

[Inner endpaper]

HEC MONTRÉAL

**Establishing Convergent Validity of Physiological Measurement Devices for
Objective Employee Experience Assessment**

par

Camila Burga Villarreal

Sylvain Sénécal

HEC Montréal

Co-directeur de recherche

Constantinos Coursaris

HEC Montréal

Co-directeur de recherche

Sciences de la gestion

Spécialisation expérience utilisateur

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)*

Janvier 2026

© Camila Burga Villarreal, 2026

Résumé

Le déploiement de capteurs physiologiques portables en milieu professionnel a devancé leur validation dans des conditions écologiquement pertinentes, créant un écart qui compromet leur adoption pour la recherche sur l'expérience employé (EE). Cette étude répond à ce défi méthodologique en évaluant la validité convergente de deux technologies émergentes de mesure de la fréquence cardiaque (FC), le Research Ring (BIOPAC Systems Inc., États-Unis), une bague portable utilisant la photopléthysmographie, et FaceReader (Noldus Technology, Pays-Bas), un système de photopléthysmographie à distance basé sur caméra, comparées au Hexoskin Smart Shirt (Carré Technologies Inc., Canada), dispositif de référence basé sur l'électrocardiogramme (ECG). À l'aide d'un devis intra-sujets, trente-cinq participants ont complété un protocole de laboratoire écologiquement valide intégrant des exigences physiques, cognitives et émotionnelles représentatives des stressseurs contemporains du milieu de travail. Les analyses de validité convergente ont révélé une forte concordance entre le Research Ring et Hexoskin pour l'ensemble des conditions expérimentales, caractérisée par des différences d'amplitude négligeables, un biais de mesure minimal et une synchronisation temporelle robuste. À l'inverse, FaceReader a démontré une validité convergente inadéquate. Ces résultats établissent le Research Ring comme une alternative valide pour la recherche terrain en psychophysiologie ambulatoire, permettant l'évaluation objective du stress continu en milieu de travail. Cette validation permet la collecte de données physiologiques fiables en milieu de travail réel et soutient une optimisation de l'expérience employé fondée sur des données objectives.

Mots-clés : Validité convergente ; surveillance de la fréquence cardiaque ; photopléthysmographie ; capteurs portables ; expérience employé ; stress professionnel ; psychophysiologie ambulatoire ; validité écologique.

Méthodes de recherche : Design expérimental intra-sujets, surveillance physiologique simultanée multi-dispositifs, protocole de tâches écologiquement pertinentes, modélisation linéaire à effets mixtes, analyse de concordance de Bland-Altman, analyse de corrélation croisée temporelle, synchronisation des données.

Abstract

The deployment of wearable physiological sensors in occupational settings has outpaced their validation under ecologically relevant conditions, creating a gap that undermines their adoption for employee experience (EX) research. This study addresses this methodological challenge by evaluating the convergent validity of two emerging heart rate (HR) monitoring technologies; the Research Ring (BIOPAC Systems Inc., USA), a wearable photoplethysmography ring, and FaceReader (Noldus Technology, Netherlands), a camera-based remote photoplethysmography, against the Hexoskin Smart Shirt (Carré Technologies Inc., Canada), an ECG-based reference device. Using a within-subjects design, thirty-five participants (18 women, 17 men; Mage = 27.0 years, SD = 8.7) completed an ecologically valid laboratory protocol incorporating physical, cognitive, and emotional task demands representative of contemporary workplace stressors. Convergent validity analyses demonstrated strong agreement between Research Ring and Hexoskin across all experimental conditions, characterized by negligible amplitude differences, minimal measurement bias, and robust temporal synchronization. In contrast, FaceReader demonstrated inadequate convergent validity, with substantial amplitude discrepancies and weak temporal synchronization relative to the reference device. These findings establish the Research Ring as a valid alternative for ambulatory psychophysiology field research, providing the empirical foundation necessary for future workplace studies to objectively assess continuous stress in workers. This validation enables researchers to collect reliable physiological data in actual work environments, ultimately informing data-driven employee experience optimization strategies based on objective evidence rather than subjective report alone.

Keywords: Convergent Validity; Heart Rate Monitoring; Photoplethysmography, Wearable Sensors, Employee Experience (EX), Occupational Stress, Ambulatory Psychophysiology, Ecological Validity.

Research Methods: Within-subjects experimental design, concurrent multi-device physiological monitoring, ecologically relevant task protocol, linear mixed-effects modeling, Bland-Altman agreement analysis, cross-correlation temporal analysis data synchronization platform.

Table of contents

Résumé	i
Abstract	ii
Table of contents	iii
List of tables	vi
List of figures	vi
List of abbreviations and acronyms	vii
Preface	viii
Acknowledgements	ix
Chapter 1: Introduction	1
1. Occupational Stress and Its Organizational Consequences	1
2. Limitations of Current Stress Measurement Approaches	2
3. The Complexity of Capturing Stress in Real Work Environments	4
4. The UX Challenge: Designing Without Physiological Insight	5
5. Physiological Measurement as a Methodological Opportunity	6
6. The Validation Gap in Wearable Psychophysiological Tools	7
7. The Present Study and Its Methodological Contributions	7
Chapter 2: Convergent Validity of Wearable and Camera-Based Heart Rate Measurement Under Simulated Workplace Conditions	10
Abstract	10
1. Introduction	10
2. Methods	13
2.1 Participants and design	13
2.2 Physiological measurement devices	14

2.3 Procedure	15
2.4. Stimulus material	17
2.5. Data collection, data synchronization and statistical analysis	20
3. Results	22
3.1 Descriptive statistics comparing the three physiological measurement devices across all task conditions	22
3.2 Visual representation of signal quality across all task conditions	23
3.3 Convergent validity between Research Ring and Hexoskin	24
3.4 Convergent validity between Hexoskin and FaceReader across all task conditions	29
3.5 Summary of device agreement and convergent validity results across all tasks	30
4. Discussion	30
4.1 Convergent validity of wearable and camera-based HR sensors	31
4.2 Sensitivity of HR to physical, cognitive, and emotional demands	31
4.3 Methodological contributions	32
4.4 Industry implications	34
4.5 Limitations and future directions	34
5. Conclusion	35
References	37
Chapter 3: Conclusion	44
1. The Convergent Validity Requirement for Workplace Psychophysiology	44
2. Objective Measurement in Previously Inaccessible Populations	44
3. Practical, Methodological, and Ethical Constraints	45
4. A Framework for Responsible Implementation	45
5. Pathway Forward: From Laboratory Validation to Field Research	46
Bibliography	47

List of tables

Table 1. Author's Contribution.....	9
Table 2. Descriptive statistics of mean heart rate across all conditions	23

