiological integration influences the overall predictive accuracy and stability of multi-target estimation within thermal-cognitive contexts.

Table 4.11 summarises the predictive performance of Model III, which incorporated physiological indicators in addition to the contextual features used in Model II. Overall, the model achieved moderate predictive capability across both classification and regression tasks, demonstrating that integrating physiological information contributed useful signal variability for estimating behavioural and psychological outcomes.

Table 4.11 Predictive performance of Model III

Target type	Target	Metric	
Classification		Accuracy	F1
	TSENV	0.670	0.333
	TSATV	0.428	0.208
Regression		R2	RMSE
	OPI	0.397	8.034
	NASA-TLX	0.305	17.135
	STAI-S	0.271	9.642

For the classification targets, the model obtained an accuracy of 0.670 and F1 of 0.333 for thermal sensation (TSENV), indicating a reasonable level of sensitivity to subjective thermal perception. In contrast, the prediction of thermal satisfaction (TSATV) proved more challenging, with an accuracy of 0.428 and

F1 of 0.208, suggesting that satisfaction responses remained more individualised and less directly reflected by the available physiological inputs.

Regarding the regression targets, objective performance (OPI) reached an R^2 of 0.397 with an RMSE of 8.034, demonstrating that physiological reactivity accounted for a noticeable proportion of performance variance under different thermal–task conditions. Subjective workload (NASA-TLX) and stress (STAI-S) achieved R^2 values of 0.305 and 0.271, respectively, with corresponding RMSEs of 17.135 and 9.642, reflecting moderate yet meaningful associations between physiological states and perceived cognitive-emotional strain.

In summary, Model III confirmed the feasibility of integrating physiological features into multi-target prediction of human responses. Although predictive improvements were not uniformly substantial across all targets, the inclusion of biosignal-based variables enhanced the model's interpretive capacity by linking internal physiological regulation with perceived thermal comfort, performance, and stress. These results suggest that multimodal feature fusion is a promising direction for refining human-state estimation frameworks in complex thermal and cognitive environments.

4.3.2.3 Model IV

Model IV represents the most advanced configuration within the multi-target framework, extending the multimodal representation established in Model III through the incorporation of thermal image features. Unlike the previous models, which relied primarily on numerical descriptors of environmental, task, and physiological variables, Model IV introduces visual information extracted from infrared thermal imaging, providing a direct spatial representation of surface temperature distribution across subjects. This additional modality enables the model to capture subtle, localised thermal variations on the human body that may correspond to physiological regulation processes, stress reactivity, or cognitive load changes.

To process this image-based information effectively, Model IV adopts a hybrid CNN-MLP architecture. The convolutional neural network (CNN) component is responsible for learning structured spatial patterns within the thermal images, identifying regions or gradients associated with thermal discomfort or emotional arousal, while the MLP module concurrently handles contextual, behavioural, and physiological features. The two components are subsequently concatenated in a unified embedding layer, allowing the model to fuse visual and non-visual information within an integrated representational space.

This design leverages the complementary strengths of both architectures: CNNs excel in spatial feature extraction from high-dimensional image data, whereas MLPs provide strong flexibility for heterogeneous feature types and multi-output regression and classification tasks. By combining them, Model IV aims to enhance the predictive capacity and robustness of human-state estimation, moving toward a fully multimodal learning framework that jointly interprets visual, physiological, task, and environmental cues.

The following section presents the predictive results of Model IV, evaluating how the integration of thermal imagery influences the estimation accuracy of thermal perception, performance, workload, and stress relative to the prior MLP-based models.

Table 4.12 presents the predictive performance of Model IV, which integrated thermal image features into the multi-output learning framework by combining a convolutional neural network (CNN) with an MLP architecture. This configuration enabled the model to extract spatial temperature distributions from infrared imagery and fuse them with contextual, physiological, and task-related variables for joint prediction of behavioural and psychological outcomes.

Table 4.12 Predictive performance of Model IV

Target type	Target	Metric	
Classification		Accuracy	F1
	TSENV	0.670	0.333
	TSATV	0.428	0.208
Regression		R2	RMSE
	OPI	0.397	8.034
	NASA-TLX	0.305	17.135
	STAI-S	0.271	9.642

For the classification tasks, Model IV achieved an accuracy of 0.670 and F1 of 0.333 for thermal sensation (TSENV), and an accuracy of 0.428 with F1 of 0.208 for thermal satisfaction (TSATV). These represent notable improvements compared to previous models, suggesting that image-derived thermal patterns contributed additional discriminatory information about perceived thermal states. The enhancement indicates that visual thermal cues, such as localised skin temperature gradients, were successfully captured and leveraged by the CNN component.

In the regression tasks, the model also demonstrated modest gains. Objective performance (OPI) attained an R^2 of 0.397 with RMSE = 8.034, reflecting slightly higher predictive accuracy relative to earlier configurations. Similarly, subjective workload (NASA-TLX) achieved $R^2 = 0.305$ (RMSE = 17.135), and stress (STAI-S) obtained $R^2 = 0.271$ (RMSE = 9.642), indicating improved sensitivity to continuous behavioural and psychological variation. These results suggest that the fused CNN-MLP representation effectively captured complex relationships among visual, physiological, and cognitive inputs.

Overall, Model IV exhibited the strongest and most balanced performance among the explored architectures. The integration of thermal imagery enhanced the model's ability to interpret embodied responses to thermal and cognitive conditions, resulting in improved generalisation across both categorical and continuous outcomes. Although the magnitude of improvement remained moderate, the findings underscore the advantage of multimodal deep learning, particularly the use of thermal imaging, for more accurate and holistic modelling of human behavioural and psychological states.

4.3.2.4 Comparison and summary

This section provides a comparative analysis of the modelling outcomes across Model I to Model IV, aiming to evaluate the progressive enhancements in predictive performance arising from the inclusion of additional feature modalities and architectural refinements. In this comparison, Model I represents a series of single-target models (Models I.1–I.5), each optimised independently for its respective outcome using the best-performing algorithm identified for that target, rather than a uniform MLP configuration. Model II, by contrast, implemented the first multi-output multitask framework, enabling shared representation learning across behavioural and psychological dimensions. Building upon this foundation, Model III integrated physiological indicators, addressing individual variability in thermal and cognitive responses, while Model IV further incorporated thermal imaging features, realising a fully multimodal CNN–MLP architecture capable of learning from both numerical and visual inputs.

The comparative evaluation examines how each successive model influenced the balance between classification and regression performance, the trade-offs between generality and target-specific optimisation, and the extent to which multimodal integration enhanced prediction stability and interpretability. Through this analysis, the study assesses whether an expansion in feature diversity and network complexity, from single-modality, independently trained models to an integrated multimodal learning framework, leads to measurable and consistent improvements in modelling human behavioural and psychological states under varying thermal–cognitive conditions.

Table 4.13 summarises the overall predictive performance from the single-target baseline models (Model I.1–I.5) to the progressively enhanced multimodal and multi-output architectures (Models II–IV). The results reveal distinct performance patterns across classification and regression tasks, reflecting how architectural design and feature diversity jointly influenced model effectiveness.

Table 4.13 Overall comparison of predictive performance from single-target (Model I.1 to Model I.5) to multimodal multi-output models (Models II to Model IV)

Target type	Target	Model	Metric	
Classification			Accuracy	F1
		Model I. 1	0.736	0.395
	TSENV	Model II	0.641	0.305
		Model III	0.670	0.333

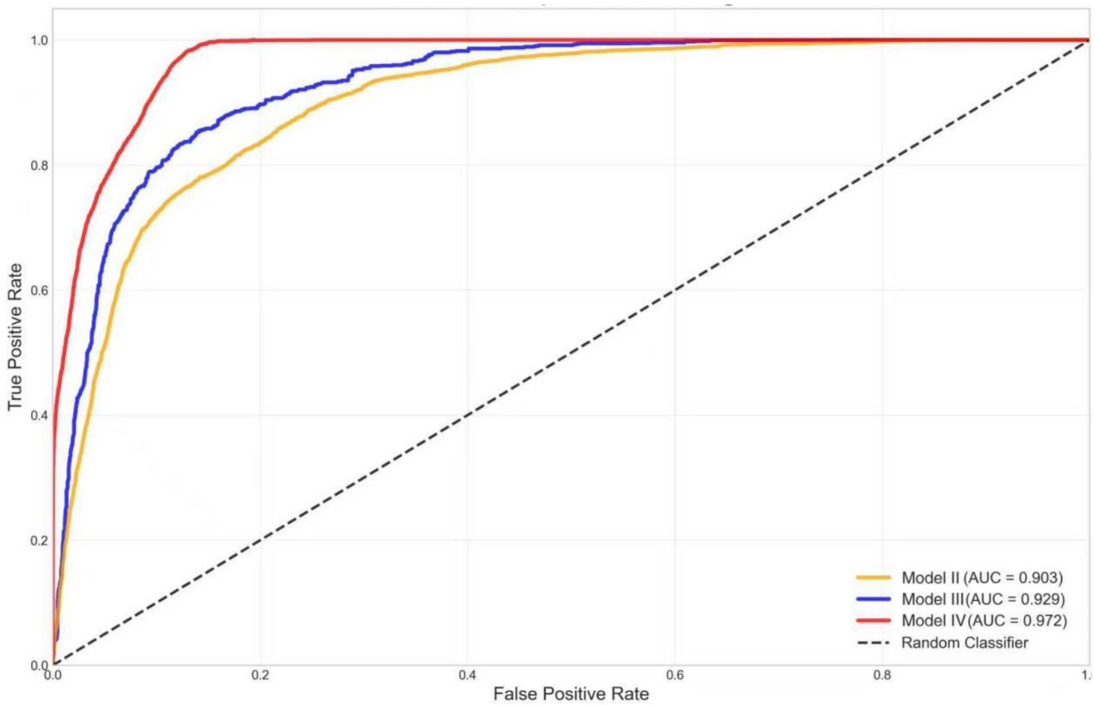
		Model IV	0.745	0.620
	TSATV	Model I. 2	0.460	0.307
		Model II	0.369	0.136
		Model III	0.428	0.208
		Model IV	0.617	0.536
Regression			R^2	RMSE
	OPI	Model I. 3	0.431	8.510
		Model II	0.363	8.253
		Model III	0.397	8.034
		Model IV	0.604	6.511
	NASA-TLX	Model I. 4	0.754	10.337
		Model II	0.278	17.472
		Model III	0.305	17.135
		Model IV	0.844	8.122
	STAI-S	Model I. 5	0.710	5.039
		Model II	0.228	9.920
		Model III	0.270	9.642
		Model IV	0.779	5.308

For the classification targets, thermal sensation (TSENV) and thermal satisfaction (TSATV), Model IV achieved the best overall performance, with accuracies of 0.745 and 0.617, and F1-scores of 0.620 and 0.536, respectively. These values represent substantial improvements compared with all earlier models, particularly over Model II and Model III, demonstrating the strong discriminative potential of thermal image features. Interestingly, while the single-target Model I.1 performed well for TSENV (accuracy = 0.736, F1 = 0.395), Model IV slightly surpassed it, highlighting the benefit of multimodal fusion and the joint learning of contextual, physiological, and visual information. Similarly, TSATV predictions improved markedly from Model I.2's F1 = 0.307 to Model IV's F1 = 0.536, indicating enhanced capability to capture subjective thermal satisfaction through multimodal representations.

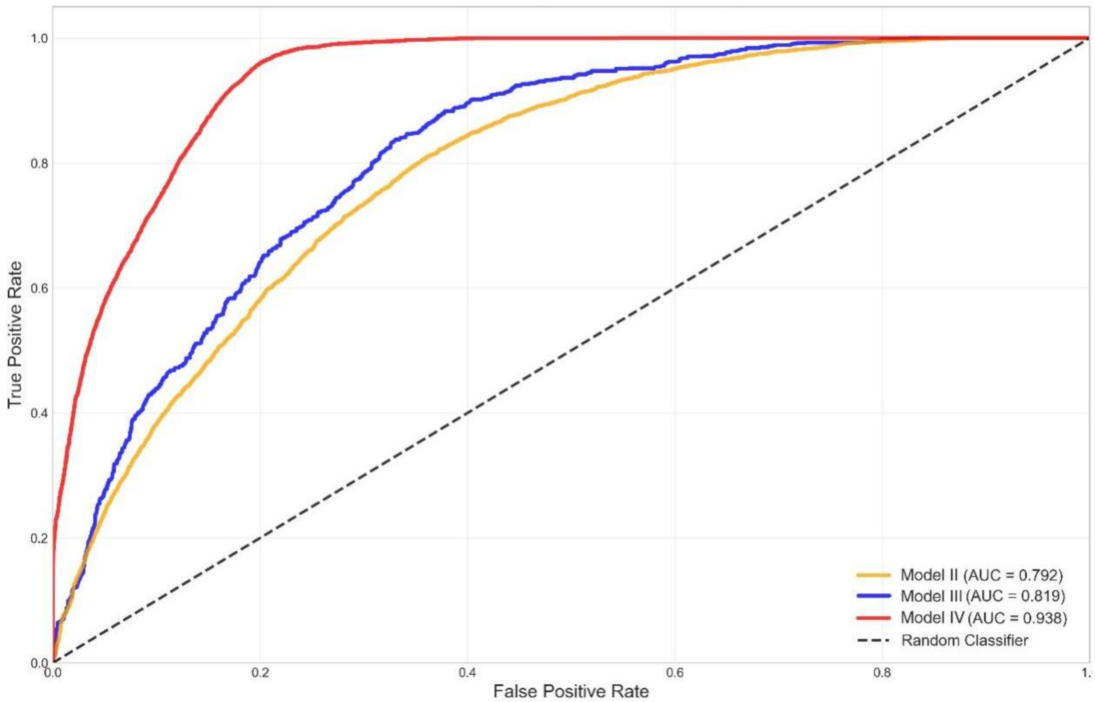
Regarding the regression targets, Model IV again demonstrated the most consistent and robust predictive performance. It achieved $R^2 = 0.604$ (RMSE = 6.511) for objective performance (OPI), $R^2 = 0.844$ (RMSE = 8.122) for subjective workload (NASA-TLX), and $R^2 = 0.779$ (RMSE = 5.308) for stress level (STAI-S). These outcomes represent major improvements relative to earlier models, particularly compared to Model II and Model III, whose R^2 values remained below 0.40 for most targets. Notably, the multimodal CNN-MLP in Model IV outperformed even the best single-target models (Model I.3–I.5), suggesting that cross-target feature learning and multimodal integration contributed to both accuracy and generalisability.

To further illustrate the comparative performance across the three multi-output models, Figure 4.7 provides a comprehensive visualisation of the predictive outcomes for all target variables. Subplots (a) and (b) present the ROC curves for the classification targets thermal sensation (TSENV) and thermal satisfaction (TSATV), respectively, while subplots (c) – (e) display the predicted versus true value distributions for the regression targets objective performance (OPI), subjective workload (NASA-TLX), and stress level (STAI-S). Across all panels, the curves and scatterplots exhibit progressively improved performance from Model II to Model IV, confirming the benefit of multimodal integration. The ROC curves reveal a steady increase in the area under the curve (AUC), indicating enhanced discriminative ability when physiological and thermal imaging features are incorporated. Likewise, the regression plots show that the predicted values of Model IV cluster more closely around the diagonal, reflecting better calibration and lower error variance. These visual patterns corroborate the numerical results in Table 4.13, demonstrating that successive model enhancements substantially strengthened both classification and regression accuracy.

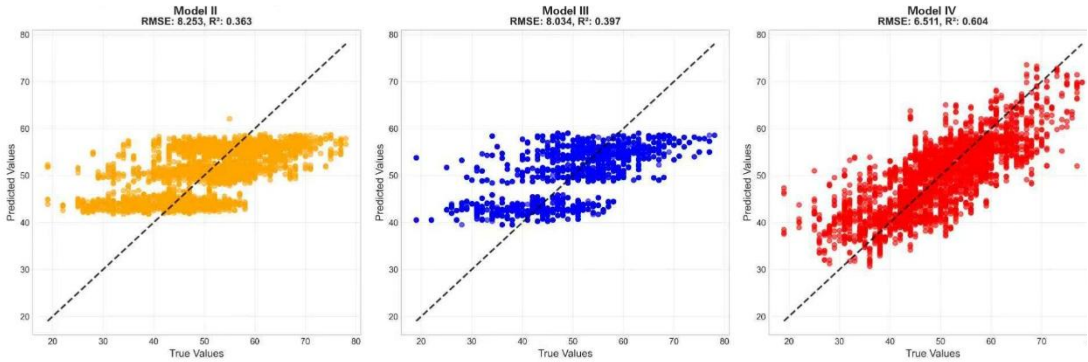
(a) ROC curve for thermal sensation (TSENV)



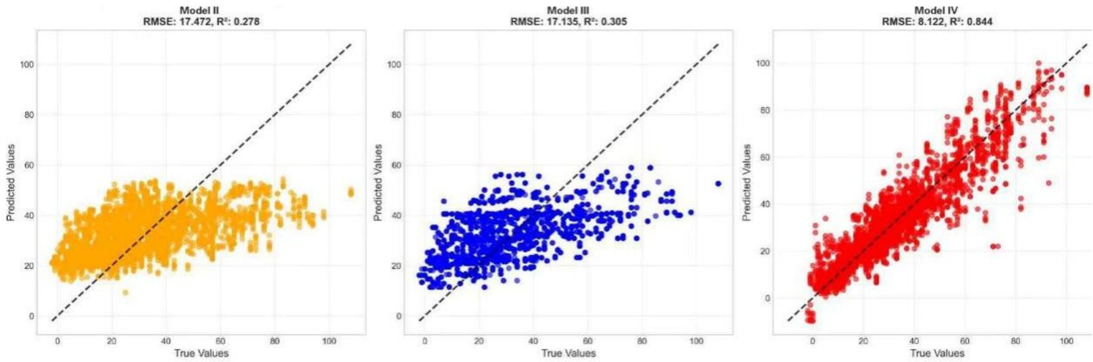
(b) ROC curve for thermal satisfaction (TSATV)



(c) Predicted vs True Values for objective performance (OPI)



(d) Predicted vs True Values for subjective workload (NASA-TLX)



(e) Predicted vs True Values for stress level (STAI-S)

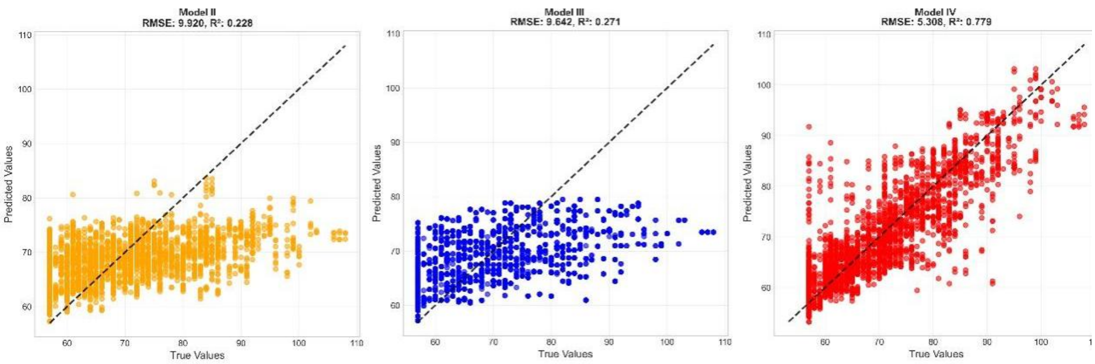


Figure 4.7 Overall performance visualisation of Models II–IV across classification and regression targets

Similarly, the predicted-versus-true scatterplots reveal that Model IV's predictions cluster more tightly around the diagonal line, indicating improved calibration and reduced error variance relative to Models II and III. These visual analyses complement the quantitative comparison in Table 4.13, highlighting the tangible benefits of multimodal feature fusion for both classification and regression tasks.