List of figures

Figure 1. Experimental procedure.....	17
Figure 2. Emotional stroop task	18
Figure 3. N-Back task	19
Figure 4. HR signal quality and temporal alignment of Hexoskin, Research Ring, and FaceReader across all tasks for one participant (representative case)	24
Figure 5. Bland-Altman plot demonstrating HR agreement between Hexoskin and Research Ring	25
Figure 6. Cross correlation between RR & HS	26
Figure 7. Mean HR across physical conditions for Hexoskin, Research Ring, and FaceReader	27
Figure 8. Mean HR across cognitive tasks for Hexoskin, Research Ring, and FaceReader.....	28
Figure 9. Mean HR across emotionally demanding tasks for Hexoskin, Research Ring, and FaceReader.....	28
Figure 10. Bland-Altman plot showing HR disagreement between Hexoskin and FaceReader across all tasks.....	29
Figure 11. Cross-correlation between Hexoskin and FaceReader across all tasks	30

List of abbreviations and acronyms

ANS: Autonomic Nervous System

HCI: Human-Computer Interaction

Heavy_M: Operating Heavy Machinery Simulator

HR: Heart Rate

HR_m: Mean Heart Rate

HS: Hexoskin

ICT: Information and Communication Technologies

RR: Research Ring

FR: FaceReader

Preface

This present work was completed as part of my Master of Science in User Experience at HEC Montreal. This thesis has been the object of evaluation and approval by the administrative management of the M.Sc. program at HEC Montreal. To ensure compliance with requirements outlined in the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, all phases of research performed as part of this thesis went through an extensive ethical review prior to beginning data collection. Approval was granted by the Research Ethics Board at HEC Montreal under the certificate number 2025-6408 (Appendix A). Accordingly, all stages of research involving humans as participants were done in an ethical way, always with informed consent obtained from participants prior to starting the experiment. Additionally, the authorization to write this thesis in the form of articles, in English, was provided by the program director. Hence, this thesis encompasses 3 chapters. While the first and last chapters introduce the topic and reiterate important research takeaways, respectively, middlemost chapters present a scientific article.

Acknowledgements

My first acknowledgements go to Sylvain Sénécal and Constantinos Coursaris, co-directors of my master's thesis. Sylvain's openness provided a space for reflection and problem-solving whenever I needed it, while Dr. C.'s rigor consistently pushed me to think more methodologically and precisely. I am also grateful to Pierre-Majorique Léger and Wietske van Osch for the structure they brought to my research and writing and for the resources they shared along the way.

I am very grateful to my parents, Gina and Cesar, for the huge sacrifice they made in immigrating to Canada to build a better future for our family, while also ensuring that our Peruvian culture continued to flourish at home. Witnessing you rebuild our lives from the ground up while studying, working, and learning two new languages has been (& will continue to be) a constant source of inspiration. *Querer es poder*. Thank you for your unconditional love and support.

I would also like to thank my partner, my brother, and my friends, Jimmy, Mateo, Léane, Nicolas, Maurane, Florence, Julia, and Ève, for being present throughout this intense research journey. Your curiosity and care made me feel supported, even when this journey felt far from your own.

A special thanks goes to Hajar, a friend I was lucky to meet through this program. She has been a true pillar throughout this journey. Whether I felt happy or uncertain about my research, Hajar understood me better than anyone else. We spent countless study sessions and FaceTime calls working side by side on our projects. Our breaks for deep talks and great laughs were much needed.

I would also like to express my gratitude to my research partner, Mehdi, with whom I collaborated at every stage of this project and whose insights helped make the research process both rigorous and enjoyable. His patience and serenity were grounding throughout our work together.

Finally, I would like to thank the Tech3Lab team, Shang Lin, François, Salima, Xavier, Frédérique, David, Alejandra, Thaddé, Flavien, Nejla and Baptiste for making me feel included as part of the Tech3Lab team. You shared valuable resources, asked challenging questions, helped troubleshoot issues as they arose, supported my analyses, and offered unique perspectives, all of which greatly enriched the quality and rigor of this research. I am also grateful for Aurélie's support during her internship at the lab; our data collection would not have been as smooth without her.

Chapter 1: Introduction

1. Occupational Stress and Its Organizational Consequences

Occupational stress is widely recognized as a major organizational challenge and a growing global societal problem (Alberdi et al., 2016; Brassey et al., 2022). It occurs when job demands do not match the person's adaptive resources (Schaufeli & Buunk, 2003). It is consistently emerging as a major risk factor for physical and mental illness, poor work performance, and low job satisfaction (Levine et al.; 2009). Stress is estimated to be responsible for 50-60% of all lost working days in enterprises and causes huge economic losses in companies (Alberdi et al., 2016).

Burnout is a unique affective response to occupational stress (Brassey et al., 2022). It develops over time as employees face continuous and prolonged exposure to work-related stress (Melamed et al., 2006), particularly when the demands of their job exceed their available resources or ability to cope (Koutsimani et al.; 2019). According to Maslach et al. (1996), burnout is a construct composed of three dimensions: (1) exhaustion, (2) cynicism, and (3) lack of professional efficacy. Exhaustion, refers to feelings of stress, particularly chronic fatigue due to excessive work demands; cynicism, refers to a detached attitude towards work in general and colleagues, which leads to a loss of interest in work; lack of professional efficiency, refers to reduced perceived efficiency, successful attainment and accomplishment both in the individual's job and the organization (Koutsimani et al., 2019). A key distinction between burnout and stress is that burnout includes the development of negative attitudes and behaviours toward the job and the organization, whereas stress is not necessarily accompanied by such attitudes and behaviours (Maslach, 1993).

Burnout prevalence in Canada is substantial: A recent BCG study revealed that 52% of the Canadian workforce reports experiencing burnout symptoms, with blue-collar workers experiencing up to 26% higher burnout rates than desk-based employees (Novacek et al., 2024). The organizational consequences are severe and quantifiable. Burned-out employees are six times more likely to intend to leave their organization within the next three to six months (Brassey et al., 2022). When they do leave, replacement costs typically range from 50% to 200% of annual salary (Dyerly, 2025), compounded by hidden costs of absenteeism, disengagement, and performance degradation (Brassey et al., 2022). Beyond higher turnover rates, burnout directly undermines

workplace safety and productivity, since it also introduces a concept called presenteeism, which means employees attend the workplace but work at diminished productivity (Alberdi et al., 2016). Regarding the individual level, stressed workers show reduced cognitive function, poorer decision-making, and increased error rates, particularly consequential in high-risk environments (Wu et al., 2019). When it comes to the organizational level, turnover rates have negative implications for several dimensions of organizational performance, including safety, productivity, and monetary outcomes (Shaw, 2011). Additionally, turnover can direct organizational attention away from core tasks like safety and maintenance, leading to increased accidents, injuries, and other failures, which ultimately lowers performance (Shaw, 2011).

Notably, this problem is not evenly distributed. Blue-collar workers, those whose roles require continuous physical presence, tend to have a lower employee experience than white-collar workers in administrative or desk-based roles (Contreras-Cruz et al., 2023). This gap is reflected in lower employee satisfaction and higher turnover rates among blue-collar workers compared to their white-collar counterparts (Contreras-Cruz et al., 2023). Contributing factors are that blue-collar roles lack the flexibility, remote work options, and autonomy that increasingly shape positive work experience in office-based jobs. Instead, these workers encounter a set of stressors, including irregular working hours, physically demanding tasks, hazardous environments, and reduced access to organizational interventions designed to mitigate burnout (flexible schedules, remote work, workspace autonomy). As a result, blue-collar workers are 1.5 times more likely to self-report high perceived stress compared to white-collar workers, a pattern attributed in part to the continuous physical activity required in their roles (Dèdelè et al., 2019).

2. Limitations of Current Stress Measurement Approaches

To address employee stress and burnout, organizations must first measure it. Yet, there is a measurement challenge: the tools currently deployed to assess employee stress, burnout, and overall experience are limited, subjective, and prone to bias, which struggle to capture the continuous, dynamic, and complex nature of work-related stress.

Currently, organizations rely primarily on three measurement categories: (1) self-report surveys, (2) interviews, and (3) organizational metrics. Each of these approaches offers value, but all suffer from important limitations. First, self-report surveys are designed to capture workers' perceived

experience, that is, the conscious interpretation a person can articulate and reflect upon. These tools typically rely on perceptual scales administered through questionnaires. However, perception is vulnerable to cognitive biases, such as recall bias, which is a type of bias that occurs when individuals, in this case, employees, do not accurately remember a past event or experience or leave out details when reporting about them (NCI, 2011). Nevertheless, self-report questionnaires remain the most widely used method for measuring constructs such as perceived stress, cognitive load, and engagement (Ortiz de Guinea et al., 2013; Anmarkrud et al., 2019; Dédelé et al., 2019). Widely adopted tools include the Perceived Stress Scale (PSS) (Cohen et al., 1983) and the Doetinchem Organizational Stress Questionnaire (VOS-D), an 81-item instrument assessing organizational stressors, psychological tension, health complaints, and social changes (Öztürk Eyimaya et al., 2021). These tools ask participants to rate experiences retrospectively.

Second, interviews offer rich qualitative insights by allowing employees to describe, in their own words, their experiences, needs, and stressors. They are particularly valuable for uncovering pain points, motivations, and contextual nuances while fostering empathy and enabling flexible, adaptive inquiry. However, interviews share the same core limitation as surveys: they rely on self-reports. Consequently, they are susceptible to common method bias, social desirability bias, interviewer effects, and subjective interpretation (Podsakoff et al., 2003; Rield et al., 2014). Therefore, interviews cannot capture emotional or physiological states that do not reach conscious awareness, a drawback when studying stress in the workplace.

Third, organizational metrics such as turnover rates, absenteeism, productivity indicators, and financial performance data are frequently used as proxy indicators to infer workplace stress. However, these measures do not assess stress directly; rather, they reflect the broader health and functioning of the organization (Shaw, 2011). From a managerial perspective, these indicators tend to command leadership attention because deteriorations in productivity, safety outcomes, or financial performance are both visible and economically consequential. Empirical research suggests, for example, that higher turnover is associated with lower productivity, reduced workplace safety, and weaker financial performance (Shaw, 2011). When these indicators begin to deteriorate, organizations frequently look for underlying causes, and employee stress or burnout is often among the factors identified. The limitation of relying on organizational metrics is that they are reactive rather than diagnostic. They capture the downstream consequences of stress once

issues have already accumulated, but they offer little insight into how stress develops or how employees experience it in real time. As a result, organizational indicators reveal that a problem exists, but not how or when it emerged.

Moreover, the dynamic nature of workplace stress further complicates reliance on reactive measurement approaches. Increasing levels of organizational change can intensify stress, and employees experiencing sustained stress often become less healthy and less effective in their roles. Reduced effectiveness can, in turn, reinforce stress, creating a self-perpetuating cycle that is difficult for organizations to interrupt (Hunt & Ulrich, 2022). This dynamic underscores a central limitation of current measurement approaches. Whether based on surveys, interviews, or organizational indicators, all these methods capture perception or downstream outcomes rather than the objective physiological or behavioral reality of workers across blue and white-collar contexts. Taken together, these approaches inform organizations about how employees interpret their capacity to cope or how stress becomes visible after the fact, yet they remain unable to access the underlying physiological and behavioral states through which work-related stress is experienced and evolves in real time.

3. The Complexity of Capturing Stress in Real Work Environments

Beyond the limitations of current measurement tools, a deeper challenge arises from the nature of contemporary work itself. The modern workplace is multidimensional and physiologically complex in ways that subjective reporting cannot fully capture (Alberdi et al., 2016). Employees do not experience stress in a single form; instead, their daily activities blend physical, cognitive, emotional, and contextual demands that interact continuously throughout the workday.

Work can be physically demanding, requiring employees to walk, lift, operate machinery, or navigate dynamic environments, among other physical demands. These activities generate physiological responses that are independent of psychological stress yet easily confounded with it when measured only through perception. However, physical demands represent only one source of physiological activation in the workplace. Many roles are also cognitively demanding, requiring sustained attention, rapid information processing, frequent task-switching, and decision-making under uncertainty. These cognitive pressures activate neural and autonomic responses that differ from those produced by physical strain.

Emotional demands add an additional layer of complexity. In many roles, employees must continuously regulate their emotions while carrying out their tasks, interacting with others, or responding to situational demands, a process known to elicit physiological responses distinct from those associated with physical exertion or cognitive load. The work environment itself further shapes these responses. Factors such as noise levels, time pressure, and spatial constraints can modulate physiological activation throughout the workday in ways that traditional measurement tools often overlook or oversimplify.

As a result, when an employee reports feeling “stressed” on a survey, it is rarely clear what has given rise to that experience. Stress may reflect physical demands, cognitive load, emotional regulation, environmental pressures, or the interaction of several of these factors as work unfolds. Without access to how the body responds across these overlapping demands in real time, organizations remain limited in their ability to identify the true sources of employee stress, making it difficult to design employee experience (EX) interventions that address underlying causes rather than surface-level symptoms.

4. The UX Challenge: Designing Without Physiological Insight

User experience (UX) methodology is grounded in a core principle: understand the user’s lived reality before designing solutions. In digital product design, UX researchers use multiple methods, such as usability testing, eye tracking, and behavioral analytics, to access dimensions of experience that users cannot fully articulate in interviews. Employee experience (EX) design should follow the same principle. However, current EX research is almost entirely dependent on what employees report about their experience or what HR systems record as outcomes. This creates a critical blind spot: UX practitioners cannot analyze what employees are experiencing physiologically as work unfolds. They lack access to the physiological indices of stress, cognitive load, and emotional regulation, the real-time physiological reality of how work feels in the body. Consequently, EX interventions are designed based on limited, subjective information. Without access to physiological data, organizations cannot distinguish between perceived stress and actual physiological burden, nor can they identify which specific workplace demands generate the greatest integrated strain across physical, cognitive, and emotional dimensions for different workers. This absence of physiological visibility is not simply a data gap; it directly constrains evidence-based EX design. When experience is inferred only from retrospective reports or

downstream metrics, the most immediate and embodied dimensions of work remain inaccessible. Addressing this blind spot therefore requires expanding EX research beyond self-report and outcome indicators toward objective measures capable of capturing how work is experienced in the body as it unfolds.