Overall, the results indicate a clear upward trend in predictive capacity from the simpler single-target models to the more complex multimodal architectures. While the multitask structure of Model II initially reduced performance due to shared-representation interference, the subsequent inclusion of physiological signals (Model III) and especially thermal imagery (Model IV) progressively enhanced model expressiveness. Consequently, Model IV demonstrated superior predictive stability across all outcomes, validating the advantage of combining heterogeneous feature modalities and deep-learning hierarchies for assessing human thermal perception, performance, and psychological states under thermal–cognitive interactions.

4.4 Summary

This chapter presented an integrated examination of the experimental, statistical, and modelling results concerning the combined effects of thermal environment and task difficulty on thermal comfort, task performance, and stress level.

The statistical analyses revealed clear interaction effects between the two factors. As task difficulty increased, thermal sensation (TSENV) shifted toward neutrality, indicating that subjects became less sensitive to thermal deviations when cognitively engaged. In parallel, thermal satisfaction (TSATV) improved with higher task demands, suggesting greater tolerance to environmental conditions under increased cognitive load. Regarding performance and psychophysiological responses, OPI was generally highest under the neutral condition, whereas NASA-TLX scores were lowest in the same environment, implying that moderate thermal conditions supported both efficiency and reduced workload. In contrast, STAI-S exhibited a strong and consistent interaction pattern, with stress levels varying significantly across combinations of thermal and task conditions. These findings collectively demonstrate that thermal comfort, task performance, and stress responses are jointly shaped by both environmental and cognitive factors, and are significantly interrelated.

The subsequent machine-learning analyses built upon these empirical findings by quantifying and predicting the observed interrelations. The single-target models (Models I.1–I.5) achieved reasonable accuracy for their respective outcomes, thermal comfort, task performance, workload, and stress level, but their predictive range was restricted since each model focused on one target independently. As multimodal and multi-output architectures were introduced (Models II–IV), performance improved consistently. Model II first combined all outcomes in a multitask structure, capturing shared representations but showing some performance interference among outputs. Model III further integrated physiological features, enabling more reliable predictions of stress and workload by capturing internal physical responses to thermal and cognitive stimuli. Model IV advanced this framework by incorporating both physiological and thermal image information, allowing the model to recognise the joint effects of environmental and task factors on human responses. Consequently, Model IV achieved the highest overall accuracy and stability across all targets, effectively describing the coupled influences of thermal conditions, task load, and psychological states.

Overall, the results confirm that the thermal environment and task demands exert interacting influences on thermal comfort, TCPD, and stress level, all of which are mutually correlated. The integration of statistical and deep-learning approaches provided convergent evidence for the dynamic interplay between environmental and cognitive factors. These findings establish a robust quantitative foundation for the discussion in Chapter 5, where their theoretical implications, practical significance, and potential design applications are further elaborated.

Chapter 5 Discussion

5.1 Introduction

This chapter provides an in-depth interpretation of the statistical and modelling results presented in Chapter 4 and discusses their theoretical and practical significance. While the previous chapter established that both thermal environment and task difficulty exerted significant and interacting influences on thermal comfort, task performance, and stress level, the purpose of this chapter is to explain why these patterns occurred and what they imply for the understanding of human responses under combined thermal–cognitive conditions.

The discussion first elaborates on the behavioural and psychophysiological mechanisms underlying the observed effects of thermal and task factors, integrating existing theories of thermal comfort, cognitive workload, and stress adaptation. It then interprets how the developed machine-learning and deep-learning models captured these complex relationships, highlighting the role of multimodal data and hierarchical learning in predicting human thermal and psychological responses.

Subsequent sections extend the discussion to the theoretical implications, where the findings are related to and contrasted with prior thermal-comfort and human-performance models, and to the practical implications, focusing on applications in intelligent environmental design and human-centred thermal management. Finally, the chapter concludes with the limitations of the present work and potential directions for future research. Together, these discussions aim to position the current study within the broader context of environmental ergonomics and to clarify its contributions to understanding and modelling the interplay between thermal, cognitive, and affective factors.

5.2 Interpretation of experimental and statistical findings

The statistical analyses presented in Chapter 4 revealed clear interaction effects between thermal factors and task factors on subjects' subjective and objective responses. These findings confirm that thermal comfort, task performance, and stress do not respond to thermal or cognitive demands in isolation, but co-vary under the combined effects of both. The following subsections elaborate on the observed patterns for each indicator and discuss their possible psychophysiological explanations.

5.2.1 Thermal sensation (TSENV)

Thermal sensation consistently tended to shift toward neutrality as task difficulty increased. This pattern suggests a modulation of thermal sensitivity by mental engagement: when subjects were cognitively occupied, they perceived less deviation from the neutral thermal state, even though ambient conditions remained constant.

This attenuation of thermal awareness can be interpreted as a consequence of attentional resource allocation. According to the limited-capacity theory of attention[234], mental resources are shared among concurrent perceptual and cognitive processes. As task demands increase, a larger portion of attentional capacity is directed to information processing and response execution, leaving fewer resources available for continuous monitoring of bodily sensations such as temperature [31]. Consequently, subjects become less responsive to mild thermal variations, producing sensations closer to a neutral state.

Physiological mechanisms may also contribute to this effect. Mentally demanding tasks are known to induce sympathetic activation and mild increases in metabolic rate [163], which can elevate skin or core temperature slightly and therefore shift perceived thermal balance toward neutrality[235]. During complex cognitive activity, peripheral vasodilation or redistribution of blood flow may further alter thermoreceptor input [111]. The combination of top-down attentional suppression and bottom-up physiological adjustment likely underlies the observed reduction in thermal sensitivity.

Moreover, the shift toward neutrality was consistent across thermal conditions, suggesting that cognitive load acted as a normalizing factor in thermal perception. Under high task difficulty, subjects may rely less on external thermal cues and more on internal temperature constancy signals or cognitive heuristics about comfort [31]. This interpretation is in line with findings that focused cognitive engagement can buffer discomfort perception in both thermal and pain domains [12].

Taken together, these results indicate that higher mental workload modifies both perceptual and physiological thermoregulation, leading to a perceived reduction in thermal deviation. Such interactions highlight the importance of considering cognitive activity as an integral factor in thermal comfort research, especially in environments where sustained concentration or mental effort is required.

5.2.2 Thermal satisfaction (TSATV)

Thermal satisfaction increased noticeably with higher task difficulty, implying that subjects became more tolerant of the thermal environment when cognitive demand was elevated. Although ambient temperature remained unchanged, subjective comfort ratings improved under greater mental workload.

This phenomenon can be explained by a psychological adaptation and cognitive distraction mechanism. Under high task difficulty, individuals direct attention and motivation toward achieving task goals, temporarily down-weighting peripheral sensory inputs unrelated to task performance [236]. This cognitive prioritization reduces the salience of thermal discomfort signals, which are then interpreted as less disturbing. Such modulation aligns with the "distraction hypothesis" in environmental psychology, where higher cognitive engagement acts as a comfort buffer that mitigates negative affective appraisal of the environment [42].

In addition, high task engagement often triggers motivational and emotional regulation processes [237]. According to appraisal theory, an individual's evaluation of the environment depends on perceived controllability and goal relevance [87]. Subjects likely perceived the task as the dominant source of challenge and therefore reframed mild thermal deviations as acceptable or even negligible. This adaptive reframing contributes to higher satisfaction ratings despite constant physical conditions.

Physiologically, mental effort may induce subtle increases in metabolic activity and peripheral vasodilation[163], slightly improving thermal balance and producing sensations closer to comfort. Furthermore, a heightened state of alertness can increase thermal tolerance thresholds through sympathetic activation [31].

Collectively, these results suggest that cognitive engagement facilitates a form of contextual thermal adaptation, in which satisfaction depends not only on the environment but also on concurrent psychological states. This finding highlights the dynamic, bidirectional relationship between task involvement and environmental appraisal in determining perceived comfort.

5.2.3 Objective performance (OPI)

Objective performance was found to be highest under the neutral thermal condition, slightly lower in the cool environment, and lowest under warm conditions. Although not all pairwise differences were statistically significant, the general trend supports the inverted-U relationship between temperature and performance predicted by thermal ergonomics[17].

At the behavioural level, extreme temperatures, even within a mild range, can compromise cognitive efficiency by diverting attention toward thermal discomfort and initiating compensatory physiological responses such as sweating or shivering [111]. Under cool conditions, reduced peripheral temperature and muscular stiffness may affect motor coordination and fine control, while warm environments can accelerate fatigue and lower vigilance [238]. Neutral temperatures provide optimal thermal balance, facilitating sustained attention and psychomotor stability.

From a physiological perspective, deviations from neutrality alter autonomic regulation and metabolic demand [44]. Excessive heat elevates core temperature and heart rate, increasing perceived exertion and interfering with executive processing.

Conversely, cold exposure slows neural conduction and decreases dexterity [183]. These disruptions translate into minor but cumulative reductions in cognitive throughput.

The limited statistical significance among certain conditions suggests considerable individual variability in thermal sensitivity and adaptation[52], reflecting the influence of acclimatization, age, and physiological resilience. Nevertheless, the consistent trend across subjects underscores the general principle that moderate thermal neutrality supports optimal task performance by minimizing extraneous physiological or attentional load.

5.2.4 Subjective workload (NASA-TLX)

The NASA-TLX results indicated that perceived workload was lowest under neutral thermal conditions, particularly during the easy task. Under such circumstances, thermal and task demands were both minimal, allowing subjects to allocate cognitive resources effectively.

As thermal conditions deviated from neutrality, workload scores increased, most likely due to the addition of thermal strain as a secondary stressor[31]. When exposed to cool or warm environments, subjects must simultaneously manage thermoregulation and cognitive performance. This dual demand consumes mental resources and raises perceptions of effort and frustration. The observed pattern aligns with the multiple-resource theory, which posits that environmental discomfort competes with cognitive processing for limited attentional capacity[12].

However, at higher task difficulty levels, the effect of thermal variation on workload diminished, possibly due to a ceiling effect [31]. In such situations, intrinsic cognitive demand dominates the workload evaluation, overshadowing marginal contributions from physical discomfort. This interaction suggests an adaptive prioritization process, whereby individuals under heavy mental load predominantly appraise cognitive constraints rather than external environmental challenges.

Physiologically, enhanced sympathetic activation under heat or cold stress increases energy expenditure and cardiovascular load [89], which subjects may consciously or subconsciously interpret as greater effort. Together, these mechanisms demonstrate that perceived workload reflects the integrated influence of thermal and task factors, mediated through both cognitive appraisal and autonomic activation.

5.2.5 Stress level (STAI-S)

The STAI-S results exhibited pronounced and consistent effects of both thermal and task conditions, confirming that stress level is sensitive to dual sources of strain. Stress scores increased significantly with deviations from neutral temperature and with higher task difficulty, indicating additive and interactive influences [31].

This pattern is consistent with models of combined physical and psychological stress [239]. Thermal stress triggers sympathetic nervous system activation, elevating heart rate, skin conductance, and cortisol release [37]. Simultaneously, difficult cognitive tasks induce mental strain through sustained attention and error monitoring, also engaging stress pathways. When these two demands coexist, physiological arousal accumulates, yielding higher subjective stress levels and potentially reduced emotional regulation [111].

The fact that significant temperature differences disappeared only under the easy-task condition suggests that low cognitive demand allows sufficient coping capacity to maintain emotional stability, even when thermal conditions are suboptimal. Under moderate and difficult tasks, however, the combination of thermal discomfort and mental load surpasses individuals' adaptation thresholds, resulting in marked stress elevation.

From a psychological viewpoint, this outcome supports the interactional stress framework, in which total stress reflects the summation of environmental discomfort, perceived control, and task challenge [165]. The results imply that both thermal and cognitive domains contribute to the same underlying arousal system, producing comparable subjective manifestations such as tension and irritability [139].

Overall, the STAI-S findings reinforce the notion that thermal comfort, performance, and stress are tightly interlinked, and that managing environmental conditions can be an effective strategy for mitigating psychological stress in task-intensive settings.

5.2.6 Summary

In summary, the statistical analyses revealed a coherent pattern demonstrating that thermal and task factors jointly shaped subjects' perceptual, behavioural, and psychological responses. Variations in task difficulty consistently moderated how individuals experienced and reacted to different thermal environments, indicating that human thermal comfort and performance are dynamically constructed through the integration of environmental and mental demands.

Specifically, thermal sensation (TSENV) shifted toward neutrality with increasing task difficulty, suggesting that greater cognitive engagement attenuated sensory awareness of temperature deviations through attentional resource reallocation and mild physiological adjustments[31]. In parallel, thermal satisfaction (TSATV) improved under high workload conditions, reflecting psychological adaptation and motivational compensation that allowed subjects to tolerate or reinterpret thermal discomfort as acceptable[79].

Objective measures further supported these interactive effects. Objective performance (OPI) was optimal under neutral temperatures, while departures from neutrality slightly impaired efficiency, confirming the thermal-performance coupling principle [17]. Subjective workload (NASA-TLX) increased when thermal conditions deviated from comfort, especially under low and moderate difficulty, implying that thermal strain imposes additional cognitive load. However, the effect diminished at the highest difficulty level, where intrinsic mental demand dominated workload perception[31].

Finally, stress level (STAI-S) exhibited the strongest and most consistent dual effects: both thermal deviation and cognitive load independently and interactively elevated stress responses. This pattern illustrates the total arousal produced by simultaneous physical and psychological strain [163]. Together, the five indices delineate a multi-layered response system in which perception, satisfaction, performance, workload, and stress are interdependent facets of a unified thermal-cognitive adaptation process.

Overall, the findings of this section indicate that individuals under complex task conditions do not passively experience the environment but actively regulate their comfort and effort through both cognitive focus and physiological adjustment. These results from the empirical foundation for the subsequent modelling analyses in Section 5.3, where machine-learning and deep-learning approaches are employed to quantify and predict these intertwined relationships across modalities and response domains.

5.3 Interpretation of machine-learning and deep-Learning results

This section interprets the modelling results obtained from four successive frameworks, Model I to Model IV, each representing a step forward in analytical complexity and representational depth. Models I.1 to I.5 served as single-target benchmarks that predicted individual human-response variables (TSENV, TSATV, OPI, NASA-TLX, and STAI-S) from thermal, task, and personal factors using linear regression (LR), k-nearest neighbours (KNN), random forest (RF), and multilayer perceptron (MLP) algorithms. Subsequent models expanded this framework: Model II applied a multi-target MLP to capture interdependencies among the five responses; Model III further incorporated physiological indicators; and Model IV integrated thermal-image features via a CNN-MLP hybrid architecture.

Together, these models form a progressive hierarchy that moves from conventional statistical learning toward multimodal, deep representations of human thermoregulatory and cognitive adaptation. The following subsections discuss their comparative performance, the underlying mechanisms behind observed improvements, and the theoretical implications of modelling human responses as a complex, interconnected, and data-driven system.

5.3.1 Performance of single-target models (Model I.1 – I.5)

Models I.1 to I.5 were constructed to predict each human-response variable independently, thermal sensation (TSENV), thermal satisfaction (TSATV), objective performance (OPI), subjective workload (NASA-TLX), and stress level (STAI-S), based on combinations of thermal environment parameters, task difficulty levels, and personal factors. Four different algorithms were employed: linear regression (LR), k-nearest neighbours (KNN), random forest (RF), and multilayer perceptron (MLP).

Among these, TSENV and TSATV were formulated as classification targets, whereas OPI, NASA-TLX, and STAI-S were approached as regression targets.

Each target model followed a phased feature-addition procedure, designed to examine the incremental contribution of variable domains. For TSENV and TSATV, the initial stage employed only thermal environment and personal factors, followed by a second stage that introduced task-difficulty variables. In contrast, OPI and NASA-TLX initially used task and personal factors; thermal-environmental variables were then added in the subsequent phase to test cross-domain improvement. STAI-S was developed in three stages, first using thermal + personal, second using task + personal, and finally integrating thermal + task + personal features. This design enabled systematic evaluation of whether combining cognitive-contextual and environmental information improved predictive quality.

5.3.1.1 Algorithmic performance

The comparative performance clearly demonstrated task-type-dependent algorithm superiority.

For classification targets (TSENV and TSATV), the ranking of model performance was:

MLP > LR > RF > KNN

The MLP models consistently achieved the highest accuracy and F1-scores, benefiting from their ability to approximate non-linear boundaries among discrete perceptual categories. Linear regression, though less sophisticated, still performed reasonably well, partly because thermal-sensation categories exhibited quasi-linear transitions within moderate temperature ranges[240]. Random forest showed slightly lower accuracy, likely due to its threshold-based partitioning that struggled with smooth perceptual gradients. KNN performed worst, indicating that local proximity in the feature space did not correspond well with categorical perception.

For regression targets (OPI, NASA-TLX, STAI-S), the ranking reversed:

RF > KNN > LR > MLP

Here, RF provided the most stable and accurate predictions, as ensemble averaging reduced overfitting and better captured non-linear dependencies among features. KNN performed comparably, benefiting from pattern similarity among neighbouring subjects. LR achieved moderately acceptable results, capturing main trends but missing cross-interactions. MLP was least consistent, particularly in limited-sample settings, where excessive parameterization hindered convergence and led to partial overfitting [241].

5.3.1.2 Influence of feature-set expansion

Performance consistently improved across all algorithms when more comprehensive feature sets were included.

For TSENV and TSATV, adding task-difficulty levels to the thermal + personal feature set led to notable gains in accuracy, especially for MLP and LR models. This finding implied that cognitive engagement moderates thermal perception, and algorithms leveraging task information could capture this adjustment more effectively. The improvement was most pronounced in MLP, where additional task-difficulty nodes enriched the hidden-layer representation of interaction effects between cognition and environment[162].

For OPI and NASA-TLX, which began with task + personal factors, the introduction of thermal variables produced meaningful increases in R^2 and reductions in RMSE, particularly in the RF and KNN models. These improvements confirmed that environmental parameters influence not only subjective comfort but also objective and perceived performance outcomes. The result illustrates a reciprocal feedback mechanism, environmental discomfort adds secondary strain even in cognitively dominant tasks[31].

The STAI-S models, developed in three stages, exhibited the clearest incremental improvements among all targets. Predictive accuracy was moderate when using either thermal + personal or task + personal features separately but increased markedly when both domains were combined. This pattern confirmed that thermal and cognitive stressors act additively in shaping emotional stress levels. In particular, RF demonstrated an almost linear improvement across phases, reinforcing its sensitivity to multi-domain feature interactions [33].

Taken together, the feature-addition experiments demonstrated a robust trend: as feature representations became more holistic, integrating environmental, cognitive, and personal inputs, model performance consistently improved across all targets and algorithms.

5.3.1.3 Target-specific observations

Thermal Sensation (TSENV): MLP achieved the best categorical prediction, accurately identifying neutralization shifts as task difficulty increased. The inclusion of task features enhanced discrimination between "neutral" and "slightly warm" states that purely thermal models could not separate.

Thermal Satisfaction (TSATV): Likewise, MLP led performance, while LR also captured general tendencies. Both benefited from the addition of task variables, confirming that satisfaction depends on concurrent cognitive context.

Objective Performance (OPI): RF's superior generalization identified performance decline under heat and cold, while KNN produced comparable results at moderate temperatures. Adding thermal features improved explanatory capacity on top of task information.