5. Physiological Measurement as a Methodological Opportunity

A complementary measurement approach has emerged: psychophysiological monitoring, the continuous, objective measurement of physiological processes, such as heart rate, heart rate variability, skin conductance, respiration, that reflect underlying cognitive, emotional, and autonomic states (Ikehara & Crosby, 2005).

This represents a methodological innovation for UX research. By integrating physiological measurement into EX research protocols, UX practitioners can access dimensions of employee experience that self-report and behavioral observation cannot provide, such as stress responses, cognitive load patterns, emotional regulation capacity, and the real-time interaction of multiple simultaneous demands. This creates new evidence base for designing work environments that respect human physiological limits.

Heart rate is particularly valuable for this purpose. It reflects cardiac activation resulting from the interplay between sympathetic and parasympathetic branches (Kumar et al., 2007), and it increases in response to physical exertion, cognitive workload, and emotional stress (Krantz & Manuck, 1984; Hao et al., 2022). This sensitivity to multiple demand types is crucial: HR captures the integrated experience of work, not separated into “physical stress” or “cognitive stress” or “emotional stress”, but as the complex, simultaneous burden that employees actually experience.

For UX practitioners designing for blue-collar workers specifically, this methodological opportunity is particularly significant. Blue-collar workers, those whose roles prevent dedicated survey time and whose work simultaneously demands physical exertion and sustained cognitive attention, are precisely those whose experience has been most invisible to traditional UX research. Physiological measurement enables access to their lived experience without requiring time for surveys, articulate self-reflection, or conscious awareness of their own states. It reveals what their bodies know about workplace demands, independent of their ability to report it.

The promise is methodologically transformative: rather than designing EX based on what employees think they experience or what organizations record as outcomes, UX practitioners could design based on objective data about what employees actually experience physiologically as work occurs. This would advance EX design from generalized assumptions about worker needs to evidence-based, physiologically-informed interventions targeted at real demands.

6. The Validation Gap in Wearable Psychophysiological Tools

This methodological opportunity cannot be realized without addressing a fundamental credibility problem. Medical-grade physiological tools (chest-worn ECG systems) provide highly validated, accurate measurements, but they are too intrusive for workplace deployment. Wearable alternatives promise the practicality this research requires: unobtrusive, continuous measurement that does not disrupt work.

The issue is that deployment of these wearables has significantly outpaced their validation (van Lier et al., 2020). Most validation studies test wearable sensors under highly controlled laboratory conditions, participants seated, at rest, in climate-controlled environments with stable lighting (Villar et al., 2015). While devices demonstrate acceptable accuracy under these static conditions, workplace realities present categorically different challenges. For instance, blue-collar workers move continuously, generating heart rate elevations through physical exertion independent of psychological stress. This creates a credibility problem for UX research. If physiological data informs workplace design decisions, that data must be reliable under the actual conditions where research occurs (Schoenmakers et al., 2025). Without validation in dynamic, realistic conditions, design insights rest on unexamined assumptions about data quality. This validation gap, between laboratory-tested devices and realistic workplace applications, must be resolved before physiological measurement becomes a trustworthy UX research method.

7. The Present Study and Its Methodological Contributions

This thesis addresses this validation gap by examining whether physiological measurement devices can maintain reliable performance under the dynamic, complex conditions where UX researchers need to use them.

Rather than validating devices in isolation, we created an ecologically relevant laboratory protocol that maintains experimental control while introducing workplace physiological complexity. We tested two novel devices; Research Ring (BIOPAC Systems Inc., USA), a wearable photoplethysmography ring; and FaceReader's rPPG (Noldus Technology, Netherlands), a camera-based remote photoplethysmography system, against Hexoskin Smart Shirt (Carré Technologies Inc., Canada), a validated ECG reference during tasks spanning three dimensions of workplace complexity: physical demands (sitting, standing, walking, machinery operation), cognitive demands (variable working memory loads), and emotional demands (emotional interference tasks).

The research question is both practical and methodologically foundational for UX: What is the convergent validity of two novel physiological devices (a PPG-based Research Ring and a camera-based rPPG system) relative to a validated ECG reference (Hexoskin) when measuring HR across conditions simulating a dynamic work environment?

This thesis makes three contributions to the literature. First, it extends PPG-ECG validity research beyond traditional stationary laboratory paradigms by incorporating physical, cognitive, and emotional tasks that approximate real occupational contexts. Second, it introduces a reproducible validation blueprint integrating three complementary workplace demand dimensions: (1) dynamic physical movement, (2) cognitive load variation, and (3) emotional interference, addressing a gap in existing guidelines that rarely specify which experimental tasks capture real-world occupational psychophysiological variability. Third, it establishes methodological foundations for ambulatory psychophysiology by bridging controlled laboratory validation with ecologically relevant task demands, providing the translational pathway necessary for researchers to confidently select physiological measurement tools for future occupational studies.

By answering the research question and delivering these contributions, this thesis establishes the empirical foundation necessary for physiological measurement to become a reliable, trustworthy method in the UX researcher's toolkit for understanding and designing employee experience.

The present research follows an article-based structure and is situated within a scientific publication trajectory. The main article presented in Chapter 2 is intended for submission to a peer-reviewed journal publishing applied psychophysiology research, including cardiovascular

psychophysiology, and examining the integration of physiological measures with behavioural responses in real-world contexts, such as the *International Journal of Psychophysiology*.

Table 1. *Author's Contribution*

Research Process	Student Contribution
Literature Review	100% Literature Review: Read over 50 scientific on previous research related to validation of wearable device for ambulatory research, cardiovascular psychophysiology and convergent validity to assess optimal methodologies, select appropriate validated psychological paradigms and identify current gaps in literature.
Research Question	80% Research Question: Formulate a comprehensive research question in response to the identified gap in the literature.
Experimental Design	30% Ethics Approval: Prepare the required documentation for the Research Ethics Board (REB). Help was provided by my co-directors of research for a final review of the application. 35% Research Protocol: Develop the research protocol to match my research objectives. Help was provided by Tech3Lab's operations team regarding the template of the research protocol. 40% Questionnaire: Develop the online Qualtrics questionnaires and formulate interview questions.
Recruitment	50% Participants Recruitment: Create inclusion and exclusion criteria for participants' recruitment. The operations team at Tech3Lab oversaw and applied the determined guidelines on the HEC panel and were also in charge of giving the compensation.
Data Collection	40% Data Collection: Moderate the experiment and conducted semi-structured interviews after the experiment. Guided research assistants in verifying physiological signals from physiological measurement devices, stimuli launch and technical setup.
Statistical Analysis	50% Data Analysis: Analyze data using SAS to test hypothesis. Led the type of statistical tests needed. Revision and format of the data file for analysis statistics was done by Tech3Lab's statistician.
Thesis Writing	100% Thesis Writing: Write this thesis by detailing all phases and implications of this project. Guidance, review and feedback were provided by research codirectors throughout the writing process.