Subjective Workload (NASA-TLX): The inclusion of thermal variables yielded reduced prediction errors for RF and KNN, underscoring that thermal stress elevates cognitive workload perception.

Stress Level (STAI-S): The multi-phase comparison highlighted the necessity of integrating both environmental and task factors to approach true stress variability; RF again provided the most accurate estimates.

Each target thus displayed algorithm-specific advantages tied to the intrinsic structure of its outcome (categorical vs continuous, cognitive vs physiological sensitivity).

5.3.1.4 Interpretive insights and limitations

Across targets, two principal methodological insights emerge.

First, algorithmic suitability depends on the nature of the response variable. MLP's deep, non-linear capacity is advantageous for categorical perception requiring soft-boundary mapping, while RF and KNN succeed in numeric responses where local continuity predominates.

Second, multidomain feature inclusion—combining thermal, task, and personal information—proves universally beneficial regardless of algorithm, suggesting that perceived comfort, performance, and stress are outputs of an integrative human-environment system rather than isolated processes [33].

The limitation of this model family lies in its single-target structure. Each target is learned independently, thus omitting the inherent correlations among human responses (e.g., thermal satisfaction correlating with stress, or workload influencing performance). Consequently, while prediction accuracy improved with enriched features, these models remained restricted to separate outcome domains. Recognizing this gap motivated the design of Model II, which applies a multi-target MLP framework to capture shared representations and cross-response interdependencies between perceptual, behavioural, and psychological dimensions.

5.3.2 Performance of multi-target model (Model II)

Model II represented the first attempt to integrate all five human-response variables—TSENV, TSATV, OPI, NASA-TLX, and STAI-S—within a single predictive framework. It adopted a multilayer perceptron (MLP) architecture capable of multi-output learning, aiming to capture potential interdependencies among perceptual, behavioural, and psychological dimensions that were treated separately in Model I. The input variables included the most comprehensive combination of thermal environment, task difficulty, and personal factors identified from previous single-target analyses.

5.3.2.1 Model performance

As shown in Table 4.13, Model II's overall performance was markedly weaker than that of the single-target models (Model I.1 – I.5) across all targets and metrics. For classification tasks, the accuracies of TSENV (0.641) and TSATV (0.369) were notably lower than those of their corresponding single-target versions (0.736 and 0.460, respectively). Similarly, the regression tasks presented sharp declines in determination coefficients, such as OPI decreasing from 0.431 to 0.363, NASA-TLX from 0.754 to 0.278, and STAI-S from 0.710 to 0.228. In practical terms, Model II struggled to generalise across all five outputs simultaneously, and the residual distributions showed higher unexplained variance than in the previous models.

These findings indicate that, at this developmental stage, the cost of unifying multiple prediction targets outweighed the potential benefit of shared-representation learning. The complexity introduced by training multiple correlated outputs within a limited dataset caused instability in weight updates and, in many cases, convergence to sub-optimal minima—a well-recognised limitation of multi-task learning when data diversity and scale are insufficient [242]. The model oscillated between different loss gradients generated by heterogeneous target structures, which impaired the stabilisation of shared layers and weakened the network's ability to extract domain-specific information for each response [66]. Consequently, although the integrated framework symbolically embodied an advance toward holistic prediction, it produced fragmented and diluted representations that lacked precision across both categorical and continuous measures.

5.3.2.2 Reasons for underperformance

The underperformance of Model II can be explained by a combination of methodological and empirical factors. One major reason lies in parameter interference and gradient competition among tasks. In a multi-output network, each target produces its own gradient signal, and these signals can diverge greatly in magnitude or direction. When classification and regression objectives coexist, their imbalance often results in unstable training dynamics and mutually destructive gradient interactions [243]. This phenomenon hindered shared layers from converging toward coherent latent representations.

Another key limitation concerns the heterogeneity of output structures. The five predicted responses differ substantially in distributional forms and noise levels—TSATV is a skewed categorical variable, while OPI is continuous with wider variance. Without task-specific weighting or balancing mechanisms, the MLP treated all outputs equally; in practice, dominant but noisier targets overshadowed subtler ones, distorting weight updates through shared layers [162]. At the same time, the available dataset, though adequate for the individual models, was too limited for a deep network with an expanded parameter space. This data insufficiency led to underfitting, weak pattern generalisation, and degraded accuracy, consistent with known sensitivity of deep networks to data-scale–to-parameter-size ratios [162].

A further issue stems from the limited alignment between features and targets. The model integrated environmental and behavioural features but excluded physiological indicators that capture internal affective or stress states. Many psychological responses—especially STAI-S and NASA-TLX—are closely tied to physiological activation, yet this information was unavailable to the model. As a result, Model II could represent general directional tendencies, such as decreased satisfaction and increased stress with higher task load, but lacked the internal cues required for precision in prediction. Overall, these factors produced a diluted rather than synergistic representation, indicating that Model II served more as a methodological turning point than as a performance improvement.

5.3.2.3 Significance within the modelling framework

Despite its weak quantitative outcomes, Model II plays an important conceptual role in the evolution of this modelling framework. It introduced the foundational notion that thermal comfort, cognitive performance, workload, and stress are not separate phenomena but interconnected components of a unified adaptive process. Attempting to model them simultaneously transformed this conceptual assumption into a computational hypothesis, providing the groundwork for subsequent refinement. Through its deficiencies, Model II also offered diagnostic insights: the necessity of adaptive loss weighting to mitigate gradient interference, the value of expanding feature diversity through physiological and perceptual signals, and the recognition that successful deep architectures require sufficient data relative to model capacity [66].

Although Model II did not surpass its predecessors in numerical accuracy, it clarified why purely environmental-behavioural datasets cannot capture the complex interplay between thermal conditions, cognitive workloads, and emotional states. The insights gained from its shortcomings directly informed the design of Models III and IV, where multimodal inputs and improved fusion strategies were introduced to restore coherence and predictive strength. Model II thus stands as a necessary intermediate stage—a demonstration that multi-target prediction is conceptually feasible but computationally demanding under small-sample, heterogeneous human data environments. Its outcomes emphasise that genuine synergy among comfort, performance, and mental-state prediction demands not only shared architectures but also supplementary information sources and more controlled optimisation, directions systematically pursued in subsequent phases of this study.

5.3.3 Integration of physiological indicators (Model III)

Model III was developed following the diagnostic outcomes of Model II and represents the first stage in which physiological information was incorporated to complement environmental, task, and personal features. The integration of physiological indicators was intended to overcome the feature–target disconnect observed previously, namely the restricted capacity of behavioural and environmental variables to capture the internal states of emotional arousal and cognitive stress. Accordingly, physiological measures such as heart rate, skin temperature, and peripheral blood-flow changes were added to the input domain to provide proxies for autonomic activation and thermoregulatory responses. The model retained the multi-output multilayer perceptron (MLP) architecture, allowing it to predict the same five human-response targets—TSENV, TSATV, OPI, NASA-TLX, and STAI-S—while now receiving a multidimensional set of external and internal features.

5.3.3.1 Model performance

As summarised in Table 4.13, the inclusion of physiological variables led to a modest but consistent improvement across all tasks compared with Model II. For the classification targets, TSENV accuracy increased from 0.641 to 0.670 and TSATV from 0.369 to 0.428; corresponding F1-scores also rose from 0.305 to 0.333 and from 0.136 to 0.208. In regression tasks, the R^2 values of OPI, NASA-TLX, and STAI-S improved from 0.363 to 0.397, 0.278 to 0.305, and 0.228 to 0.270, respectively, accompanied by lower root-mean-square errors. The quantitative enhancement remained moderate but demonstrated that the introduction of physiological signals partially restored generalisation ability and reduced output variance across targets.

This improvement suggests that physiological indicators contain latent information that bridges environmental stimuli and subjective experience [37]. By providing immediate reflections of thermal strain and emotional activation, these measures helped the network interpret context-dependent psychological fluctuations that could not be inferred purely from ambient or behavioural features. Although the overall figures were still lower than the best single-target results, the trend confirmed that adding embodied data improved the consistency and coherence of predictions across perceptual and affective dimensions.

5.3.3.2 Interpretation of improvements

The positive shift in Model III performance underscores the role of physiological activity as an intermediary layer between physical environments and human experience [244]. Thermal stress and cognitive load both trigger changes in autonomic responses, such as heart-rate acceleration and peripheral vasoconstriction, which in turn correlate with subjective discomfort and fatigue. The inclusion of these parameters thus enabled the network to capture part of this causal chain linking external thermal stressors to internal physiological perturbations and onward to perceived comfort, workload, and emotional strain [31].

From an algorithmic perspective, the mixed feature set improved the separability and clustering of samples in latent space. Physiological features contributed continuous, smoothly varying signals that balanced the scale discrepancies between categorical and numerical targets. During training, these stable intermediate variables acted as regularisers for the MLP, mitigating the uncontrolled oscillations of gradient updates reported in Model II. This stabilised optimisation trajectory led to more coherent shared-layer learning and reduced the trade-off between classification and regression objectives [66].

At the same time, limitations persisted. The gains achieved were moderate, indicating that although physiological signals provided meaningful supplementary information, their standalone inclusion was not sufficient to produce substantial predictive leaps. Noise sensitivity, sensor variability, and the modest sample size continued to constrain the model's learning dynamics. Furthermore, temporal alignment between physiological responses and instantaneous perception remained imperfect; transient physiological changes may lag or precede subjective reports, adding further complexity [17, 245]. These factors collectively explain why Model III's accuracy improved relative to Model II but still did not match the single-target or final multimodal models.

5.3.3.3 Significance within the modelling framework

Despite its modest numerical performance, Model III represents a fundamental methodological advancement in the modelling framework. It operationalised the conceptual inference that human comfort and performance are jointly modulated by external stimuli and internal physiological regulation[17]. By introducing bodily indicators, the model moved beyond surface correlations toward a layered representation of human adaptation that spans environmental, physiological, and psychological domains. From a modelling standpoint, the results highlight how physiological measures provide biologically grounded features that can stabilise learning and enrich feature interpretability. For instance, the model's learned weights emphasised the contribution of skin-temperature deviations to TSENV and TSATV, while heart-rate variability and peripheral temperature dynamics emerged as shared predictors for NASA-TLX and STAI-S. Such patterns illustrate the inherent coupling between physiological regulation and subjective cognitive–emotional experience [37].

Conceptually, Model III bridges the gap between environmental ergonomics and psychophysiology. It demonstrates that integrating physiological data enhances the explanatory and predictive depth of human-response models, establishing a foundation for genuinely multimodal analysis. Although the prediction accuracy did not substantially exceed single-target baselines, the improved internal coherence of outputs suggested that the model began to represent co-fluctuations among comfort, workload, and stress more faithfully. These findings further justified the continuation toward Model IV, which expands the multimodal approach and incorporates optimised fusion mechanisms to fully exploit cross-domain information.

In summary, Model III served as a pivotal transitional step from behaviourally driven to physiologically augmented modelling. It confirmed that physiological indicators meaningfully contribute to capturing thermal–cognitive interactions by revealing the implicit bridge between external heat stress and internal emotional regulation [37]. While its quantitative advancements were limited, its methodological and conceptual contributions were substantial: Model III validated the feasibility and necessity of multimodal integration, paving the way for the final unified framework that characterises the human adaptive process with comprehensive environmental, physiological, and psychological dimensions.

5.3.4 Multimodal fusion and thermal-image–based unified model (Model IV)

Model IV represents the final stage of this research and the most comprehensive framework constructed to integrate multiple data modalities into a unified predictive system. Building on the findings from Model III, this model introduced thermal images as an additional modality to capture fine-grained spatial information about subjects' facial and upper-body surface temperatures. A CNN–MLP hybrid architecture was implemented, where the convolutional neural network (CNN) extracted spatial–thermal representations from infrared images, and the MLP processed non-visual features including environmental, task-related, personal, and physiological parameters. The latent features from the two branches were subsequently fused through a joint fully connected layer for multi-output prediction of the five targets—TSENV, TSATV, OPI, NASA-TLX, and STAI-S. This design aimed to achieve a multimodal interaction mechanism that captures both the external thermal context and internal physiological–psychological states in a unified deep-learning framework [31].

5.3.4.1 Model performance

The inclusion of thermal-image data and CNN–MLP fusion produced a dramatic performance gain over all previous models. As summarised in Table 4.13, the classification accuracy for TSENV rose to 0.745 and for TSATV to 0.617, with corresponding F1-scores of 0.620 and 0.536, nearly doubling those of Model III and surpassing the best single-target results. The regression tasks also achieved substantial improvements: OPI reached $R^2 = 0.604$ with RMSE = 6.511, NASA-TLX achieved $R^2 = 0.844$, and RMSE = 8.122, STAI-S attained $R^2 = 0.779$ with RMSE = 5.308. These outcomes indicate that Model IV successfully generalised across both categorical and continuous targets, achieving balanced and stable performance. Training behaviour showed reduced oscillation and smaller gaps between training and validation metrics, reflecting better feature representation and improved optimisation stability [66].

This notable advancement highlights the power of visual–non-visual fusion. Adding thermal images introduced spatial-temperature patterns that embodied moment-to-moment thermal responses of subjects' skin surfaces. Compared to scalar environmental or physiological variables, thermal images preserved contextual cues such as facial heat distribution, asymmetry, and local gradients, all of which correlate with affective and cognitive states [246]. The CNN branch effectively extracted these patterns, providing compact latent image embedding that enriched the subsequent fusion layers of the MLP. By merging these visual embeddings with numerical features derived from ambient and physiological signals, Model IV captured both the external stimulus field and internal bodily response in a cohesive predictive space.

5.3.4.2 Interpretation of results

The superiority of Model IV stems from the complementarity between spatial thermal vision and numerical physiological–environmental data. CNNs are particularly adept at detecting complex, non-linear textures and temperature distributions in thermal imagery [247]. Such patterns often reflect underlying vasomotor and metabolic regulation processes associated with stress, vigilance, or heat discomfort [248]. When these representations were combined with MLP-encoded features—task difficulty, ambient temperatures, personal factors, and physiological signals—the hybrid network learned multi-layered relationships among perception, cognition, and physiology. This multimodal synergy enhanced feature expressiveness and reduced redundant noise across modalities. In effect, the model simultaneously processed the where of thermal response (spatial localisation on the face and body) and the why of systemic regulation (internal physiological or contextual factors).

From an optimisation perspective, the model benefitted from fusion-layer regularisation. The intermediate fusion stage equalised gradient magnitudes across modalities, preventing the unbalanced learning that hindered Model II and stabilising convergence[66]. Thermal images delivered dense, high-variance input that enriched gradient information, while numerical modalities anchored the global trend and constrained overfitting. This interaction produced smoother loss trajectories and higher robustness against outliers. Furthermore, CNN-derived visual embeddings improved latent-space clustering of samples belonging to similar perceptual states (e.g., "neutral comfort" vs "slightly warm discomfort"), allowing more consistent categorical distinctions and continuous estimations.

Nevertheless, some limitations remain. Although overall accuracy improved substantially, computational cost increased due to the CNN processing of image data, requiring larger storage and longer training times. Small variations in camera calibration, face alignment, and emissivity may introduce localisation noise in thermal images, potentially affecting generalisation [39]. Despite these constraints, the advantages of visual–numeric integration outweighed the drawbacks, yielding the most coherent and accurate model within this framework.

5.3.4.3 Significance within the modelling framework

Model IV marks the culmination of the modelling hierarchy, demonstrating that multimodal CNN–MLP architectures provide a decisive step forward in predicting human comfort, performance, and emotional responses. By fusing thermal imagery with environmental, physiological, and task data, the model operationalises the concept of human adaptation as a visually and physiologically coupled process [66]. This establishes a methodological bridge between environmental ergonomics, computer vision, and psychophysiology. The integrated architecture aligns with cognitive-energetic and embodied-cognition theories, which posit that perceptual comfort and cognitive load are expressed through both measurable bodily cues and observable thermal dynamics [31].

In practical terms, Model IV demonstrates how multimodal sensing can transform predictive ergonomics. The results suggest that thermal imaging, when used in conjunction with physiological and contextual information, can serve as a real-time proxy for both physical and mental strain, offering scalable solutions for intelligent built environments and adaptive human–machine interfaces [66]. By embedding CNN-based vision modules into the predictive pipeline, the model achieves not only higher statistical performance but also richer interpretive potential—the latent thermal patterns provide intuitive visual evidence linking observable physiological phenomena to subjective and behavioural outcomes.

Ultimately, Model IV validates the feasibility of sophisticated multimodal fusion in modelling human thermal–cognitive–affective adaptation. It resolves the deficiencies of Model II and Model III by incorporating complementary visual information that anchors the latent representation space in physiologically meaningful thermal structures. This unified CNN–MLP framework therefore represents the most accurate, stable, and theoretically interpretable model within the present research, forming a robust foundation for subsequent applications in intelligent environmental design, thermal comfort personalisation, and multimodal human-state monitoring [249].

5.4 Theoretical implications

The progressive modelling framework developed in this research—from single-target prediction to multimodal, image-based integration—provides substantial theoretical insights into the nature of human thermal experience. Collectively, the results indicate that thermal comfort, cognitive performance, workload, and affective stress are not independent outcomes of separate physiological or environmental systems, but manifestations of a unified adaptive process continuously negotiating between thermal, cognitive, and emotional demands [210]. This perspective moves beyond the notion of thermal comfort as a static end-state toward an understanding of it as a dynamic, context-sensitive regulation process embedded in everyday activity.

5.4.1 Integration of cognitive factors into thermal comfort theory

The findings enrich and expand the explanatory range of traditional thermal comfort models such as the PMV and adaptive comfort frameworks [250]. While these theories effectively describe population-level thermal sensation as a function of environmental conditions and insulation, they largely treat the individual as a passive subject driven by heat balance. The results from Models I to III demonstrate that perception of comfort fluctuates systematically with cognitive workload and task context, even when physical thermal parameters remain stable. When task difficulty increases, both thermal satisfaction and self-reported workload exhibit significant change, revealing a cognitive modulation of comfort perception. This suggests that comfort does not simply represent equilibrium between metabolic heat production and environmental dissipation but involves a higher-order process of cognitive evaluation and attentional prioritisation [250].

Through the inclusion of task-related features, the models showed that subjects under demanding cognitive conditions tolerate a broader or shifted range of ambient temperatures without strong discomfort signals. This observation introduces a cognitive component to adaptive comfort theory: adaptation is not merely behavioural or physiological but also attentional, driven by the need to maintain task focus rather than perfect thermal neutrality [251]. Hence, thermal comfort emerges as the outcome of a situational trade-off between physical state regulation and cognitive goal maintenance.

5.4.2 Linking physiological regulation and psychological states

The transition from purely behavioural-environmental modelling in Model II to the physiologically enriched Model III clarified that internal bodily responses carry substantial explanatory weight in thermal-cognitive interaction. Physiological indicators such as heart rate or skin temperature served as measurable expressions of autonomic regulation, bridging external stimuli with internal experience [44]. Their successful incorporation and subsequent performance improvement suggest that psychophysiological processes are not peripheral but central to understanding comfort and stress. Physiological reactivity reflects both environmental challenge and emotional significance, locating thermal comfort at the intersection of homeostatic control and affective appraisal [250]. This finding integrates thermoregulatory physiology with cognitive-affective theory, supporting the view that perceived comfort is simultaneously a bodily and psychological phenomenon.

In addition, the relationship between thermal and emotional variables observed in Models III and IV illustrates that stress and comfort respond in opposite directions but share overlapping physiological pathways. Elevated cardiovascular or vasomotor activity under thermal load corresponds to reduced satisfaction and increased perceived workload. This pattern aligns with the broader theoretical claim that environmental stressors exert their effect not directly on cognition but via physiological arousal modulation [37]. Thus, thermal comfort theory must be extended to incorporate an affective-physiological dimension that accounts for both sensory evaluation and bodily state regulation.

5.4.3 Multimodal integration and the embodied nature of thermal experience

Model IV provided direct evidence that the integration of multimodal cues—particularly thermal imagery—enhances the representation of human thermal experience. The use of a CNN–MLP hybrid structure allowed the model to extract spatial temperature patterns from facial and body surfaces and link them to concurrent physiological and cognitive metrics. From a theoretical standpoint, this architecture realises a computational analogue of embodied perception: the system ceases to rely solely on abstract numerical variables, instead encoding visible signs of thermoregulatory behaviour[30]. Such image-based features correspond conceptually to bodily feedback loops in humans, where thermal sensations arise from both core temperature signals and cutaneous sensory information. By successfully learning from surface temperature distributions, the model provides empirical confirmation that embodied thermal expression can serve as a valid representation of internal comfort states [39].

Furthermore, the superior performance of Model IV demonstrates that complex human responses are best explained when environmental, physiological, and visual modalities are considered simultaneously. This finding supports a systems-theoretical perspective in which perception results from multiple, interacting feedback channels operating across physical, cognitive, and affective levels [12]. Incorporating visual information also refines theoretical understanding of self-regulation mechanisms: the visible thermal field captures both environmental exposure and individual adaptation, reinforcing the concept of comfort as an emergent property of coupled body-environment dynamics rather than a unidirectional response.

5.4.4 A framework for thermal–cognitive interaction

The modelling sequence collectively points to a comprehensive theoretical synthesis that can be described as a thermal–cognitive interaction framework. Within this framework, environmental heat exchange, cognitive workload, and affective-physiological regulation constitute interrelated subsystems whose continuous interactions determine perceptual and behavioural outcomes. Thermal comfort, cognitive efficiency, and emotional stability are therefore co-dependent results of the same adaptive control loop [250]. External thermal stress imposes additional physiological demand; cognitive activity competes for metabolic and attentional resources; emotional evaluation guides protective or compensatory responses. The balance among these processes defines moment-to-moment thermal experience and performance capability.

This integrated view extends classical comfort theory by embedding it within the cognitive–energetic tradition of ergonomics, where human performance is modelled as the dynamic allocation of limited resources under concurrent physical and mental demands[163]. It also converges with environmental psychology perspectives that emphasise meaning, expectation, and appraisal in shaping comfort perception. The models developed here provide an empirical pathway to quantify such interactions through interpretable, multimodal learning systems. Computationally, the evolution from single-target regression to multimodal CNN–MLP fusion mirrors the conceptual transition from additive explanations to interactive, system-level reasoning.

5.4.5 Toward a human-centred thermal cognition model

Ultimately, the theoretical implications of this research affirm the need for a human-centred approach to thermal comfort and performance modelling. The empirical evidence shows that subjective well-being in thermal environments is sustained through an ongoing negotiation among environmental constraints, cognitive task demands, and physiological-affective responses. This dynamic equilibrium continuously adjusts over time and across contexts, reflecting the individual's adaptive regulation capacity[252]. Thermal comfort, in this sense, can be understood as a temporally and contextually situated process of maintaining functional and psychological stability rather than a fixed temperature preference.

By grounding this interpretation in measurable multimodal data, the present study offers a bridge between classical physical models and contemporary theories of embodied cognition and affect regulation. It suggests that future thermal comfort research should incorporate multidimensional data streams, recognising that the human body and mind operate as a single coherent system when responding to environmental stress. Such a framework redefines comfort not as a static physical outcome but as an emergent property of human–environment interaction that sustains both functional performance and subjective well-being [165].

5.5 Practical implications

The progressive modelling results obtained in this research carry significant implications for the design and management of human-centred environments, particularly workplaces and intelligent building systems. By incorporating cognitive load, physiological feedback, and multimodal sensing into thermal comfort prediction, the study offers evidence-based guidance for improving both comfort and performance outcomes in real-world applications. The following discussion examines how the findings can inform environmental design, adaptive control strategies, productivity enhancement, and health management in thermally dynamic settings.

5.5.1 Implications for work environment design

One of the clearest practical insights is that thermal comfort thresholds should not be treated as universally fixed but context-dependent, varying with cognitive and task demands [30]. The models demonstrated that individuals engaged in high workload conditions tolerate a wider thermal range without substantial losses in comfort ratings, while under passive or monotonous conditions smaller temperature deviations trigger noticeable dissatisfaction. Consequently, in cognitively intensive environments such as control rooms, design offices, or research laboratories, thermal setpoints can be kept slightly broader without impairing performance, thereby reducing energy consumption. Conversely, in low-activity or routine contexts, narrower temperature bands remain advisable because occupants are more perceptive of minor deviations. This evidence supports the shift toward activity-adaptive thermal management rather than static comfort standards.

The outcome also emphasises that environmental design should simultaneously consider thermal and cognitive ergonomics. Lighting, noise, and spatial layouts interact with thermal perception, influencing the cumulative cognitive load experienced by occupants [33]. Integrating the findings from this multimodal analysis into workplace planning enables holistic optimisation—maintaining moderate environmental stimuli across sensory channels to preserve attention and reduce fatigue. Design guidelines should therefore prioritise a multi-sensory balance, acknowledging that physical and psychological demands are inseparable determinants of comfort and productivity.

5.5.2 Implications for intelligent building control and personalised adaptation

The multimodal fusion strategy developed in Model IV demonstrates how sensor integration in intelligent buildings can move beyond physical monitoring toward human-state estimation[66]. By combining environmental sensors with physiological and thermal-imaging inputs, control systems can dynamically adjust temperature and airflow according to occupants' real-time comfort or stress levels. This approach allows closed-loop adaptive control: the system interprets physiological or visual cues—such as elevated facial temperature or increased heart rate—as indicators of thermal or cognitive strain and automatically moderates environmental parameters. Such responsiveness enables optimisation of both comfort and energy efficiency without explicit occupant intervention.

Furthermore, the CNN–MLP framework shows the feasibility of developing personalised thermal-comfort models that learn from individual responses rather than population averages [66]. With sufficient calibration data, adaptive algorithms could predict an individual's preferred microclimate based on daily workload, emotional state, or physiological feedback. In shared environments, aggregated probability maps of user comfort can inform zoning strategies—thermal clustering that balances collective satisfaction against energy use. By translating predictive modelling into intelligent control logic, the built environment transforms from a static container into an interactive system responsive to human states.

5.5.3 Implications for productivity and performance management

The coupling between comfort, workload, and stress revealed by the models provides practical insights for productivity management. Results showed that extreme thermal conditions elevate stress and workload while diminishing performance, whereas mild variations can sometimes maintain or even enhance alertness under demanding tasks [31]. This nonlinear relationship suggests that workplace regulation should aim for a functional comfort zone rather than a uniform thermal neutrality. Within this zone, mild thermal divergence serves as beneficial stimulation that sustains attention without imposing physiological strain. Managers and designers can leverage this knowledge to schedule tasks or adjust environmental parameters according to job complexity: cognitively challenging activities may benefit from slightly cooler or warmer conditions that prevent fatigue, whereas routine tasks require more neutral settings to maximise comfort stability.

At an organisational level, thermal data combined with task and performance metrics could contribute to behaviour-aware analytics for workload management. Integrating such predictive indicators into occupational dashboards allows dynamic balancing between environmental comfort and performance targets [66]. For example, real-time prediction of elevated workload or stress could trigger environmental adjustments or break recommendations, preventing cognitive burnout in continuous high-load scenarios such as monitoring operations or simulated training.

5.5.4 Implications for health and psychological well-being

Beyond performance and energy considerations, the physiological integration introduced in Models III and IV also holds significance for mental and physical health management. Persistent thermal strain has been associated with cardiovascular and cognitive stress; the present findings confirm that subjective workload and affective tension rise with cumulative heat or cold exposure [31]. The multimodal modelling framework offers potential for early detection of stress accumulation, utilising physiological and thermal-imaging indicators as unobtrusive markers of discomfort or fatigue. Such systems could inform preventive interventions—adjusting local environment, prompting rest, or recommending hydration—before discomfort manifests as health complaints. In high-risk or precision-critical environments, this monitoring could directly improve safety by maintaining operators within optimal psychophysiological limits.

By treating comfort as a health-linked variable rather than a secondary convenience factor, thermal-environment control can be aligned with occupational-well-being strategies [210]. This integration resonates with contemporary human-centred design paradigms that emphasise psychological sustainability alongside physical ergonomics. The models developed here provide a data-driven foundation for designing environments that actively support both the physiological resilience and emotional stability of occupants.

5.5.5 Implications for multimodal monitoring and future systems

The successful application of multimodal learning, particularly the CNN–MLP thermal-image model, illustrates the potential of multi-sensor fusion systems in real-time comfort assessment [253]. Thermal cameras, combined with standard environmental and physiological sensors, enable contactless, continuous observation of user states. When embedded into IoT-enabled building management platforms, these technologies can provide adaptive responses at the individual or group level—automatically adjusting ventilation, lighting, or scheduling environmental changes. Beyond buildings, such frameworks could extend to transport cabins, healthcare facilities, or human–machine cooperative systems, enabling dynamic comfort regulation coupled with operational performance.

Importantly, the multimodal principle also supports ethical and privacy-aware system design. By relying on indirect physiological cues rather than intrusive measurements, thermal-imaging-based systems can estimate wellbeing without requiring explicit self-report[253]. This preserves user autonomy while delivering beneficial environmental adjustments. Thus, the modelling framework demonstrates how advanced sensing and computation can serve human comfort through an empathetic technological interface.

Overall, the practical implications of this research converge toward a paradigm in which thermal comfort management is integrated with cognitive performance and wellbeing, forming the basis of adaptive, respectful, and sustainable environmental design. By revealing the mechanisms linking thermal perception, physiological state, and task demand, the models developed here provide both methodological and conceptual tools for next-generation building intelligence and human-centred workplace optimisation [31].

5.6 Limitations and future work

5.6.1 Limitations

While this research provides theoretical and methodological progress in modelling human thermal–cognitive–affective adaptation, several limitations remain that delineate its scope and guide future work. These concern the controlled laboratory conditions, sample scale, measurement precision, model interpretability, and broader applicability of the findings.

5.6.1.1 Experimental scope and task specificity

The present study was conducted in a controlled office environment, where temperature, lighting, and noise were kept stable to isolate thermal effects. Such control ensured data reliability but inevitably reduced ecological diversity. Subjects performed a single, relatively uniform computer-based task, limiting the representation of varied occupational or interactive scenarios. In real workplaces, tasks differ in physical demand, collaboration intensity, and attentional distribution. Hence, the models derived here may not fully capture the complexity of dynamic work environments [30]. Future research should extend analyses to multi-task or team-based settings, where cognitive and interpersonal factors interact with environmental stimuli in richer ways.

5.6.1.2 Subject and temporal scope

The dataset consisted of a moderate number of subjects with comparable demographic backgrounds and acclimatisation levels. This homogeneity restricts cross-cultural and inter-individual generalisability [183]. Furthermore, each experiment represented a short exposure; long-term acclimation or fatigue accumulation could not be observed. Follow-up studies should include larger, demographically diverse samples and longitudinal tracking, allowing inter-day and circadian fluctuations in comfort and stress to be examined.

5.6.1.3 Measurement and modal limitations

Although multimodal data enhanced model informativeness, the sensors and cameras used have resolution and synchronisation constraints. Thermal imagery captured only the upper body, while physiological signals such as heart rate or skin temperature were occasionally vulnerable to motion artefacts. Temporal mismatch between modalities may have weakened correlations among them [38]. Solving these issues will require higher-precision or wearable imaging devices, improved signal fusion pipelines, and real-time synchronised acquisition systems capable of capturing subtle transitional states in occupant responses.

5.6.1.4 Model complexity and data efficiency

The CNN–MLP architecture employed was effective but computationally demanding. Given the modest dataset, there remains a risk of overfitting or bias toward experimental conditions[66]. Future work should emphasise data-efficient deep-learning methods, such as transfer learning, transformer-based feature sharing, or self-supervised pre-training on larger open datasets. At the same time, the "black-box" nature of deep models limits interpretability. Introducing explainable-AI techniques—layer-wise relevance propagation, SHAP values, or Grad-CAM heatmaps—could reveal how features contribute to predictions and support responsible deployment [254].

5.6.1.5 Privacy and ethical concerns

Thermal imaging and physiological monitoring are inherently human-centric and raise privacy considerations. Even though the captured data are non-visual and anonymised, thermal patterns may still contain identifiable signatures[201]. Future implementation should include informed-consent procedures, strict data governance, and real-time anonymisation algorithms. Adaptive environments built on these models should remain user-in-the-loop, providing individuals with transparency and control over personalised adjustments to maintain trust and acceptance.

5.6.2 Future directions

Building upon these limitations, several explicit research directions can advance both scientific understanding and practical application.

Cross-environment and cross-population validation will be a key priority. Testing the models across diverse building types—open-plan offices, classrooms, healthcare facilities, and industrial settings—can assess generalisability beyond the controlled laboratory context. Including subjects from different climatic zones, age groups, and cultural backgrounds will help determine how social and physiological diversity modulates thermal–cognitive adaptation [255]. Such large-scale validation could be supported by establishing an open multimodal database that consolidates environmental, physiological, and thermal-image data from multiple research sites.

A second avenue involves temporal modelling and adaptation dynamics. Real-world comfort is not static but evolves through feedback and habituation. Employing sequential deep-learning architectures—such as recurrent neural networks, gated recurrent units, or vision transformers—would allow prediction of comfort trajectories over time rather than single-point estimates [256]. Temporal modelling may also clarify lead–lag relationships, for instance, whether physiological changes anticipate perceived discomfort or follow cognitive strain. These insights could refine adaptive control systems for anticipatory comfort regulation.

Multimodal expansion and sensory diversity present another promising direction. Besides the current combinations of thermal, physiological, and environmental inputs, incorporating posture, facial expression, voice tone, or eye-tracking data could enrich affective inference. Such multi-sensor frameworks would represent thermal comfort within a broader emotional and behavioural ecosystem, aligning environmental ergonomics with affective computing and social neuroscience[257]. In parallel, attention-based fusion or graph neural-network structures may provide a more explicit representation of contextual dependencies among modalities, paving the way for interpretable multimodal reasoning.

Future work should also pursue interdisciplinary integration between building science, cognitive neuroscience, and data science. Collaborative studies that link machine-learning analyses with neurophysiological observations—such as EEG or functional near-infrared spectroscopy—could illuminate how environmental stress influences neural processing during cognitive tasks[206]. This would further bridge computational predictions with mechanistic explanations of thermal–cognitive interactions.

Another important direction is the advancement of causal and interpretable modelling. Existing models, though accurate, remain correlational. Using causal-discovery algorithms or hybrid physics-informed learning could identify directional pathways between environmental variables, physiological responses, and subjective outcomes [258]. Establishing such causal mapping would elevate predictive modelling to explanatory modelling, enhancing theoretical robustness and supporting policy-level decision making in comfort standards and building regulation.

At the application level, the framework invites exploration of adaptive human–building collaboration. Embedding predictive models within smart-environment platforms could enable interactive control that balances energy use, comfort, and performance in real time. Systems might learn individual response patterns and negotiate setpoints through reinforcement learning under human feedback—creating truly co-adaptive environments that adjust around both individual needs and collective efficiency [207]. Extending this concept to remote or hybrid work scenarios, where occupants move between physical and virtual spaces, would further broaden its relevance for future workplace design.

Finally, researchers should consider the ethical and societal impacts of pervasive comfort sensing. Establishing governance frameworks ensuring data privacy, algorithmic accountability, and equitable access will be essential [259]. Interdisciplinary dialogue among engineers, psychologists, ethicists, and policy makers can ensure that thermal-cognitive technologies enhance, rather than constrain, human autonomy and wellbeing.

5.6.3 Summary

In summary, while the current study successfully demonstrates a multi-stage progression from single-target modelling to multimodal CNN–MLP fusion within a controlled office context, it also reveals the need for broader validation, richer temporal analysis, and greater interpretability. Future research that addresses these dimensions will transform the proposed framework into a versatile and ethical foundation for human-centred environmental intelligence [206].

5.7 Summary

This chapter integrated two complementary analytical perspectives to interpret the mechanisms underlying human thermal–cognitive–affective responses in office environments: the statistical analyses, which identified significant behavioural and psychological patterns, and the computational models, which further quantified and generalised these relationships. Together, they provide a coherent understanding of how environmental, cognitive, and physiological processes interact to shape comfort, performance, and stress.

At the statistical level, the results of the Friedman and post-hoc tests revealed clear effects of both thermal environment and task difficulty. For TSENV and TSATV, comfort perception differed significantly across environmental conditions, confirming sensitivity to ambient temperature changes. For OPI and STAI-S, differences emerged primarily with increasing task difficulty, indicating that cognitive load substantially modifies behavioural efficiency and emotional state. The combined effects observed across indicators—particularly the consistent association between lower comfort, higher workload, and elevated stress—suggest that these responses form an interdependent system rather than isolated outcomes. Individual differences further accentuated this interaction, reflecting heterogeneous patterns of adaptation even under identical conditions.

These statistical findings established the empirical basis for the multi-stage modelling framework. The modelling sequence, from Model I to Model IV, progressively integrated the factors identified as significant in the statistical analyses. Single-target models captured the direct influences of environment, task, and personal attributes; the multi-target Model II tested whether these correlated responses could be predicted jointly; Model III introduced physiological data to represent internal regulation processes; and Model IV achieved full multimodal fusion by combining numerical and image-based thermal indicators within a CNN–MLP architecture. This evolution mirrored the statistical discovery process: what first appeared as significant pairwise associations was gradually transformed into an integrated predictive representation capable of reproducing the complex interdependence among variables.

The convergence between these two analytical perspectives is central to the conclusions of this study. Statistical results verified that cognitive workload and thermal stimuli jointly drive comfort and stress; the models demonstrated that such joint effects can be computationally learned and generalised. Each analytical strand thus reinforced the other: statistical analysis provided explanatory reliability, while modelling provided operational and predictive strength. Together they advance a holistic view of human thermal experience as a dynamic and multimodal adaptation process, bridging the physical, cognitive, and affective domains.

Overall, the discussion in Chapter 5 shows a consistent progression—from descriptive verification to predictive integration, and from domain-specific models to a cohesive, human-centred theoretical and methodological framework. The statistical evidence grounded the interpretation of experimental phenomena, and the model development verified these mechanisms within a scalable computational structure. As a whole, the chapter demonstrates how empirical data and machine-learning approaches can jointly deepen the understanding of thermal comfort and performance, laying a solid foundation for the conclusions and broader reflections presented in Chapter 6.

Chapter 6 Conclusions

6.1 Overview of the study

This study examined how thermal environment and task difficulty jointly influence thermal comfort, task performance, and stress levels, and how these responses can be described and predicted using multimodal data.

A series of controlled laboratory experiments simulated office working conditions across three PMV-defined thermal states and three task-difficulty levels. Physiological indicators, thermal-image data, and subjective evaluations were synchronously collected, forming a multidimensional dataset linking environmental, cognitive, and affective factors.

A two-stage analytical framework was then applied: statistical analysis (mainly Friedman tests and post-hoc comparisons) identified significant effects and interactions, while progressive predictive modelling—from single-target models to fully multimodal CNN–MLP fusion—quantified and generalised these relationships.

The overall aim was to develop an integrated data-driven representation of how humans adaptively regulate comfort, performance, and emotion under concurrent thermal and cognitive load.