Chapter 2: Convergent Validity of Wearable and Camera-Based Heart Rate Measurement Under Simulated Workplace Conditions

Abstract

The growing use of physiological sensors in workplace research has created new opportunities for ambulatory psychophysiology, yet the validity of commercial devices under realistic task conditions remains uncertain. This study evaluated the convergent validity of two novel heart rate (HR) measurement devices, the Research Ring (BIOPAC Systems Inc., USA), which uses photoplethysmography (PPG), and FaceReader's rPPG module (Noldus Technology, Netherlands), against a chest-worn electrocardiographic reference, the Hexoskin Smart Shirt (Carré Technologies Inc., Canada). Hexoskin was selected because its ECG-derived HR has been validated in controlled indoor settings with stable electrode contact, conditions that aligned with our laboratory protocol. Thirty-five participants completed physical, cognitive, and emotionally demanding tasks representative of blue and white-collar work, while all devices recorded HR concurrently. Mean HR was computed per condition, and convergent validity was assessed through linear mixed-effects models, Bland-Altman analyses, and cross-correlation functions. Under these ecologically relevant laboratory conditions, the Research Ring produced HR estimates comparable to Hexoskin, whereas FaceReader's rPPG did not achieve acceptable validity. These findings provide a methodological foundation for future research in ambulatory psychophysiology and workplace monitoring.

Keywords: Convergent validity; Heart rate; Electrocardiogram (ECG), Photoplethysmography (PPG); Remote photoplethysmography (rPPG); Wearable sensors; Ambulatory psychophysiology; Ecologically relevant tasks; Occupational health monitoring; Cognitive load.

1. Introduction

Every day, workers across diverse industries operate in demanding environments where sustained attention and cognitive performance are critical. Blue-collar workers, defined as individuals whose work requires physical skill rather than office-based activity, face both mental and physical challenges (Cambridge Dictionary, 2025). In warehouses and logistics centers, forklift operators navigate busy spaces, while delivery drivers manage complex routes under time pressure. On

factory floors, assembly-line workers must maintain concentration through repetitive and safety-critical tasks. White-collar workers, defined by Cambridge Dictionary as individuals whose work primarily requires cognitive rather than physical effort, encounter comparable mental demands (2025): air traffic controllers continuously monitor multiple aircraft, financial traders make split-second decisions, and software developers sustain attention through extended coding sessions. Across these occupational groups, sustained cognitive load can lead to fatigue, elevated stress, and increased risk of error, ultimately affecting well-being, safety, and performance.

Understanding and monitoring workers' cognitive states could therefore play a vital role in enhancing both workplace safety and employee experience. Real-time measurement of cognitive load and stress in dynamic work environments requires continuous physiological monitoring that is valid and non-disruptive. Psychophysiology offers a promising approach: by measuring physiological activity continuously and objectively, it enables researchers to infer underlying cognitive and emotional processes as individuals perform complex tasks (Bañuelos-Lozoya et al., 2021). Among physiological systems, cardiovascular psychophysiology has proven particularly informative for assessing responses to cognitive and emotional demands (Obrist, 2012). This domain examines cardiac activity, blood flow, and blood pressure as indicators of autonomic nervous system activation. Heart rate (HR), defined as the number of times a person's heart beats in a minute, reflects overall cardiac activation resulting from the interplay between sympathetic and parasympathetic branches (Kumar et al., 2007). Heart rate acceleration can be due to increased sympathetic drive or parasympathetic withdrawal (Krantz & Manuck, 1984). More specifically, HR increases with physical exertion, cognitive workload, and emotional stress, making it a robust index of arousal and mental effort (Hao et al., 2022). In occupational contexts, elevated HR relative to baseline indicates that workers are experiencing heightened physical or cognitive demands (Akbar et al., 2019). Consequently, cardiac activity is the most widely used physiological indicator of mental workload (Rowe et al., 1998; Kumar et al., 2007).

A major methodological challenge, however, lies in ensuring that novel physiological sensors, often in the form of wearable or camera-based devices, provide valid measurements comparable to established reference systems (van Lier et al., 2020). More specifically, wearable devices based on photoplethysmography (PPG) and camera-based tools using remote photoplethysmography (rPPG) must be validated against reference devices (RD) to determine whether they can reliably

capture HR signals under conditions that reflect real-world workplace contexts rather than only in controlled, motion-restricted laboratory settings. Validity demonstrated solely in static, no-movement environments offer limited practical industry value if the same technology fails to perform accurately when users move, interact with tools, or engage in cognitively and emotionally demanding tasks typical of modern workplaces.

Establishing convergent validity is therefore essential before adopting new physiological devices for research or applied monitoring. Convergent validity is achieved when two instruments intended to measure the same construct yield highly correlated and statistically consistent results under the same conditions (Straub et al., 2004). In the context of this study, the focus is on HR measurements. Without demonstrated convergent validity between novel physiological measurement devices and reference systems in ecologically relevant conditions, large-scale adoption of emerging technologies entails financial and operational risks. Ensuring validity beyond the laboratory is thus a prerequisite for recommending these tools to industrial partners seeking to do research on employee experience.

Accordingly, this study addresses the following research question: What is the convergent validity of two novel physiological devices (a PPG-based Research Ring and a camera-based rPPG system) relative to a validated ECG reference (Hexoskin) when measuring HR across conditions simulating a dynamic work environment?

The present study addresses this need by evaluating the convergent validity of two novel HR measurement tools, the Research Ring (BIOPAC Systems Inc.), a wearable PPG-based sensor, and FaceReader's rPPG (Noldus Technology, Netherlands), against a validated reference device, the Hexoskin Smart Shirt (Carré Technologies Inc., Canada). By testing these instruments concurrently across ecologically relevant conditions of the workplace, including physical activity (sitting, standing, walking, and operating a heavy-machinery simulator), cognitively demanding tasks (N-Back), and emotionally evocative tasks (Emotional Stroop), this present study evaluates whether HR signals from novel sensors correspond to those obtained from a reference device. Examining convergence across these varied task categories allows assessment of each device's sensitivity not only to physical exertion but also to subtler physiological changes associated with mental workload and emotional engagement.