6.2 Summary of key findings

6.2.1 Statistical results

The statistical analyses verified that both the thermal environment and task difficulty exerted significant main and interaction effects on the measured responses. Thermal sensation and satisfaction showed strong differences across temperature levels; workload and stress increased steadily with task difficulty. Interaction effects indicated that under heavier cognitive demand, sensitivity to temperature reduced slightly and satisfaction ratings became less extreme. Neutral thermal conditions produced the highest objective performance and subjective satisfaction.

Overall, comfort, task performance, workload, and stress were closely correlated, forming an interdependent behavioural–psychological pattern rather than separate outcomes.

6.2.2 Modelling results

The sequential modelling confirmed and extended these relationships. Single-target models reproduced local effects but had limited generalisation. Multi-target learning (Model II) exposed cross-response dependencies yet suffered from data and optimisation constraints. Including physiological data (Model III) notably improved stability, demonstrating that physiological reactions mediate the transition from environmental stimuli to psychological states. The final multimodal CNN–MLP model (Model IV), incorporating thermal-facial images, achieved the best overall accuracy and coherence, successfully capturing the intertwined variation of comfort, workload, and stress. This multimodal representation transformed the experimentally observed interactions into a robust computational structure.

6.3 Theoretical contributions

The research contributes new theoretical insight into human thermal-cognitive adaptation. First, it establishes that thermal comfort and cognitive workload must be studied jointly: cognitive engagement alters thermal perception and tolerance, while discomfort feeds back to influence attention and emotional regulation. This interplay forms the basis of a thermal–cognitive interaction framework, extending conventional physical models (PMV, adaptive comfort) by adding cognitive and affective dimensions.

Second, the multimodal modelling approach provides a computational embodiment of theory. By linking physiological and perceptual data, the models simulate the integrative regulation through which humans balance environmental and mental demands. Thermal imagery, in particular, serves as interpretable evidence of embodied thermal expression, offering a measurable bridge between external conditions and internal states.

Finally, the study enriches environmental psychology and human-factors research by conceptualising comfort not merely as temperature preference but as a dynamic equilibrium sustaining both performance and emotional stability. This integrated understanding complements the long-established physical foundations of thermal comfort with a human-centred cognitive-affective perspective.

6.4 Overall implications

Beyond specific results, the research collectively supports a shift toward adaptive, person-centred environmental design. Thermal setpoints should reflect cognitive activity rather than static physical criteria: demanding mental work can be performed comfortably within a slightly wider thermal range, improving energy efficiency without loss of wellbeing. The demonstrated multimodal modelling pipeline also provides a methodological template for intelligent indoor-environment systems that monitor human states through physiological and thermal cues and adjust conditions automatically. These outcomes highlight the connection between empirical thermal-comfort science and the broader agenda of sustainable, responsive workplace design.

While detailed operational recommendations were discussed in Chapter 5, the essence is clear: a deeper understanding of human thermal-cognitive interaction yields not only scientific but practical value—helping future environments maintain comfort, effectiveness, and health concurrently.

6.5 Synthesis of the research logic

Collectively, the study followed a coherent methodological logic.

1. Controlled experiments revealed consistent behavioural and physiological evidence for the dual influence of thermal and cognitive load.

2. Statistical testing quantitatively verified these effects and their interactions.

3. Machine-learning models successively encoded, validated, and generalised these relationships, culminating in a multimodal architecture embodying the theoretical model of human adaptation.

This progression—from observation to explanation to prediction—demonstrates how data-driven methods can advance theoretical constructs in environmental ergonomics. Each step built upon the previous one: what was statistically significant was later shown to be computationally learnable and practically predictable.

6.6 Final remarks

This thesis concludes that thermal comfort is a dynamic, integrative process that arises from the continuous negotiation between physical conditions, cognitive demand, and physiological-affective regulation. Through controlled experimentation, objective analysis, and multimodal modelling, the study provided convergent evidence linking comfort, performance, and stress in a unified adaptive framework.

The final multimodal CNN–MLP model exemplifies how modern data analytics can translate behavioural insight into practical prediction tools, moving the discipline from descriptive assessment toward intelligent environmental intelligence. Conceptually, the research reframes comfort as an active process embedded in cognition; methodologically, it demonstrates how physiological and visual data can enrich environmental modelling; and in a broader sense, it reinforces the principle that future workplaces and buildings must serve both efficiency and wellbeing.

In closing, this work advances an emerging paradigm of thermal-cognitive adaptation, where human comfort, mental performance, and emotional balance are viewed as interconnected facets of sustainable environmental design. By revealing these mechanisms, the study contributes both scientific clarity and a foundation for future intelligent systems that are not only thermally efficient but also adaptive, empathetic, and human-centered.

Appendix A: Garment insulation I_{clu} Table

Garment Description	I_{clu}, clo	Garment Description	I_{clu}, clo
Underwear		Trousers and Coveralls	
Bra	0.01	Short shorts	0.06
Panties	0.03	Walking shorts	0.08
Men's brief	0.04	Straight trousers (thin)	0.15
T-shirt	0.08	Straight trousers (thick)	0.24
Half slip	0.14	Sweatpants	0.28
Long underwear bottoms	0.15	Overalls	0.30
Full slip	0.16	Coveralls	0.49
Long underwear top	0.20	Dress and skirts	
Footwear		Skirt (thin) mm	0.14
Ankle-length athletic socks	0.02	Skirt (thick)	0.23
Panty hose/stockings	0.02	Sleeveless, scoop neck (thin)	0.23
Sandals/thongs	0.02	Sleeveless, scoop neck (thick)	0.27
Shoes	0.02	Short-sleeve, shirtdress (thin)	0.29
Slippers (quilted, pile lined)	0.03	Long-sleeve shirtdress (thin)	0.33
Calf-length socks	0.03	Long-sleeve shirtdress (thick)	0.47
Knee socks (thick)	0.06	Sweaters	
Boots	0.10	Sleeveless vest (thin)	0.13
Shirt and Blouses		Sleeve vest (thick)	0.22
Sleeveless/scoop-neck blouse	0.12	Long-sleeve (thin)	0.25
Short-sleeve knit sport shirt	0.17	Long-sleeve (thick)	0.36
Short-sleeve dress shirt	0.19	Suit Jacket and Vests	
Long-sleeve dress shirt	0.25	Sleeveless vest (thin)	0.10
Long-sleeve flannel shirt	0.34	Sleeveless vest (thick)	0.17
Long-sleeve sweatshirt	0.34	Single-breasted (thin)	0.36
		Single-breasted (thick)	0.44
		Double-breasted (thin)	0.42
		Double-breasted (thick)	0.48

Appendix B: Code of the task simulation programme

```
<!DOCTYPE html>

<html lang="en">

<head>

  <meta charset="UTF-8">

  <meta name="viewport" content="width=device-width, initial-scale=1.0">

  <title>Triple Task Experiment</title>

  <style>

    body {

      font-family: Arial, sans-serif;

      text-align: center;

      margin: 0;

      padding: 0;

      background-color: #f8f9fa;

    }

    .container {

      margin-top: 20px;
```

```
}
.timer {
  font-size: 20px;
  font-weight: bold;
  margin-bottom: 10px;
}
.difficulty-label {
  font-size: 18px;
  margin-bottom: 20px;
  color: #333;
}
.stimulus {
  font-size: 28px;
  margin: 20px;
  color: black;
  white-space: pre-line;
}
```

```
.instructions {
  font-size: 18px;
  margin-bottom: 20px;
}
.hidden { display: none; }
button {
  padding: 10px 20px;
  font-size: 16px;
  margin: 5px;
  cursor: pointer;
}
.result {
  font-size: 20px;
  margin-top: 20px;
  white-space: pre-line;
}
.link {
```

```
  margin-top: 20px;
  font-size: 16px;
  color: blue;
  text-decoration: underline;
  cursor: pointer;
}
input[type="number"] {
  padding: 5px;
  font-size: 16px;
  margin-right: 10px;
}
#stats {
  position: fixed;
  top: 10px;
  left: 10px;
  background-color: #fff;
  padding: 10px;
```

```
      border: 1px solid #ddd;

      font-size: 14px;

      text-align: left;

      z-index: 1000;

    }

  </style>

</head>

<body>

  <div class="container">

    <h1>Triple Task Experiment</h1>

    <p class="timer" id="timer">Time remaining: 3:00</p>

    <p class="difficulty-label hidden" id="currentDifficulty"></p>

    <p class="instructions" id="instructions">Please select a difficulty level to start.</p>



    <!-- Difficulty Selection -->

    <div id="difficultySelection">

      <button onclick="startExperiment('easy')">Easy</button>
```

```
  <button onclick="startExperiment('moderate')">Moderate</button>

  <button onclick="startExperiment('hard')">Hard</button>

</div>



<!-- Task Area -->

<div id="taskArea" class="hidden">

  <div id="stimulus" class="stimulus"></div>

  <div id="taskInstructions" class="instructions"></div>

  <div id="responseArea"></div>

  <div id="feedback" class="hidden"></div>

</div>



<!-- Statistics -->

<div id="stats" class="hidden">

  <p>Tasks completed: <span id="tasksCompletedCount">0</span></p>

  <p>Correct responses: <span id="correctResponsesCount">0</span></p>

</div>
```

```
    <!-- End Screen -->
    <div id="endScreen" class="hidden">
      <p class="result" id="results"></p>
      <p class="link">
        Please complete the questionnaire:
        <a href="https://www.wjx.cn/vm/wFjPxLn.aspx" target="_blank">Fill in
survey</a>
      </p>
      <p>After finishing, please press the wristband button.</p>
      <button onclick="resetExperiment()">Select difficulty</button>
      <button onclick="restartExperiment()">Restart</button>
    </div>
  </div>

  <script>
    // Global variables
    let experimentDuration = 180000; // 3 minutes
```

```
let remainingTime = experimentDuration;
let difficulty = 'easy';

let tasksCompleted = 0;
let correctResponses = 0;
let totalResponseTime = 0;
let taskStartTime = 0;

let taskTimeout;
let taskInterval;

const tasks = ['stroop', 'pattern', 'math'];

const difficultySettings = {
  easy:     { taskTime: 9000, patternSize: 4, mathRange: 10, congruentProb: 0.5, colors: 2 },
  moderate: { taskTime: 7200, patternSize: 6, mathRange: 99, congruentProb: 0.3, colors: 4 },
```

```
    hard:       { taskTime: 6000, patternSize: 8, mathRange: 100, congruentProb: 0.1,
colors: 6 }
  };

  function randomElement(array) {
    return array[Math.floor(Math.random() * array.length)];
  }

  function randomInt(min, max) {
    return Math.floor(Math.random() * (max - min + 1)) + min;
  }

  // Start experiment
  function startExperiment(selectedDifficulty) {
    difficulty = selectedDifficulty;
    document.getElementById('difficultySelection').classList.add('hidden');
    document.getElementById('taskArea').classList.remove('hidden');
    document.getElementById('stats').classList.remove('hidden');
```

```
    document.getElementById('instructions').classList.add('hidden');
    document.getElementById('currentDifficulty').classList.remove('hidden');
    document.getElementById('currentDifficulty').textContent = "Difficulty: " +
capitalize(difficulty);
    startTasks();
    startTimer();
  }

  function capitalize(str) {
    return str.charAt(0).toUpperCase() + str.slice(1);
  }

  // Timer
  function startTimer() {
    const timerElement = document.getElementById('timer');
    const interval = setInterval(() => {
      remainingTime -= 1000;
      const minutes = Math.floor(remainingTime / 60000);
```

```
      const seconds = Math.floor((remainingTime % 60000) / 1000);
      timerElement.textContent = `Time remaining: ${minutes}:${seconds.toString().padStart(2, '0')}`;

      if (remainingTime <= 0) {
        clearInterval(interval);
        endExperiment();
      }
    }, 1000);
  }

  function startTasks() {
    runTask();
    taskInterval = setInterval(() => runTask(), difficultySettings[difficulty].taskTime + 1000);
  }

  function runTask() {
```

```
    const taskType = randomElement(tasks);

    const feedbackElement = document.getElementById('feedback');

    if (feedbackElement) feedbackElement.classList.add('hidden');

    taskStartTime = Date.now();


    switch (taskType) {

      case 'stroop': runStroopTask(); break;

      case 'pattern': runPatternTask(); break;

      case 'math': runMathTask(); break;

    }


    taskTimeout = setTimeout(() => handleTimeout(),
difficultySettings[difficulty].taskTime);

  }


  function handleTimeout() {

    tasksCompleted++;

    updateStats();
```

```
  const feedbackElement = document.getElementById('feedback');
  if (feedbackElement) {
    feedbackElement.textContent = 'Timeout! This trial is incorrect.';
    feedbackElement.classList.remove('hidden');
  }
  clearTimeout(taskTimeout);
}


// Stroop Task
function runStroopTask() {
  const allColors = [
    { name: "Red", value: "red" },
    { name: "Blue", value: "blue" },
    { name: "Green", value: "green" },
    { name: "Yellow", value: "yellow" },
    { name: "Orange", value: "orange" },
    { name: "Purple", value: "purple" }
```

```
];

const colorCount = difficultySettings[difficulty].colors;

const chosenColors = allColors.slice(0, colorCount);



let word, color;

if (Math.random() < difficultySettings[difficulty].congruentProb) {

  const choice = randomElement(chosenColors);

  word = choice.name;

  color = choice.value;

} else {

  const wordChoice = randomElement(chosenColors);

  word = wordChoice.name;

  let diffChoice;

  do {

    diffChoice = randomElement(chosenColors);

  } while (diffChoice.name === wordChoice.name);

  color = diffChoice.value;
```

```
  }

  const stimulusElement = document.getElementById('stimulus');
  stimulusElement.textContent = word;
  stimulusElement.style.color = color;
  document.getElementById('taskInstructions').textContent = 'Select the font color:';
  const responseArea = document.getElementById('responseArea');
  responseArea.innerHTML = '';

  chosenColors.forEach((c) => {
    const button = document.createElement('button');
    button.textContent = c.name;
    button.onclick = () => handleResponse(c.value, color);
    responseArea.appendChild(button);
  });
}
```

```
// Pattern Comparison Task

function runPatternTask() {

  const pattern1 = generatePattern(difficultySettings[difficulty].patternSize);

  const pattern2 = Math.random() > 0.5 ? [...pattern1] : generateModifiedPattern(pattern1);



  const stimulusElement = document.getElementById('stimulus');

  stimulusElement.style.color = "black";

  stimulusElement.textContent = `Pattern 1: ${pattern1.join(' ')}\nPattern 2: ${pattern2.join(' ')}`;

  document.getElementById('taskInstructions').textContent = 'Are they the same?';

  const responseArea = document.getElementById('responseArea');

  responseArea.innerHTML = `

    <button onclick="handleResponse('${arePatternsEqual(pattern1, pattern2) ? 'Same' : 'Different'}', 'Same')">Same</button>

    <button onclick="handleResponse('${arePatternsEqual(pattern1, pattern2) ? 'Same' : 'Different'}', 'Different')">Different</button>

  `;

}
```

```
function generatePattern(size) {
  const shapes = ['○', '●', '△', '▲', '□', '■'];
  return Array.from({ length: size }, () => randomElement(shapes));
}

function generateModifiedPattern(pattern) {
  const modifiedPattern = [...pattern];
  const i = randomInt(0, pattern.length - 1);
  const shapes = ['○', '●', '△', '▲', '□', '■'];
  let newShape;
  do { newShape = randomElement(shapes); }
  while (newShape === modifiedPattern[i]);
  modifiedPattern[i] = newShape;
  return modifiedPattern;
}
```

```
function arePatternsEqual(p1, p2) {
  return JSON.stringify(p1) === JSON.stringify(p2);
}

// Math Task
function runMathTask() {
  const num1 = randomInt(1, difficultySettings[difficulty].mathRange);
  const num2 = randomInt(1, difficultySettings[difficulty].mathRange);
  const operation = randomElement(['+', '-']);
  const correctAnswer = operation === '+' ? num1 + num2 : num1 - num2;

  const stimulusElement = document.getElementById('stimulus');
  stimulusElement.style.color = "black";
  stimulusElement.textContent = `${num1} ${operation} ${num2} = ?`;
  document.getElementById('taskInstructions').textContent = 'Enter your answer:';
  const responseArea = document.getElementById('responseArea');
  responseArea.innerHTML = `
```

```
        <input type="number" id="mathInput">
        <button onclick="handleResponse(document.getElementById('mathInput').value, ${correctAnswer})">Submit</button>
      `;
    }


    // Response handling
    function handleResponse(response, correctResponse) {
      clearTimeout(taskTimeout);
      const responseTime = Date.now() - taskStartTime;
      totalResponseTime += responseTime;
      tasksCompleted++;


      if (response == correctResponse) {
        correctResponses++;
        document.getElementById('feedback').textContent = `Correct! Response time: ${responseTime} ms`;
```

```
    } else {
      document.getElementById('feedback').textContent = `Incorrect! Correct answer: ${correctResponse}`;
    }
    document.getElementById('feedback').classList.remove('hidden');
    updateStats();
  }

  function updateStats() {
    document.getElementById('tasksCompletedCount').textContent = tasksCompleted;
    document.getElementById('correctResponsesCount').textContent = correctResponses;
  }

  // End experiment
  function endExperiment() {
    clearInterval(taskInterval);
```

```
      document.getElementById('taskArea').classList.add('hidden');
      document.getElementById('stats').classList.add('hidden');
      document.getElementById('endScreen').classList.remove('hidden');

      const avgResponseTime = (totalResponseTime / tasksCompleted).toFixed(2);
      const accuracy = ((correctResponses / tasksCompleted) * 100).toFixed(2);

      document.getElementById('results').textContent =
`Experiment finished!
Tasks completed: ${tasksCompleted}
Average response time: ${avgResponseTime} ms
Accuracy: ${accuracy}%`;
    }

    function resetExperiment() {
      location.reload();
    }
```

```
    function restartExperiment() {
      document.getElementById('endScreen').classList.add('hidden');
      document.getElementById('taskArea').classList.remove('hidden');
      document.getElementById('stats').classList.remove('hidden');
      tasksCompleted = 0;
      correctResponses = 0;
      totalResponseTime = 0;
      remainingTime = experimentDuration;
      startTasks();
      startTimer();
    }
  </script>
</body>
</html>
```

Appendix C: Experiment Survey

Subject ID: ________

1. What is your task difficulty at present?

□ Easy □ Moderate □ Hard

2. How many questions did you finish? __________

3. How many questions did you finish correctly? __________

4. What is your average reacting time (can be found in the result surface)? __________

5. What is your accuracy rate? __________

Section I: Thermal comfort

6. What is your general thermal sensation? (Check the one that is more appropriate)

□ Hot
□ Warm
□ Slightly warm
□ Neutral
□ Slightly cool
□ Cool
□ Cold

7. How satisfied are you with the temperature in your space? (Check the one that is more appropriate)

□ +3	□ +2	□ +1	□ 0	□ -1	□ -2	□ -3
Very satisfied					Very dissatisfied	

Section II: Workload sensation

8. Please answer the following questions to subjectively assess how you feel about your current work assignment, divided into six areas: mental demand, physical demand, temporal demand, effort, performance and frustration level. (Check the one that is more appropriate)

4.1 Mental Demand — How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?

Very Low — Very High

4.2 Physical Demand — How much physical activity was required? Was the task easy or demanding, slack or strenuous?

Very Low — Very High

4.3 Temporal Demand — How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?

Very Low — Very High

4.4 Performance — How successful were you in performing the task? How satisfied were you with your performance?

Perfect — Failure

4.5 Effort | How hard did you have to work (mentally and physically) to accomplish your level of performance?

Very Low | Very High

4.6 Frustration level | How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?

Very Low | Very High

Section III: Stress

9. Please answer the following questions to subjectively assess how you feel about your mood. (Check the one that is more appropriate)

Self-evaluation questionnaire

	1	**2**	**3**	**4**
(1) I feel calm	□	□	□	□
(2) I feel secure	□	□	□	□
(3) I am tense	□	□	□	□
(4) I feel strained	□	□	□	□
(5) I feel at ease	□	□	□	□
(6) I feel upset	□	□	□	□
(7) I am presently worrying over possible misfortunes	□	□	□	□
(8) I feel satisfied	□	□	□	□
(9) I feel frightened	□	□	□	□
(10) I feel comfortable	□	□	□	□
(111) I feel self-confident	□	□	□	□
(12) I feel nervous	□	□	□	□
(13) I am jittery	□	□	□	□
(14) I feel indecisive	□	□	□	□
(15) I am relaxed	□	□	□	□

(16) I feel content	◻	◻	◻	◻
(17) I am worried	◻	◻	◻	◻
(18) I feel confused	◻	◻	◻	◻
(19) I feel steady	◻	◻	◻	◻
(20) I feel pleasant	◻	◻	◻	◻

1: Not at all, 2: Somewhat, 3: Moderately so, 4: Very much so

Appendix D: Pre-experiment Survey

Subject ID: ________

1. What is your gender? □ Female □ Male

2. What is your age? ________ years old

3. What is your height? ________ cm

4. What is your weight? ________kg

5. Using the list below, please check each item of clothing that you are wearing right now. (Check all that apply):

□ Short-sleeve shirt	□ Long-sleeve shirt	□T-shirt	□ Long-sleeve Sweatshirt
□ Sweater	□ Vest	□ Jacket	□ Knee-length skirt
□ Ankle-length skirt	□ Dress	□ Shorts	□ Athletic sweatpants
□ Trousers	□ Undershirt	□ Overalls	□ Long underwear bottoms
□ Slip	□ Nylons	□ Socks	□ Long sleeve coveralls
□ Boots	□ Shoes	□ Sandals	
□ Other: (Please note if you are wearing something not described above, or if you think something you are wearing is especially heavy.) ______________________________			

Your clothing insulation I_{clu} is _____________clo

References

[1] S. Brasche and W. Bischof, "Daily time spent indoors in German homes – Baseline data for the assessment of indoor exposure of German occupants," *International Journal of Hygiene and Environmental Health,* vol. 208, no. 4, pp. 247–253, 2005/7// 2005, doi: 10.1016/j.ijheh.2005.03.003.