The findings have implications for both psychophysiological research and applied workplace monitoring. For researchers in Human-Computer Interaction (HCI), validated physiological measurements for ambulatory research enable the collection of reliable physiological data in naturalistic settings, thereby increasing ecological validity. For practitioners in industrial and organizational contexts, establishing convergent validity provides confidence that new monitoring tools yield accurate and interpretable measures of cognitive and emotional load. By validating novel physiological measurement devices under conditions that mirror authentic workplace demands, this study establishes an empirical foundation for the reliable deployment of next-generation sensors in organizational and field research.

2. Methods

2.1 Participants and design

Thirty-five participants (18 women, 17 men; $M = 27$ years, $SD = 8.7$; range = 18-60 years) completed the experiment. Inclusion criteria required fluency in French or English, aged 18 years or older, no reported skin allergies or sensitivities, and no history of epilepsy. Participants were recruited from our institution's participant pool and received financial compensation (CAD \$40) for their time. The study protocol was approved by our institution's Research Ethics Board (certificate #2025-6408).

The study employed a one-factor within-subjects design where each participant completed all workplace-relevant conditions, including varying cognitive load, exposure to emotional interference stimuli, different physical demands (sitting, standing, walking), and simulated heavy-machinery operation, while wearing all physiological measurement devices simultaneously. This within-subjects approach was deliberately selected to optimize convergent validity assessment by ensuring that physiological responses recorded by each device were obtained under identical psychological and physical conditions, thereby isolating device-related variance. Moreover, exposing participants to overlapping cognitive, emotional, and physical demands was intended to approximate ecologically valid workplace environments, where motion artifacts, task switching, and concurrent demand sources are typical features of real work environments. This approach enabled two key advantages: (1) intra-individual comparison of physiological and behavioral

responses across conditions, and (2) control for inter-individual variability in baseline physiology and task sensitivity. The total session duration was approximately 120 minutes.

2.2 Physiological measurement devices

Heart rate (HR) was selected as the primary physiological outcome because it reflects cardiac activation resulting from the dynamic interplay between sympathetic and parasympathetic branches of the autonomic nervous system (Kumar et al., 2007). HR increases in response to physical exertion, cognitive workload, and emotional stress (Krantz & Manuck, 1984; Hao et al., 2022), making it particularly suitable for a research protocol involving simultaneous and varying demand types. In the present study, where participants were exposed to multitasking conditions combining cognitive load, emotional interference, and task switching, HR served as an integrated indicator of the physiological burden associated with task demands.

Three physiological measurement devices were used to capture heart rate data. All three devices recorded simultaneously throughout the experimental session. Simultaneous acquisition across modalities enabled direct within-subject comparison of cardiac signals under identical task demands, thereby strengthening the internal validity of convergent analyses. HR was calculated as beats per minute and averaged across each condition to produce mean HR values for subsequent statistical comparison.

2.2.1 Hexoskin Smart Shirt (Carré Technologies Inc., Canada), reference device (RD)

The Hexoskin Smart Shirt (HS; Carré Technologies Inc., Canada) served as the reference device, capturing ECG signals at 256 Hz via three integrated textile electrodes (two chest, one lateral rib cage). Its validity has been established across multiple studies: Villar et al. (2015) reported narrow limits of agreement (-3.2 to +2.9 bpm) with clinical ECG during lying, sitting, standing, and walking; Cherif et al. (2018) demonstrated near-perfect concordance ($\rho_c = 0.98-0.99$) with 12-lead ECG during maximal exercise testing; and Schoenmakers et al. (2025) rated it highly suitable for short-term laboratory research. This validation profile confirms Hexoskin's accuracy for ECG-based HR measurement in controlled indoor settings with stable electrode contact, conditions matching our study design. Participants' thorax, navel, and chest circumference were measured post-consent to ensure optimal vest fit and electrode stability.

2.2.2 Research Ring (BIOPAC Systems Inc., USA): Novel wearable HR measurement device

The Research Ring (RR; BIOPAC Systems Inc., USA) recorded photoplethysmogram (PPG) signals at 400 Hz via an optical sensor positioned on the index finger. PPG technology detects blood volume changes through light absorption/reflection (green, red, or infrared wavelengths), providing a non-invasive method for deriving cardiac measures including heart rate (Alian & Shelley, 2014). While PPG sensors have shown strong correlation with ECG in stationary conditions (Volpes et al., 2024), their performance under movement and varying cognitive-emotional states requires validation (Poh et al., 2010). The inclusion of the Research Ring was methodologically motivated by its applied relevance as a lightweight wearable suitable for real-world occupational deployment. The device placement (index finger) was selected for optimal and standardized signal quality across participants (BIOPAC Systems, Inc., 2025). Hand assignment was randomized across participants to control for dominance effects.

2.2.3 FaceReader (Noldus Technology, Netherlands): Novel Non-Contact HR Measurement Device

FaceReader's remote photoplethysmography (rPPG) module (FR; Noldus Technology, Netherlands) extracted HR from facial video recordings by detecting blood volume changes in facial capillaries. This camera-based approach (Logitech C922, 1080p, positioned 50 cm from face) enables contactless cardiac monitoring through analysis of subtle color variations in facial skin (Wang et al., 2018; Yang et al., 2016). While rPPG technology has shown promise in controlled settings (Yu et al., 2011; Poh et al., 2010; Yang et al., 2016), its sensitivity to motion artifacts and lighting variations presents challenges for dynamic applications (Poh et al., 2010). The inclusion of a rPPG measurement device enabled evaluation of whether camera-based cardiac estimation can provide viable physiological monitoring in contexts where wearable devices are impractical, intrusive, or operationally constrained. The system recorded continuously across all seated task conditions; the walking task was excluded due to camera tracking constraints. Lighting conditions were standardized across participants to minimize environmental variability.

2.3 Procedure

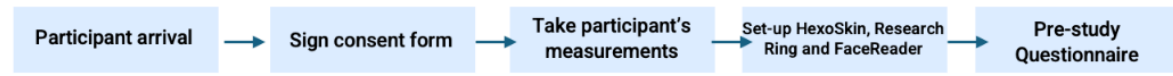
The experimental session followed a standardized protocol to simulate realistic work demands across both blue-collar and white-collar contexts. The deliberate integration of cognitive, emotional, and physical task demands was intended to elicit differentiated autonomic responses,

enabling evaluation of whether device convergence remained stable across different sources of physiological activation. Upon arrival, participants provided written informed consent and completed a demographic questionnaire. Anthropometric measurements (thorax, navel, and chest circumference) were recorded to ensure proper fit of the Hexoskin Smart Shirt (Carré Technologies Inc., Canada), the reference physiological measurement device. Participants were then equipped with the Research Ring (BIOPAC Systems Inc., USA), the novel physiological measurement device being tested.