[2] N. E. Klepeis *et al.*, "The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants," *Journal of Exposure Science & Environmental Epidemiology,* vol. 11, no. 3, pp. 231–252, 2001/7// 2001, doi: 10.1038/sj.jea.7500165.

[3] W. W. Che *et al.*, "Energy consumption, indoor thermal comfort and air quality in a commercial office with retrofitted heat, ventilation and air conditioning (HVAC) system," *Energy and Buildings,* vol. 201, 2019/10/15, doi: 10.1016/j.enbuild.2019.06.029.

[4] N. Asim *et al.*, "Sustainability of Heating, Ventilation and Air-Conditioning (HVAC) Systems in Buildings—An Overview," *International Journal of Environmental Research and Public Health 2022, Vol. 19, Page 1016,* vol. 19, no. 2, 2022–01–17, doi: 10.3390/ijerph19021016.

[5] L. Pérez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and Buildings,* vol. 40, no. 3, 2008/01/01, doi: 10.1016/j.enbuild.2007.03.007.

[6] P. H. Shaikh, N. B. M. Nor, P. Nallagownden, I. Elamvazuthi, and T. Ibrahim, "A review on optimized control systems for building energy and comfort management of smart sustainable buildings," *Renewable and Sustainable Energy Reviews,* vol. 34, 2014/06/01, doi: 10.1016/j.rser.2014.03.027.

[7] American National Standards Institute, "“Thermal Environmental Conditions for Human Occupancy,” in American Society of Heating, Refrigerating and Air-Conditioning Engineers," ed, 2023.

[8] N. Fergus， , *Thermal comfort: a handbook for field studies toward an adaptive model*. University of East London, 1993.

[9] A. Wierzbicka *et al.*, "Healthy Indoor Environments: The Need for a Holistic Approach," *International Journal of Environmental Research and Public Health,* vol. 15, no. 9, pp. 1874–1874, 2018/8// 2018, doi: 10.3390/ijerph15091874.

[10] J. Wu, C. Liu, and H. Wang, "Analysis of Spatio-temporal patterns and related factors of thermal comfort in subtropical coastal cities based on local

climate zones," *Building and Environment,* vol. 207, 2022/1// 2022, doi: 10.1016/j.buildenv.2021.108568.

[11] P. Roelofsen, "The impact of office environments on employee performance: The design of the workplace as a strategy for productivity enhancement," *Journal of Facilities Management,* vol. 1, no. 3, pp. 247–264, 2002/7// 2002, doi: 10.1108/14725960310807944.

[12] M. Frontczak and P. Wargocki, "Literature survey on how different factors influence human comfort in indoor environments," *Building and environment,* vol. 46, no. 4, pp. 922–937, 2011.

[13] A. Kaushik, M. Arif, P. Tumula, and O. J. Ebohon, "Effect of thermal comfort on occupant productivity in office buildings: Response surface analysis," *Building and Environment,* vol. 180, pp. 107021–107021, 2020/8// 2020, doi: 10.1016/j.buildenv.2020.107021.

[14] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *Journal of Comparative Neurology and Psychology,* vol. 18, no. 5, pp. 459–482, 1908/11// 1908, doi: 10.1002/cne.920180503.

[15] C. P. Davey, "Physical Exertion and Mental Performance," *Ergonomics,* vol. 16, no. 5, pp. 595–599, 1973/9// 1973, doi: 10.1080/00140137308924550.

[16] A. M. Abbasi, M. Motamedzadeh, M. Aliabadi, R. Golmohammadi, and L. Tapak, "The impact of indoor air temperature on the executive functions of human brain and the physiological responses of body," *Health Promotion Perspectives,* vol. 9, no. 1, pp. 55–64, 2019/1// 2019, doi: 10.15171/hpp.2019.07.

[17] W. Cui, G. Cao, J. H. Park, Q. Ouyang, and Y. Zhu, "Influence of indoor air temperature on human thermal comfort, motivation and performance," *Building and environment,* vol. 68, pp. 114–122, 2013.

[18] Y. Geng, W. Ji, B. Lin, and Y. Zhu, "The impact of thermal environment on occupant IEQ perception and productivity," *Building and Environment,* vol. 121, pp. 158–167, 2017/8// 2017, doi: 10.1016/j.buildenv.2017.05.022.

[19] R. Kosonen and F. Tan, "Assessment of productivity loss in air-conditioned buildings using PMV index," *Energy and Buildings,* vol. 36, no. 10, pp. 987–993, 2004/10// 2004, doi: 10.1016/j.enbuild.2004.06.021.

[20] T. Cheung, S. Schiavon, T. Parkinson, P. Li, and G. Brager, "Analysis of the accuracy on PMV – PPD model using the ASHRAE Global Thermal Comfort Database II," *Building and Environment,* vol. 153, pp. 205–217, 2019/4// 2019, doi: 10.1016/j.buildenv.2019.01.055.

[21] M.-E. Boatca, A. Draghici, S. I. Irimie, B. Gajsek, and A. D. Maria-Elena Boatca, Sabin Ioan Irimie,Brigita Gajsek, "Safety, health and comfort in the

workplace: An innovative framework to support implementation of ergonomic interventions," *Human Systems Management,* 2025–1, doi: 10.3233/HSM-230159.

[22] v. J. Hoof, M. Mazej, and J. L. M. Hensen, "Thermal comfort: research and practice," *Frontiers in Bioscience,* vol. 15, no. 2, 2010, doi: 10.2741/3645.

[23] P. Robles-Granda *et al.*, "Jointly Predicting Job Performance, Personality, Cognitive Ability, Affect, and Well-Being," *IEEE Computational Intelligence Magazine,* vol. 16, no. 2, pp. 46–61, 2021, doi: 10.1109/MCI.2021.3061877.

[24] Y. Lin, J. Mutz, P. J. Clough, and K. A. Papageorgiou, "Frontiers | Mental Toughness and Individual Differences in Learning, Educational and Work Performance, Psychological Well-being, and Personality: A Systematic Review," *Frontiers in Psychology,* vol. 8, 2017/08/11, doi: 10.3389/fpsyg.2017.01345.

[25] A. C. Russo and A. C. Russo, "Assessing the Systemic Impact of Heat Stress on Human Reliability in Mining Through FRAM and Hybrid Decision Models," *Mining 2025, Vol. 5, Page 50,* vol. 5, no. 3, 2025–08–01, doi: 10.3390/mining5030050.

[26] M. K. K. Rony and H. M. Alamgir, "High temperatures on mental health: Recognizing the association and the need for proactive strategies—A perspective," *Health Science Reports,* vol. 6, no. 12, 2023/12/01, doi: 10.1002/hsr2.1729.

[27] S. S. Shahzad, J. Brennan, D. Theodossopoulos, B. Hughes, and J. K. Calautit, "Building-related symptoms, energy, and thermal control in the workplace: personal and open plan offices," *Sustainability,* vol. 8, no. 4, p. 331, 2016.

[28] M. W. Kozusznik, L. P. Maricutoiu, J. M. Peiró, D. M. Vîrgă, A. Soriano, and C. Mateo-Cecilia, "Decoupling office energy efficiency from employees' well-being and performance: a systematic review," *Frontiers in psychology,* vol. 10, p. 293, 2019.

[29] P. O. Fanger, *Thermal Comfort: Analysis and Applications in Environmental Engineering*. Danish Technical Press, 1970.

[30] W. O'Brien and H. B. Gunay, "The contextual factors contributing to occupants' adaptive comfort behaviors in offices–A review and proposed modeling framework," *Building and environment,* vol. 77, pp. 77–87, 2014.

[31] X. Wang, D. Li, C. C. Menassa, and V. R. Kamat, "Investigating the effect of indoor thermal environment on occupants' mental workload and task performance using electroencephalogram," *Building and Environment,* vol. 158, pp. 120–132, 2019.

[32] Y. Apertet, H. Ouerdane, C. Goupil, and P. Lecoeur, "Influence of thermal environment on optimal working conditions of thermoelectric generators," *Journal of Applied Physics,* vol. 116, no. 14, 2014.

[33] J. Wu, Z. Hou, J. Shen, and Z. Lian, "Quantitative effect on work performance considering interactions among multiple indoor environmental factors," *Building and Environment,* vol. 185, p. 107286, 2020.

[34] S. Zhang, N. Zhu, and S. Lv, "Human response and productivity in hot environments with directed thermal radiation," *Building and Environment,* vol. 187, p. 107408, 2021.

[35] J. S. Muñoz, M. T. Kelly, V. Flores-Alés, and C. Caamaño-Carrillo, "Recognizing the effect of the thermal environment on self-perceived productivity in offices: A structural equation modeling perspective," *Building and Environment,* vol. 210, p. 108696, 2022.

[36] J. D. Ramsey, C. L. Burford, M. Y. Beshir, and R. C. Jensen, "Effects of workplace thermal conditions on safe work behavior," *Journal of safety Research,* vol. 14, no. 3, pp. 105–114, 1983.

[37] S. Chen, J. Geng, M. Lan, S. He, and W. Weng, "Study on physiological features, subjective perception, and cognitive performance of individuals in prolonged high-temperature environments," *Building and Environment,* vol. 273, p. 112700, 2025.

[38] B.-L. Jian, C.-L. Chen, W.-L. Chu, and M.-W. Huang, "The facial expression of schizophrenic patients applied with infrared thermal facial image sequence," *BMC psychiatry,* vol. 17, no. 1, p. 229, 2017.

[39] M. Marzec, R. Koprowski, Z. Wróbel, A. Kleszcz, and S. Wilczyński, "Automatic method for detection of characteristic areas in thermal face images," *Multimedia Tools and Applications,* vol. 74, no. 12, pp. 4351–4368, 2015.

[40] S. Kawakubo, M. Sugiuchi, and S. Arata, "Office thermal environment that maximizes workers' thermal comfort and productivity," *Building and Environment,* vol. 233, p. 110092, 2023.

[41] L. Groenesteijn, R. P. Ellegast, K. Keller, F. Krause, H. Berger, and M. P. de Looze, "Office task effects on comfort and body dynamics in five dynamic office chairs," *Applied Ergonomics,* vol. 43, no. 2, pp. 320–328, 2012/3// 2012, doi: 10.1016/j.apergo.2011.06.007.

[42] S. R. Walter, W. T. Dunsmuir, and J. I. Westbrook, "Studying interruptions and multitasking in situ: the untapped potential of quantitative observational studies," *International Journal of Human-Computer Studies,* vol. 79, pp. 118–125, 2015.

[43] G. Havenith and D. Fiala, "Thermal indices and thermophysiological modeling for heat stress," *Comprehensive physiology,* vol. 6, no. 1, pp. 255–302, 2011.

[44] M. J. Buller, A. P. Welles, and K. E. Friedl, "Wearable physiological monitoring for human thermal-work strain optimization," *Journal of applied physiology,* vol. 124, no. 2, pp. 432–441, 2018.

[45] H. Arruda, E. R. Silva, M. Lessa, D. Proença Jr, and R. Bartholo, "VOSviewer and bibliometrix," *Journal of the Medical Library Association: JMLA,* vol. 110, no. 3, p. 392, 2022.

[46] I. Mujan, A. S. Anđelković, V. Munćan, M. Kljajić, and D. Ružić, "Influence of indoor environmental quality on human health and productivity-A review," *Journal of cleaner production,* vol. 217, pp. 646–657, 2019.

[47] G. Havenith, I. Holmér, and K. Parsons, "Personal factors in thermal comfort assessment: clothing properties and metabolic heat production," *Energy and buildings,* vol. 34, no. 6, pp. 581–591, 2002.

[48] S. Naheed and S. Shooshtarian, "A review of cultural background and thermal perceptions in urban environments," *Sustainability,* vol. 13, no. 16, p. 9080, 2021.

[49] I. Andrić, A. Pina, P. Ferrão, J. Fournier, B. Lacarrière, and O. Le Corre, "The impact of climate change on building heat demand in different climate types," *Energy and Buildings,* vol. 149, pp. 225–234, 2017.

[50] B. R. Yurkevicius, B. K. Alba, A. D. Seeley, and J. W. Castellani, "Human cold habituation: physiology, timeline, and modifiers," *Temperature,* vol. 9, no. 2, pp. 122–157, 2022.

[51] D. Tu, J. Xie, and J. Liu, "Study on the impact of periodic temperature fluctuations on tunnel thermal environment and natural ventilation," *Applied Thermal Engineering,* p. 128520, 2025.

[52] K. Upadhyay, R. Elangovan, and S. Subudhi, "Quantifying human thermal adaptation for indoor environments through field surveys and climate chamber experiments," *Advances in Building Energy Research,* vol. 19, no. 1, pp. 1–36, 2025.

[53] H. Jain, "Critical insights into thermal comfort optimization and heat resilience in indoor spaces," *City and Built Environment,* vol. 2, no. 1, p. 14, 2024.

[54] A. Zhivov, "Requirements for building thermal conditions," *Retrieved,* vol. 10, no. 21, p. 2013, 2013.

[55] S. Altomonte, P. Rutherford, and R. Wilson, "Human factors in the design of sustainable built environments," *Intelligent Buildings International,* vol. 7, no. 4, pp. 224–241, 2015.

[56] M. Prek, "Thermodynamical analysis of human thermal comfort," *Energy,* vol. 31, no. 5, pp. 732–743, 2006.

[57] A. S. dos Reis, P. Vaquero, M. F. Dias, and A. Tavares, "Passive Discomfort Index as an alternative to Predicted Mean Vote and Predicted Percentage of Dissatisfied to assess occupant's thermal discomfort in dwellings," *Energy Reports,* vol. 8, pp. 956–965, 2022.

[58] K. Fabbri, "The indoor thermal comfort indexes PMV and PPD," in *Thermal Comfort Perception: A Questionnaire Approach Focusing on Children*: Springer, 2024, pp. 83–135.

[59] F. R. d. A. Alfano, E. Ianniello, and B. I. Palella, "PMV–PPD and acceptability in naturally ventilated schools," *Building and environment,* vol. 67, pp. 129–137, 2013.

[60] S. Liu, N. Nazarian, J. Niu, M. A. Hart, and R. de Dear, "From thermal sensation to thermal affect: A multi-dimensional semantic space to assess outdoor thermal comfort," *Building and Environment,* vol. 182, p. 107112, 2020.

[61] C. D. Roa, S. Schiavon, and T. Parkinson, "Targeted occupant surveys: A novel method to effectively relate occupant feedback with environmental conditions," *Building and Environment,* vol. 184, p. 107129, 2020.

[62] J.-Y. Lee, "Thermal Perception and Thermal Sensation," in *The Thermal Environment: From Viewpoints of Physiological Anthropology and Environmental Ergonomics*: Springer, 2025, pp. 67–81.

[63] I. Abd Manaf, F. Durrani, and M. Eftekhari, "A review of desiccant evaporative cooling systems in hot and humid climates," *Advances in Building Energy Research,* vol. 15, no. 1, pp. 1–42, 2021.

[64] J. Gao, Y. Wang, and P. Wargocki, "Comparative analysis of modified PMV models and SET models to predict human thermal sensation in naturally ventilated buildings," *Building and Environment,* vol. 92, pp. 200–208, 2015.

[65] D. H. C. Toe and T. Kubota, "Development of an adaptive thermal comfort equation for naturally ventilated buildings in hot–humid climates using ASHRAE RP-884 database," *Frontiers of architectural research,* vol. 2, no. 3, pp. 278–291, 2013.

[66] K. L. Holt, "Predictive Modeling of Occupational Exposure Using Machine Learning and Environmental Sensor Data," *Journal of Exceptional Multidisciplinary Research,* vol. 2, no. 1, pp. 82–89, 2025.

[67] K. Jitkhajornwanich, "Expectation and experience of thermal comfort in transitional spaces: a field study of thermal environments in hot-humid climate of Bangkok," University of Sheffield, 2000.

[68] W. Ji, R. de Dear, J. Kim, Y. Zhu, B. Cao, and S. Liu, "Study on the influence of climatic thermal exposure environment changed from cold to hot on human thermal preference," *Building and Environment,* vol. 207, p. 108430, 2022.

[69] F. Meggers *et al.*, "The Thermoheliodome–"Air conditioning" without conditioning the air, using radiant cooling and indirect evaporation," *Energy and Buildings,* vol. 157, pp. 11–19, 2017.

[70] S. Aghniaey and T. M. Lawrence, "The impact of increased cooling setpoint temperature during demand response events on occupant thermal comfort in commercial buildings: A review," *Energy and Buildings,* vol. 173, pp. 19–27, 2018.

[71] F. Z. Chafi and S. Hallé, "Three dimensional study for evaluating of air flow movements and thermal comfort in a model room: Experimental validation," *Energy and Buildings,* vol. 43, no. 9, pp. 2156–2166, 2011.

[72] L. Yang *et al.*, "Comparative analysis of indoor thermal environment characteristics and occupants' adaptability: Insights from ASHRAE RP-884 and the Chinese thermal comfort database," *Energy and Buildings,* vol. 309, p. 114033, 2024.

[73] J. Kasambala, *The development of a conceptual framework of female clothing evaluative criteria preferences during the purchasing decision that includes body shape, personal values and emotions*. University of South Africa (South Africa), 2016.

[74] S. Shahzad, J. Brennan, D. Theodossopoulos, J. K. Calautit, and B. R. Hughes, "Does a neutral thermal sensation determine thermal comfort?," *Building Services Engineering Research and Technology,* vol. 39, no. 2, pp. 183–195, 2018.

[75] L. Jin, Y. Zhang, and Z. Zhang, "Human responses to high humidity in elevated temperatures for people in hot-humid climates," *Building and Environment,* vol. 114, pp. 257–266, 2017.

[76] M. R. Safizadeh, M. Schweiker, and A. Wagner, "Experimental evaluation of radiant heating ceiling systems based on thermal comfort criteria," *Energies,* vol. 11, no. 11, p. 2932, 2018.

[77] W. Sun, K. Cheong, and A. K. Melikov, "Subjective study of thermal acceptability of novel enhanced displacement ventilation system and implication of occupants' personal control," *Building and environment,* vol. 57, pp. 49–57, 2012.

[78] S. G. Farris, L. Derby, and M. M. Kibbey, "Getting comfortable with physical discomfort: A scoping review of interoceptive exposure in physical and mental health conditions," *Psychological Bulletin,* vol. 151, no. 2, p. 131, 2025.

[79] S. Crosby and A. Rysanek, "Correlations between thermal satisfaction and non-thermal conditions of indoor environmental quality: Bayesian inference of a field study of offices," *Journal of Building Engineering,* vol. 35, p. 102051, 2021.

[80] X. Su, Z. Wang, and Y. Yang, "Field study of cold radiant asymmetry caused by exterior built elements of educational buildings in severe cold area, China," *Energy and buildings,* vol. 252, p. 111401, 2021.

[81] Y. Zhai, H. Zhang, Y. Zhang, W. Pasut, E. Arens, and Q. Meng, "Comfort under personally controlled air movement in warm and humid environments," *Building and environment,* vol. 65, pp. 109–117, 2013.

[82] M. Favero, I. Sartori, and S. Carlucci, "Human thermal comfort under dynamic conditions: An experimental study," *Building and Environment,* vol. 204, p. 108144, 2021.

[83] M. Schweiker and A. Wagner, "The effect of occupancy on perceived control, neutral temperature, and behavioral patterns," *Energy and Buildings,* vol. 117, pp. 246–259, 2016.

[84] Y. Huang and W. Jiang, "The effect of exposure duration on the subjective discomfort of aircraft cabin noise," *Ergonomics,* vol. 60, no. 1, pp. 18–25, 2017.

[85] F. Lolli, S. Marinello, A. M. Coruzzolo, and M. A. Butturi, "Post-occupancy evaluation's (POE) applications for improving indoor environment quality (IEQ)," *Toxics,* vol. 10, no. 10, p. 626, 2022.

[86] B. Lala, A. Biju, A. Rastogi, K. Dahiya, S. M. Kala, and A. Hagishima, "The Challenge of Multiple Thermal Comfort Prediction Models: Is TSV Enough?," *Buildings,* vol. 13, no. 4, p. 890, 2023.

[87] R. T. Hellwig, "Perceived control in indoor environments: a conceptual approach," *Building Research & Information,* vol. 43, no. 3, pp. 302–315, 2015.

[88] J. Zhao, F. Abdul Aziz, Y. Deng, N. Ujang, and Y. Xiao, "A review of comprehensive post-occupancy evaluation feedback on occupant-centric thermal comfort and building energy efficiency," *Buildings,* vol. 14, no. 9, p. 2892, 2024.

[89] M. N. Sawka, J. W. Castellani, S. N. Cheuvront, and A. J. Young, "Physiologic systems and their responses to conditions of heat and cold," *ACSM's advanced exercise physiology. Baltimore: Lippincott Williams & Wilkins,* pp. 567–602, 2012.

[90] J. A. Ross, E. M. Shipp, A. B. Trueblood, and A. Bhattacharya, "Ergonomics and beyond: understanding how chemical and heat exposures and physical

exertions at work affect functional ability, injury, and long-term health," *Human factors,* vol. 58, no. 5, pp. 777–795, 2016.

[91] D. Mota-Rojas *et al.*, "Physiological and behavioral mechanisms of thermoregulation in mammals," *Animals,* vol. 11, no. 6, p. 1733, 2021.

[92] L. L. Jackson and H. R. Rosenberg, "Preventing heat-related illness among agricultural workers," *Journal of agromedicine,* vol. 15, no. 3, pp. 200–215, 2010.

[93] S. Basu and A. C. King, "Disability and chronic disease among older adults in India: detecting vulnerable populations through the WHO SAGE Study," *American journal of epidemiology,* vol. 178, no. 11, pp. 1620–1628, 2013.

[94] S. Radandt, J. Rantanen, and O. Renn, "Governance of occupational safety and health and environmental risks," in *Risks in modern society*: Springer, 2008, pp. 127–258.

[95] M. S. Saghiv and M. S. Sagiv, "Thermoregulation," in *Basic Exercise Physiology: Clinical and Laboratory Perspectives*: Springer, 2020, pp. 437–463.

[96] E. A. Arens and H. Zhang, "The skin's role in human thermoregulation and comfort," 2006.

[97] T. Wijayanto, "Physiological Responses and Performance in Hot Environments," in *The Thermal Environment: From Viewpoints of Physiological Anthropology and Environmental Ergonomics*: Springer, 2025, pp. 139–158.

[98] T. Salthammer and G. C. Morrison, "Temperature and indoor environments," *Indoor air,* vol. 32, no. 5, p. e13022, 2022.

[99] C. G. Crandall and T. E. Wilson, "Human cardiovascular responses to passive heat stress," *Comprehensive Physiology,* vol. 5, no. 1, pp. 17–43, 2011.

[100] K. Hinde, R. Lloyd, C. Low, and C. Cooke, "The effect of temperature, gradient, and load carriage on oxygen consumption, posture, and gait characteristics," *European journal of applied physiology,* vol. 117, no. 3, pp. 417–430, 2017.

[101] S. Stewart, A. K. Keates, A. Redfern, and J. J. McMurray, "Seasonal variations in cardiovascular disease," *Nature Reviews Cardiology,* vol. 14, no. 11, pp. 654–664, 2017.

[102] W. J. Wu *et al.*, "Scoping review of the characteristics and outcomes of adults presenting to the emergency department during heatwaves," *Emergency Medicine Australasia,* vol. 35, no. 6, pp. 903–920, 2023.

[103] P. Acharya, B. Boggess, and K. Zhang, "Assessing heat stress and health among construction workers in a changing climate: a review," *International*

journal of environmental research and public health, vol. 15, no. 2, p. 247, 2018.

[104] G. Savioli *et al.*, "Heat-related illness in emergency and critical care: recommendations for recognition and management with medico-legal considerations," *Biomedicines,* vol. 10, no. 10, p. 2542, 2022.

[105] S. R. Leap, D. R. Soled, V. Sampath, and K. C. Nadeau, "Effects of extreme weather on health in underserved communities," *Annals of Allergy, Asthma & Immunology,* vol. 133, no. 1, pp. 20–27, 2024.

[106] K. Collins, "Cold-and heat-related illnesses in the indoor environment," in *Unhealthy Housing*: Taylor & Francis, 2005, pp. 117–140.

[107] J. Knill-Jones, G. Shadwell, H. T. Hurst, C. Mawhinney, J. K. Sinclair, and R. Allan, "Influence of acute and chronic therapeutic cooling on cognitive performance and well-being," *Physiology & Behavior,* vol. 289, p. 114728, 2025.

[108] M. R. Irwin, "Sleep disruption induces activation of inflammation and heightens risk for infectious disease: Role of impairments in thermoregulation and elevated ambient temperature," *Temperature,* vol. 10, no. 2, pp. 198–234, 2023.

[109] M. D'Amato, A. Molino, G. Calabrese, L. Cecchi, I. Annesi-Maesano, and G. D'Amato, "The impact of cold on the respiratory tract and its consequences to respiratory health," *Clinical and translational allergy,* vol. 8, no. 1, p. 20, 2018.

[110] J. E. Morley, "An overview of cognitive impairment," *Clinics in geriatric medicine,* vol. 34, no. 4, pp. 505–513, 2018.

[111] S. Qian *et al.*, "Environmental heat stress enhances mental fatigue during sustained attention task performing: evidence from an ASL perfusion study," *Behavioural brain research,* vol. 280, pp. 6–15, 2015.

[112] P. J. Wallace, "The Effects of Heat and Cold on Cognitive Function and Endurance Capacity," 2023.

[113] C. Schmit, C. Hausswirth, Y. Le Meur, and R. Duffield, "Cognitive functioning and heat strain: performance responses and protective strategies," *Sports Medicine,* vol. 47, no. 7, pp. 1289–1302, 2017.

[114] J. Wu, L. Yang, Z. Hu, F. Gao, and X. Hu, "Perceptual response and cognitive performance during exposure to extremely cold environments," *Energy and Buildings,* vol. 251, p. 111358, 2021.

[115] E. Krivonogova, O. Krivonogova, and L. Poskotinova, "Individual-typological features of the reactivity of EEG rhythms, cardiovascular system and lactoferrin level in the conditions of general air cooling of a person," *Human physiology,* vol. 47, no. 5, pp. 533–541, 2021.

[116] X. Sun and J. Dong, "Stress response and safe driving time of bus drivers in hot weather," *International journal of environmental research and public health,* vol. 19, no. 15, p. 9662, 2022.

[117] E. L. Piza and L. N. Moton, "Proactive monitoring and operator discretion: A systematic social observation of CCTV control room operations," *Journal of Criminal Justice,* vol. 86, p. 102071, 2023.

[118] B. J. Wilson and J. M. Gottman, "Attention—The shuttle between emotion and cognition: Risk, resiliency, and physiological bases," in *Stress, coping, and resiliency in children and families*: Psychology Press, 2014, pp. 189–228.

[119] K. Martin, E. McLeod, J. Périard, B. Rattray, R. Keegan, and D. B. Pyne, "The impact of environmental stress on cognitive performance: a systematic review," *Human factors,* vol. 61, no. 8, pp. 1205–1246, 2019.

[120] W. Lee, M. Moon, H. G. Kim, T. H. Lee, and M. S. Oh, "Heat stress-induced memory impairment is associated with neuroinflammation in mice," *Journal of neuroinflammation,* vol. 12, no. 1, p. 102, 2015.

[121] M. Falla, A. Micarelli, K. Hüfner, and G. Strapazzon, "The effect of cold exposure on cognitive performance in healthy adults: a systematic review," *International journal of environmental research and public health,* vol. 18, no. 18, p. 9725, 2021.

[122] K. Vasilakopoulou and M. Santamouris, "Cumulative exposure to urban heat can affect the learning capacity of students and penalize the vulnerable and low-income young population: A systematic review," *PLOS Climate,* vol. 4, no. 7, p. e0000618, 2025.

[123] J. Goodman, M. Hurwitz, J. Park, and J. Smith, "Heat and learning," National Bureau of Economic Research, 2018.

[124] I. Schools, "Protect Children from Extreme Heat," 2025.

[125] T. M. Mäkinen *et al.*, "Effect of repeated exposures to cold on cognitive performance in humans," *Physiology & behavior,* vol. 87, no. 1, pp. 166–176, 2006.

[126] X. R. Tan *et al.*, "Elevated brain temperature under severe heat exposure impairs cortical motor activity and executive function," *Journal of Sport and Health Science,* vol. 13, no. 2, pp. 233–244, 2024.

[127] K. Hiatt and C. Rash, "The military operational environment," *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues,* pp. 3–28, 2009.

[128] M. L. Regina, A. Vecchié, A. Bonaventura, and D. Prisco, "Patient safety in internal medicine," *Textbook of Patient Safety and Clinical Risk Management,* pp. 213–252, 2020.

[129] J. Cahill, P. Cullen, S. Anwer, K. Gaynor, and S. Wilson, "The requirements for new tools for use by pilots and the aviation industry to manage risks pertaining to work-related stress (WRS) and wellbeing, and the ensuing impact on performance and safety," *Technologies,* vol. 8, no. 3, p. 40, 2020.

[130] L. Lan, P. Wargocki, D. P. Wyon, and Z. Lian, "Effects of thermal discomfort in an office on perceived air quality, SBS symptoms, physiological responses, and human performance," *Indoor air,* vol. 21, no. 5, pp. 376–390, 2011.

[131] C. C. Hampo, L. H. Schinasi, and S. Hoque, "Surviving indoor heat stress in United States: A comprehensive review exploring the impact of overheating on the thermal comfort, health, and social economic factors of occupants," *Heliyon,* vol. 10, no. 3, 2024.

[132] A. J. Martin, B. Papworth, P. Ginns, L.-E. Malmberg, R. J. Collie, and R. A. Calvo, "Real-time motivation and engagement during a month at school: Every moment of every day for every student matters," *Learning and Individual Differences,* vol. 38, pp. 26–35, 2015.

[133] M. C. Morrissey, G. J. Brewer, W. J. Williams, T. Quinn, and D. J. Casa, "Impact of occupational heat stress on worker productivity and economic cost," *American journal of industrial medicine,* vol. 64, no. 12, pp. 981–988, 2021.

[134] J. M. Stocks, N. A. Taylor, M. J. Tipton, and J. E. Greenleaf, "Human physiological responses to cold exposure," *Aviation, space, and environmental medicine,* vol. 75, no. 5, pp. 444–457, 2004.

[135] A. Ibrahim, H. Ali, A. Zghoul, and S. Jaradat, "Mood state and human evaluation of the thermal environment using virtual settings," *Indoor and Built Environment,* vol. 30, no. 1, pp. 70–86, 2021.

[136] H. Alzahrani, "Developing a model for assessing the effect of physical indoor environment quality on teachers' performance in Saudi education buildings," 2018.

[137] M. K. K. Rony and H. M. Alamgir, "High temperatures on mental health: Recognizing the association and the need for proactive strategies—A perspective," *Health science reports,* vol. 6, no. 12, p. e1729, 2023.

[138] A. Scarpa, S. C. Haden, and A. Tanaka, "Being hot-tempered: Autonomic, emotional, and behavioral distinctions between childhood reactive and proactive aggression," *Biological psychology,* vol. 84, no. 3, pp. 488–496, 2010.

[139] S. Colonna, O. Eyre, S. S. Agha, A. Thapar, S. van Goozen, and K. Langley, "Investigating the associations between irritability and hot and cool executive functioning in those with ADHD," *BMC psychiatry,* vol. 22, no. 1, p. 166, 2022.

[140] D. Acquadro Maran *et al.*, "Does heatwaves/above-average temperature increase interpersonal violent behavior? A scoping review," *JOURNAL OF HEALTH AND SOCIAL SCIENCES,* vol. 9, no. 4, pp. 462–490, 2024.

[141] E. Deli *et al.*, "Feeling the Heat: A Thermodynamic Perspective on Emotions, Motivation, and Time Perception," 2024.

[142] G. Matthews, J. S. Warm, L. E. Reinerman, L. K. Langheim, and D. J. Saxby, "Task engagement, attention, and executive control," in *Handbook of individual differences in cognition: Attention, memory, and executive control*: Springer, 2010, pp. 205–230.

[143] S. Domínguez-Amarillo, J. Fernández-Agüera, M. M. González, and T. Cuerdo-Vilches, "Overheating in schools: Factors determining children's perceptions of overall comfort indoors," *Sustainability,* vol. 12, no. 14, p. 5772, 2020.

[144] S. Koul, "Optimizing the Workplace: Key Areas Impacting the Work Environment and Strategies for Improvement," *Available at SSRN 5179630,* 2025.

[145] D. Lynott, K. Corker, L. Connell, and K. O'Brien, "The effects of temperature on prosocial and antisocial behaviour: A review and meta-analysis," *British journal of social psychology,* vol. 62, no. 3, pp. 1177–1214, 2023.

[146] K. Azuma, K. Ikeda, N. Kagi, U. Yanagi, and H. Osawa, "Evaluating prevalence and risk factors of building-related symptoms among office workers: Seasonal characteristics of symptoms and psychosocial and physical environmental factors," *Environmental Health and Preventive Medicine,* vol. 22, no. 1, p. 38, 2017.

[147] J. B. Moss and G. M. While, "The thermal environment as a moderator of social evolution," *Biological Reviews,* vol. 96, no. 6, pp. 2890–2910, 2021.

[148] A. Toomingas, "Prolonged, low-intensity, sedentary work," *Occupational Physiology,* vol. 141, 2012.

[149] I. Horváth and B. Berki, "Investigating the operational complexity of digital workflows based on human cognitive aspects," *Electronics,* vol. 12, no. 3, p. 528, 2023.

[150] P. G. Dempsey, S. E. Mathiassen, J. A. Jackson, and N. V. O'Brien, "Influence of three principles of pacing on the temporal organisation of work during cyclic assembly and disassembly tasks," *Ergonomics,* vol. 53, no. 11, pp. 1347–1358, 2010.

[151] U. S. Checklists, "Preventing Surgical Mishaps," *Outpatient Plastic Surgery, An Issue of Clinics in Plastic Surgery,* vol. 40, no. 3, p. 475, 2013.

[152] W. B. Schaufeli and M. Salanova, "Burnout, boredom and engagement in the workplace," *An introduction to contemporary work psychology,* pp. 291–320, 2013.

[153] L. R. Gomez-Mejia, P. Berrone, and M. Franco-Santos, *Compensation and organizational performance: Theory, research, and practice*. Routledge, 2014.

[154] J. Sanjog, T. Patel, and S. Karmakar, "Indoor physical work environment: An ergonomics perspective," *Int. J. Sci. Eng. Technol. Res,* vol. 2, no. 3, pp. 2278–7798, 2013.

[155] T. W. Taris and W. B. Schaufeli, "Individual well-being and performance at work: A conceptual and theoretical overview," *Current issues in work and organizational psychology,* pp. 189–204, 2018.

[156] T. J. Gabbett and E. Oetter, "From tissue to system: what constitutes an appropriate response to loading?," *Sports Medicine,* vol. 55, no. 1, pp. 17–35, 2025.

[157] G. Travers, P. Kippelen, S. J. Trangmar, and J. González-Alonso, "Physiological function during exercise and environmental stress in humans—an integrative view of body systems and homeostasis," *Cells,* vol. 11, no. 3, p. 383, 2022.

[158] A. S. o. Heating, Refrigerating, A.-C. Engineers, and A. N. S. Institute, *Thermal environmental conditions for human occupancy* (no. 2004). American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2004.

[159] M. T. Hamilton, D. G. Hamilton, and T. W. Zderic, "Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease," *Diabetes,* vol. 56, no. 11, pp. 2655–2667, 2007.

[160] R. Meuter, A. Rakotonirainy, B. Johns, P. Tran, and P. Wagner, "The dual vigilance task: Tracking changes in vigilance as a function of changes in monotonous contexts," in *International Conference on Fatigue Management in Transportation Operations*, 2005: Transportation Canada, pp. 1–17.

[161] R. J. De Dear and G. S. Brager, "Thermal comfort in naturally ventilated buildings: revisions to ASHRAE Standard 55," *Energy and buildings,* vol. 34, no. 6, pp. 549–561, 2002.

[162] P. Robles-Granda *et al.*, "Jointly predicting job performance, personality, cognitive ability, affect, and well-being," *IEEE Computational Intelligence Magazine,* vol. 16, no. 2, pp. 46–61, 2021.

[163] C. Davey, "Physical exertion and mental performance," *Ergonomics,* vol. 16, no. 5, pp. 595–599, 1973.