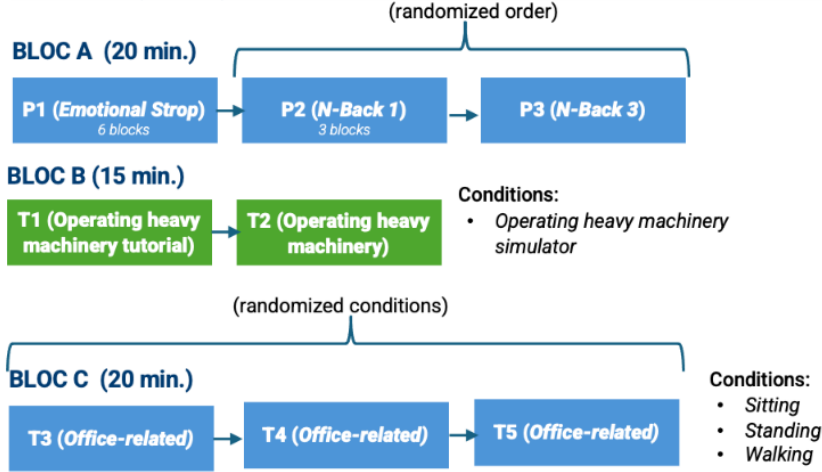
The experimental procedure consisted of three blocks (Fig. 1). In Block A, the Psychological Paradigm Block, participants completed the Emotional Stroop and N-Back paradigms while seated at a desk. Within the N-Back task, the order of the 1-back and 3-back load conditions was randomized across participants to control for sequence effects. Block B, Heavy Machinery Simulation Block, comprised two tasks (T1, T2) designed to simulate blue-collar heavy machinery operation, implemented through Farming Simulator V22 with a Logitech driving wheel and joystick setup. Block C, the Information Search Tasks Block, comprised three office-related information search tasks (T3, T4, T5) performed under different ergonomic conditions: sitting at a standard desk, standing at an elevated desk, and walking along a rectangular circuit in the laboratory. While the task sequence remained fixed (T3→T4→T5) to maintain comparable cognitive demands, the assignment of ergonomic conditions to each task was counterbalanced across participants through randomization (Appendix B), ensuring each condition-task combination was equally represented across the sample.

Figure 1. *Experimental procedure*

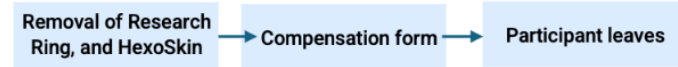
WELCOME AND SETUP (30 min.)



USER TEST (60 min.)



CONCLUSION (5 min.)

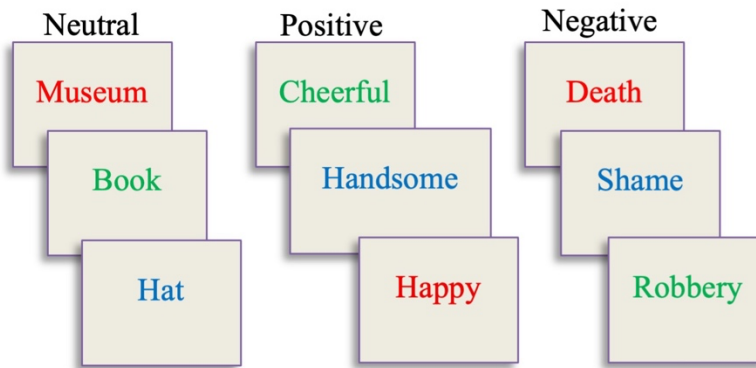


2.4. Stimulus material

2.4.1. Emotional interference task

The Emotional Stroop paradigm, a variation of the classic Stroop task, was used to induce emotional interference through processing of emotionally salient stimuli, allowing assessment of whether devices captured comparable HR responses to emotional processing (Williams et al., 2021; Ben-Haim et al., 2016; Song et al., 2017). This task engages autonomic responses distinct from those elicited by purely cognitive or physical demands, providing an additional dimension for device validation. Participants were instructed to identify the font color of words from three affective categories (neutral, positive, and negative). Words were presented in incongruent colors, requiring participants to inhibit automatic word reading tendencies (Fig. 2). Each trial consisted of a word presentation followed by an inter trial interval, and participants completed six blocks in total. This task was included to test whether the novel devices and reference system showed convergent validity in detecting HR changes induced by emotional interference.

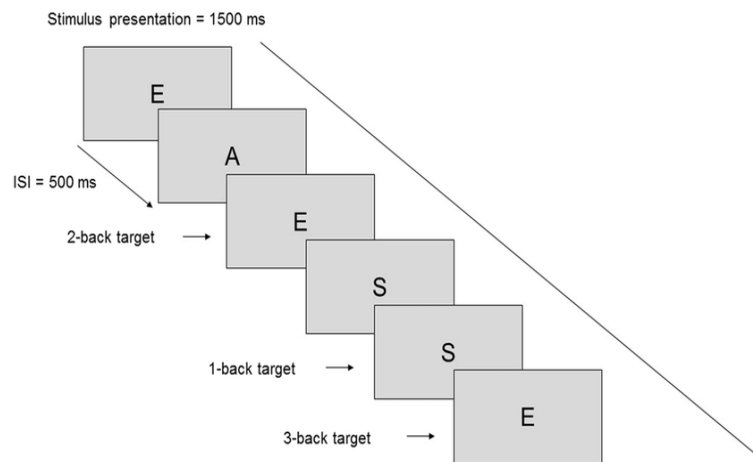
Figure 2. *Emotional stroop task*



2.4.2. Cognitive load task

The N-Back paradigm was used to induce varying levels of cognitive load through working memory demands, as it requires simultaneous monitoring, maintenance, and updating of information (Kirchner, 1958; Jaeggi et al., 2010). This task is one of the most popular working memory measures in psychophysiological research due to its ability to systematically manipulate cognitive load by varying N, with higher loads (e.g., 3-back) producing measurable performance decrements and increased autonomic activation compared to lower loads (e.g., 1-back) (Jaeggi et al., 2010; Sasuka et al., 2024). Participants viewed sequential letters and indicated via keyboard whether the current letter matched the one presented n positions earlier (Fig. 3). Two load conditions were administered: 1-back (low load) and 3-back (high load), selected to create a clear contrast in working memory demands while maintaining task feasibility (Jaeggi et al., 2010). Each trial included a stimulus presentation and an inter stimulus interval. Participants completed three blocks of 20 trials per load condition, presented in alternating order. This task was included to test whether the novel devices and reference system showed convergent validity in detecting HR changes induced by increasing cognitive load.

Figure 3. *N-Back task*



2.4.3. Heavy machinery simulator tasks

Two heavy machinery simulator scenarios simulated blue-collar operational demands typical of logistics and transportation environments. The simulator was operated using a gaming controller while participants remained seated.

In the tutorial scenario (T1), participants practiced basic operations including vehicle driving, braking, steering, and crane control to pick up and release a shipping container. This scenario familiarized participants with simulator controls. In the operational scenario (T2), participants performed a complete realistic operation requiring navigation to container location, crane operation to pick up container, driving to designated trailer location while maintaining container stability, and accurate placement of container onto trailer bed. Task completion time were recorded.