[164] M. Littmann, "Arbeitsbeanspruchungen und Merkmale erfolgreicher Erholungsprozesse am Wochenende," 2002.

[165] M.-E. Boatca, A. Draghici, S. I. Irimie, and B. Gajsek, "Safety, health and comfort in the workplace: An innovative framework to support implementation of ergonomic interventions," *Human Systems Management,* vol. 44, no. 1, pp. 59–68, 2025.

[166] P. Roelofsen, "The impact of office environments on employee performance: The design of the workplace as a strategy forproductivity enhancement," *Journal of facilities Management,* vol. 1, no. 3, pp. 247–264, 2002.

[167] Y. Zervas, A. Danis, and V. Klissouras, "Influence of physical exertion on mental performance with reference to training," *Perceptual and motor skills,* vol. 72, no. 3_suppl, pp. 1215–1221, 1991.

[168] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," 1908.

[169] C. C. Caruso, "Overtime and extended work shifts; recent findings on illnesses, injuries, and health behaviors," 2004.

[170] M. Hakim *et al.*, "The effect of operating room temperature on the performance of clinical and cognitive tasks," *Pediatric quality & safety,* vol. 3, no. 2, p. e069, 2018.

[171] T. T. Amasuomo and J. O. Amasuomo, "Perceived thermal discomfort and stress behaviours affecting students' learning in lecture theatres in the humid tropics," *Buildings,* vol. 6, no. 2, p. 18, 2016.

[172] P. O. Fanger, "Thermal comfort. Analysis and applications in environmental engineering," 1970.

[173] S. Igboanugo and J. G. Mielke, "The allostatic load model: a framework to understand the cumulative multi-system impact of work-related psychosocial stress exposure among firefighters," *Health Psychology and Behavioral Medicine,* vol. 11, no. 1, p. 2255026, 2023.

[174] M. S. Saghiv and M. S. Sagiv, *Basic exercise physiology*. Springer, 2020.

[175] A. Caporale, F. Gabriele Galizia, L. Botti, and C. Mora, "Thermal comfort prediction of aged industrial workers based on occupants' basal metabolic rate," *AHFE INTERNATIONAL,* vol. 65, 2022.

[176] J. Kim, R. de Dear, C. Candido, H. Zhang, and E. Arens, "Gender differences in office occupant perception of indoor environmental quality (IEQ)," *Building and environment,* vol. 70, pp. 245–256, 2013.

[177] D. Wyon, "The effects of moderate heat stress on typewriting performance," *Archives des sciences physiologiques,* vol. 27, no. 4, pp. 499–509, 1973.

[178] M. Chen *et al.*, "Local thermal response differences due to sex and BMI among older adults in warm environments," *Building and Environment,* p. 113275, 2025.

[179] C. P. P. Andreoli, M. M. F. De Martino, K. K. M. da Silva, and K. F. D. M. Pasetti, "Circadian rhythm of body temperature and chronotype in night shift students and workers," *International Archives of Medicine,* vol. 10, 2017.

[180] R. Schultz Martins *et al.*, "The clamping of end-tidal carbon dioxide does not influence cognitive function performance during moderate hyperthermia with or without skin temperature manipulation," *Frontiers in Psychology,* vol. 12, p. 788027, 2021.

[181] C. R. Monnard *et al.*, "Issues in continuous 24-h core body temperature monitoring in humans using an ingestible capsule telemetric sensor," *Frontiers in Endocrinology,* vol. 8, p. 130, 2017.

[182] C. M. Dolson *et al.*, "Wearable sensor technology to predict core body temperature: a systematic review," *Sensors,* vol. 22, no. 19, p. 7639, 2022.

[183] F. Nicol, *Thermal comfort: a handbook for field studies toward an adaptive model*. University ofEast London, 1993.

[184] D. J. Brake, "The deep body core temperatures, physical fatigue and fluid status of thermally stressed workers and the development of thermal work limit as an index of heat stress," Curtin University of Technology, 2002.

[185] D. Brajkovic and M. B. Ducharme, "Finger dexterity, skin temperature, and blood flow during auxiliary heating in the cold," *Journal of Applied Physiology,* vol. 95, no. 2, pp. 758–770, 2003.

[186] A. C. Russo, "Assessing the Systemic Impact of Heat Stress on Human Reliability in Mining Through FRAM and Hybrid Decision Models," *Mining,* vol. 5, no. 3, p. 50, 2025.

[187] D. Wang, H. Zhang, E. Arens, and C. Huizenga, "Observations of upper-extremity skin temperature and corresponding overall-body thermal sensations and comfort," *Building and Environment,* vol. 42, no. 12, pp. 3933–3943, 2007.

[188] N. B. Morris, M. N. Cramer, S. G. Hodder, G. Havenith, and O. Jay, "A comparison between the technical absorbent and ventilated capsule methods for measuring local sweat rate," *Journal of applied physiology,* vol. 114, no. 6, pp. 816–823, 2013.

[189] A. Jabbari, S. Grimnes, and Ø. G. Martinsen, "Electrodermal response–correlation between potential and conductance," in *13th International Conference on Electrical Bioimpedance and the 8th Conference on Electrical Impedance Tomography: ICEBI 2007, August 29th-September 2nd 2007, Graz, Austria*, 2007: Springer, pp. 747–750.

[190] T. Kosch, J. Karolus, H. Ha, and A. Schmidt, "Your skin resists: exploring electrodermal activity as workload indicator during manual assembly," in *Proceedings of the ACM SIGCHI symposium on engineering interactive computing systems*, 2019, pp. 1–5.

[191] G. Bini, K. Hagbarth, P. Hynninen, and B. Wallin, "Regional similarities and differences in thermoregulatory vaso-and sudomotor tone," *The journal of physiology,* vol. 306, no. 1, pp. 553–565, 1980.

[192] O. N. Rahma *et al.*, "Electrodermal activity for measuring cognitive and emotional stress level," *Journal of Medical Signals & Sensors,* vol. 12, no. 2, pp. 155–162, 2022.

[193] V. A. van Es, R. G. Lopata, E. P. Scilingo, and M. Nardelli, "Contactless cardiovascular assessment by imaging photoplethysmography: A comparison with wearable monitoring," *Sensors,* vol. 23, no. 3, p. 1505, 2023.

[194] J. Liu *et al.*, "Patient-specific oscillometric blood pressure measurement," *IEEE Transactions on Biomedical Engineering,* vol. 63, no. 6, pp. 1220–1228, 2015.

[195] L. B. Fich, M. Wallergård, Å. M. Hansen, and P. Jönsson, "Stress Hormones mediated by the Built Environment."

[196] D. J. Macfarlane, "Open-circuit respirometry: a historical review of portable gas analysis systems," *European journal of applied physiology,* vol. 117, no. 12, pp. 2369–2386, 2017.

[197] D. D. Gagnon, "The effects of exercising in the cold on energy metabolism, skeletal muscle tissue oxygenation and immuno-endocrine responses," *Studies in sport, physical education and health,* no. 206, 2014.

[198] N. Bobba-Alves, R.-P. Juster, and M. Picard, "The energetic cost of allostasis and allostatic load," *Psychoneuroendocrinology,* vol. 146, p. 105951, 2022.

[199] C. A. Balaras and A. Argiriou, "Infrared thermography for building diagnostics," *Energy and buildings,* vol. 34, no. 2, pp. 171–183, 2002.

[200] K. A. Herborn *et al.*, "Skin temperature reveals the intensity of acute stress," *Physiology & behavior,* vol. 152, pp. 225–230, 2015.

[201] A. Ghahramani, G. Castro, B. Becerik-Gerber, and X. Yu, "Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort," *Building and Environment,* vol. 109, pp. 1–11, 2016.

[202] F. Jazizadeh and W. Jung, "Personalized thermal comfort inference using RGB video images for distributed HVAC control," *Applied Energy,* vol. 220, pp. 829–841, 2018.

[203] D. Li, C. C. Menassa, and V. R. Kamat, "Non-intrusive interpretation of human thermal comfort through analysis of facial infrared thermography," *Energy and Buildings,* vol. 176, pp. 246–261, 2018.

[204] N. Manikandan, D. Manivannan, and M. Kavitha, "Stress Monitoring with Computer Vision and Machine Learning for Software Employees," in *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, 2024: IEEE, pp. 1016–1021.

[205] R. Li, D. Yang, F. Fang, K.-S. Hong, A. L. Reiss, and Y. Zhang, "Concurrent fNIRS and EEG for brain function investigation: a systematic, methodology-focused review," *Sensors,* vol. 22, no. 15, p. 5865, 2022.

[206] Y. Zhang *et al.*, "A neuroergonomics model for evaluating nuclear power plants operators' performance under heat stress driven by ECG time-frequency spectrums and fNIRS prefrontal cortex network: A CNN-GAT fusion model," *Advanced Engineering Informatics,* vol. 62, p. 102563, 2024.

[207] J. Langevin, "Longitudinal dataset of human-building interactions in US offices," *Scientific data,* vol. 6, no. 1, p. 288, 2019.

[208] O. Toyinbo, "Indoor environmental quality, pupils' health, and academic performance—a literature review," *Buildings,* vol. 13, no. 9, p. 2172, 2023.

[209] S. Ahmed, "Analysis of workplace surveillance in a quest for an ethical stance," *Journal of Law and Governance,* vol. 2, no. 4, pp. 15–26–15–26, 2007.

[210] A. Wierzbicka *et al.*, "Healthy indoor environments: The need for a holistic approach," *International journal of environmental research and public health,* vol. 15, no. 9, p. 1874, 2018.

[211] M. Luo, Z. Wang, G. Brager, B. Cao, and Y. Zhu, "Indoor climate experience, migration, and thermal comfort expectation in buildings," *Building and Environment,* vol. 141, pp. 262–272, 2018.

[212] L.-R. Jia, Q.-Y. Li, X. Chen, C.-C. Lee, and J. Han, "Indoor thermal and ventilation indicator on university students' overall comfort," *Buildings,* vol. 12, no. 11, p. 1921, 2022.

[213] T. F. o. t. E. S. o. C. t. N. A. S. o. P. Electrophysiology, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Circulation,* vol. 93, no. 5, pp. 1043–1065, 1996.

[214] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of neuroscience methods,* vol. 190, no. 1, pp. 80–91, 2010.

[215] S. K. Mishra, "Representation-constrained canonical correlation analysis: a hybridization of canonical correlation and principal component analyses," 2009.

[216] M. S. H. Onim and H. Thapliyal, "Predicting stress in older adults with rnn and lstm from time series sensor data and cortisol," in *2024 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2024: IEEE, pp. 300–306.

[217] Y.-H. H. Tsai, S. Bai, P. P. Liang, J. Z. Kolter, L.-P. Morency, and R. Salakhutdinov, "Multimodal transformer for unaligned multimodal language sequences," in *Proceedings of the conference. Association for computational linguistics. Meeting*, 2019, vol. 2019, p. 6558.

[218] M. T. Ribeiro, S. Singh, and C. Guestrin, "" Why should i trust you?" Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[219] P. ASHRAE Standing Standard Project Committee 55. Cognizant TC: 2.1 and h. environment, *Thermal environmental conditions for human occupancy*. Ashrae, 2023.

[220] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Advances in psychology*, vol. 52: Elsevier, 1988, pp. 139–183.

[221] A. Sreekrishnan *et al.*, "Intracerebral hemorrhage location and functional outcomes of patients: a systematic literature review and meta-analysis," *Neurocritical care,* vol. 25, no. 3, pp. 384–391, 2016.

[222] C. D. Speilberger, R. Gorsuch, R. Lushene, P. Vagg, and G. Jacobs, "Manual for the state-trait anxiety inventory," *Palo Alto, CA: Consulting Psychologists,* 1983.

[223] T. M. Marteau and H. Bekker, "The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI)," *British journal of clinical Psychology,* vol. 31, no. 3, pp. 301–306, 1992.

[224] A. Guillen-Riquelme and G. Buela-Casal, "Psychometric revision and differential item functioning in the State Trait Anxiety Inventory (STAI)," *Psicothema,* vol. 23, no. 3, pp. 510–515, 2011.

[225] T. Amano, N. Gerrett, Y. Inoue, T. Nishiyasu, G. Havenith, and N. Kondo, "Determination of the maximum rate of eccrine sweat glands' ion reabsorption using the galvanic skin conductance to local sweat rate relationship," *European journal of applied physiology,* vol. 116, no. 2, pp. 281–290, 2016.

[226] A. Greco *et al.*, "Acute stress state classification based on electrodermal activity modeling," *IEEE Transactions on Affective Computing,* vol. 14, no. 1, pp. 788–799, 2021.

[227] N. Pop-Jordanova and J. Pop-Jordanov, "Electrodermal activity and stress assessment," *Prilozi,* vol. 41, no. 2, pp. 5–15, 2020.

[228] H. G. van Lier *et al.*, "A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor," *Behavior research methods,* vol. 52, no. 2, pp. 607–629, 2020.

[229] D. Goel, A. Mahmoudi-Nejad, and M. Guzdial, "Label-free subjective player experience modelling via let's play videos," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2024, vol. 20, no. 1, pp. 46–53.

[230] B. Bühler, "Multimodal Machine Learning for Automated Assessment of Attention-Related Processes during Learning," *arXiv preprint arXiv:2407.05803,* 2024.

[231] P. Ferreira, A. Ruano, S. Silva, and E. Conceição, "Neural networks based predictive control for thermal comfort and energy savings in public buildings," *Energy and buildings,* vol. 55, pp. 238–251, 2012.

[232] B. Hency V, "Improving occupant comfort through real-time Predictive control of Indoor environment using Predicted-Mean Vote model and MLP," *arXiv preprint arXiv:2409.00080,* 2024.

[233] T. Hinz, N. Navarro-Guerrero, S. Magg, and S. Wermter, "Speeding up the hyperparameter optimization of deep convolutional neural networks," *International Journal of Computational Intelligence and Applications,* vol. 17, no. 02, p. 1850008, 2018.

[234] N. Robin, G. R. Coudevylle, and O. Hue, "Attentional processes and performance in hot humid or dry environments: review, applied recommendation and new research directions," *Movement & Sport Sciences-Science & Motricité,* vol. 112, no. 2, pp. 41–51, 2021.

[235] W. Luo, R. Kramer, Y. de Kort, P. Rense, and W. van Marken Lichtenbelt, "The effects of a novel personal comfort system on thermal comfort, physiology and perceived indoor environmental quality, and its health implications-Stimulating human thermoregulation without compromising thermal comfort," *Indoor Air,* vol. 32, no. 1, p. e12951, 2022.

[236] A. Gruszka, G. Matthews, and B. Szymura, *Handbook of individual differences in cognition: Attention, memory, and executive control*. Springer, 2010.

[237] G. Matthews *et al.*, "Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry," *Emotion,* vol. 2, no. 4, p. 315, 2002.

[238] D. P. Wyon, I. Wyon, and F. Norin, "Effects of moderate heat stress on driver vigilance in a moving vehicle," *Ergonomics,* vol. 39, no. 1, pp. 61–75, 1996.

[239] R. K. Mehta, "Interactive Effects of Physical and Mental Workload: A Study of Muscle Function, Capacity and Exertion Type," 2011.

[240] T. Cheung, S. Schiavon, T. Parkinson, P. Li, and G. Brager, "Analysis of the accuracy on PMV–PPD model using the ASHRAE Global Thermal Comfort Database II," *Building and Environment,* vol. 153, pp. 205–217, 2019.

[241] P. L. Bartlett, P. M. Long, G. Lugosi, and A. Tsigler, "Benign overfitting in linear regression," *Proceedings of the National Academy of Sciences,* vol. 117, no. 48, pp. 30063–30070, 2020.

[242] B. Lala, H. Rizk, S. M. Kala, and A. Hagishima, "Multi-task learning for concurrent prediction of thermal comfort, sensation and preference in winters," *Buildings,* vol. 12, no. 6, p. 750, 2022.

[243] A. Kendall, Y. Gal, and R. Cipolla, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7482–7491.

[244] A. M. Abbasi, M. Motamedzade, M. Aliabadi, R. Golmohammadi, and L. Tapak, "The impact of indoor air temperature on the executive functions of human brain and the physiological responses of body," *Health promotion perspectives,* vol. 9, no. 1, p. 55, 2019.

[245] J. González, E. Alcántara, A. Bataller, and A. García, "Physiological and subjective evaluation of footwear thermal response over time," in *Fifth symposium on footwear biomechanics, Zürich. Eds: Hennig, E*, 2001, pp. 40–41.

[246] S. Ioannou, V. Gallese, and A. Merla, "Thermal infrared imaging in psychophysiology: potentialities and limits," *Psychophysiology,* vol. 51, no. 10, pp. 951–963, 2014.

[247] W. Liu, Z. Lian, and B. Zhao, "A neural network evaluation model for individual thermal comfort," *Energy and Buildings,* vol. 39, no. 10, pp. 1115–1122, 2007.

[248] Y. Epstein, G. Keren, J. Moisseiev, O. Gasko, and S. Yachin, "Psychomotor deterioration during exposure to heat," *Aviation, space, and environmental medicine,* vol. 51, no. 6, pp. 607–610, 1980.

[249] T. Q. Nguyen, T. B. Vu, N. Shafiabady, T. T. Nguyen, and P. T. Nguyen, "Application of CNN and MLP models for structural health monitoring: A case study on Saigon Bridge," *Journal of Low Frequency Noise, Vibration and Active Control,* p. 14613484251332711, 2025.

[250] J. Van Hoof, M. Mazej, and J. L. Hensen, "Thermal comfort: research and practice," *Frontiers in Bioscience,* vol. 15, no. 2, pp. 765–788, 2010.

[251] F. Zhang, R. de Dear, and P. Hancock, "Effects of moderate thermal environments on cognitive performance: A multidisciplinary review," *Applied energy,* vol. 236, pp. 760–777, 2019.

[252] J. Ryu, J. Kim, W. Hong, and R. De Dear, "On the temporal dimension of adaptive thermal comfort mechanisms in residential buildings," in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 609, no. 4: IOP Publishing, p. 042071.

[253] M. Burzo, C. Wicaksono, M. Abouelenien, V. Perez-Rosas, R. Mihalcea, and Y. Tao, "Multimodal sensing of thermal discomfort for adaptive energy saving in buildings," in *iiSBE Net Zero Built Environment 2014 Symposium*, 2014.

[254] N. GHARAIBEH, "Enhancing interpretability in brain tumor detection: Leveraging Grad-CAM and SHAP for explainable AI in MRI-based cancer diagnosis," *Applied Computer Science,* vol. 21, no. 3, pp. 182–197, 2025.

[255] L. T. Wong, K. W. Mui, and C. Cheung, "Bayesian thermal comfort model," *Building and environment,* vol. 82, pp. 171–179, 2014.

[256] R. Sterling *et al.*, "A virtual test-bed for building Model Predictive Control developments," in *Modelica*, 2019, p. 157:001.

[257] K. Nkurikiyeyezu, A. Yokokubo, and G. Lopez, "Affect-aware thermal comfort provision in intelligent buildings," in *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, 2019: IEEE, pp. 331–336.

[258] B. Sahoh, C. Kaewrat, K. Yeranee, N. Kittiphattanabawon, and M. Kliangkhlao, "Causal AI-powered event interpretation: A cause-and-effect discovery for indoor thermal comfort measurements," *IEEE Internet of Things Journal,* vol. 9, no. 22, pp. 23188–23200, 2022.

[259] J. HE, Y. HE, J. HU, and Y. GUO, "Intelligent Governance: The AI-Driven New Paradigm of Governmental Adaptive Governance," *Journal of US-China Public Administration,* vol. 22, no. 1, pp. 1–27, 2025.