2.4.4. Information search tasks

Three information search tasks simulated white-collar office work in a transportation and logistics environment. Using an iPad with a Qualtrics-based interface, participants accessed a simulated operations database to identify the correct Purchase Order (PO) number based on pickup location, destination, and scheduled pickup time.

Standard retrieval tasks (T3 and T5): Baseline information search tasks requiring sustained attention and systematic information processing. The correct PO was present in the dropdown menu Stressor condition task (T4): An error-induction task in which the correct PO number was

deliberately absent from the dropdown menu, simulating a real-world data integrity issue. This manipulation was designed to induce cognitive stress.

Information search tasks were performed under three ergonomic conditions counterbalanced across participants: (1) sitting at a standard desk, (2) standing at an elevated desk surface, and (3) walking through a designated circuit within the laboratory. This ergonomic manipulation allowed systematic variation of physical activity while maintaining comparable cognitive demands, thereby enabling examination of motion-related influences on cross-device signal convergence while minimizing confounding from task structure. Task completion time and number of errors were recorded as behavioral performance metrics.

2.5. Data collection, data synchronization and statistical analysis

2.5.1. Cardiovascular data

Continuous cardiovascular data were recorded simultaneously throughout the experimental session using all three devices. The Hexoskin Smart Shirt captured electrocardiogram (ECG) signals via integrated textile electrodes embedded in the vest (Montes et al., 2018). The Research Ring (BIOPAC Systems Inc., USA) measured photoplethysmography (PPG) signals through an optical sensor positioned on the index finger, detecting peripheral blood volume changes with each cardiac cycle (Yang et al., 2016; Munten et al., 2023). The FaceReader (Noldus Technology, Netherlands) extracted remote photoplethysmography (rPPG) signals from facial video recordings, providing a contactless estimate of cardiac activity. The hand on which participants wore the Research Ring was randomized across participants.

All devices recorded continuously and concurrently across every experimental condition. HR was derived as the number of beats per minute (bpm) based on each device's native detection algorithm. For each participant, HR values were aggregated by task condition (e.g., 1-back, 3-back, T1–T5, and emotional Stroop blocks) to yield one mean HR value per device per condition. This structure enabled within-subject comparisons of mean heart rate (HR_m), defined as the average number of heart beats per minute derived from the continuous HR signal recorded by each device over the same temporal window.

Mean heart rate (HR_m) was used as the primary cardiovascular outcome, indexing overall cardiac activation and physiological arousal. Increases in HR_m were interpreted as reflecting higher sympathetic activation and reduced parasympathetic activity in response to physical, cognitive, or emotional demands. HR_m was therefore computed separately for each task condition and device source (Hexoskin, Research Ring, FaceReader) to allow convergent validity analyses across physiological measurement modalities.

2.5.2. Data synchronization

All cardiovascular data from the three devices, Hexoskin Smart Shirt (Carré Technologies Inc., Canada), Research Ring (BIOPAC Systems Inc., USA), and FaceReader remote photoplethysmography (Noldus Technology, Netherlands), were synchronized using Cobalt Photobooth (Tech3Lab, Canada), a proprietary software platform designed for multi-modal psychophysiological data integration (Léger et al., 2019). Raw data files were prepared according to device-specific formats. The synchronization protocol utilized NTP Bluetooth timestamps and event markers from Cobalt Capture V1 (Tech3Lab, Canada) to align all signals to a common timeline, compensating for device-specific latencies and sampling rate differences (Léger et al., 2019). This synchronization enabled direct statistical comparison of HR values across devices throughout all experimental conditions.

2.5.3. Statistical analysis

All statistical analyses were performed to assess convergent validity between the three HR measurement devices: Hexoskin Smart Shirt (Carré Technologies Inc., Canada), Research Ring (BIOPAC Systems Inc., USA), and FaceReader (Noldus Technology, Netherlands). Descriptive statistics (mean, standard deviation) were calculated for HR across all experimental conditions for each device to characterize cardiovascular responses to task demands. Prior to inferential analyses, HR time series were visually inspected at the participant level to verify temporal correspondence, detect potential outliers, and assess signal quality across devices.

Convergent validity was evaluated using three complementary approaches. First, linear mixed-effects regression models with random intercepts were conducted to compare least-squares means of HR amplitude between device pairs across all experimental conditions. Device was entered as the fixed effect, and p-values were adjusted using the Holm-Bonferroni method to control for

multiple comparisons. Second, Bland-Altman analysis quantified agreement between devices by calculating mean bias and 95% limits of agreement (mean difference $\pm 1.96 \times \text{SD}$), with acceptable tolerance defined as ± 10 bpm based on established psychophysiological research standards (van Lier et al., 2020). Third, cross-correlation functions were computed to assess temporal synchronization between HR time series, with Pearson correlation coefficients and lag values indicating the degree of temporal concordance.

To examine device sensitivity to task demands, separate linear mixed-effects regressions were conducted for each device. For physical conditions (sitting, standing, walking, heavy-machinery operation), pairwise comparisons were made between conditions. For cognitive (n-back 1, n-back 3) and emotional interference (Emotional Stroop blocks) tasks, mean HR was compared against baseline. Main effects of condition were tested, followed by pairwise contrasts with Holm-Bonferroni adjustments. All statistical models were implemented using SAS OnDemand for Academics and SAS Studio, a web-based development environment supporting cloud-based data processing and model execution.

3. Results

3.1 Descriptive statistics comparing the three physiological measurement devices across all task conditions

Table 2 presents the descriptive statistics for mean HR across all task conditions and the baseline for each device. Mean HR increased from sitting to walking and heavy-machinery operation, consistent with expected cardiovascular responses to physical effort. During the cognitive (n-back 1, n-back 3) and emotional-stimulus tasks (blocks), HR values remained lower than those observed during physical activity, with cognitive tasks showing modest elevations relative to baseline and emotional-stimulus blocks remaining comparatively stable. Together, these descriptive values provide the initial context for examining signal quality and device-level consistency in the subsequent analyses.

Table 2. *Descriptive statistics of mean heart rate across all conditions*

Condition	Source	n	Mean HR (bpm)	SD
Baseline	FR	32	59.13	7.01
	HS	33	74.06	10.58
	RR	33	74.51	8.95
Sitting	FR	34	60.01	9.72
	HS	35	70.58	10.78
	RR	30	72.30	8.94
Standing	FR	33	59.53	10.13
	HS	35	79.48	10.98
	RR	31	77.62	8.77
Walking	FR	.	.	.
	HS	35	84.24	11.03
	RR	32	83.92	9.38
Operating a heavy machinery simulator	FR	34	56.74	5.44
	HS	35	75.46	9.72
	RR	35	74.83	7.48
1-Back	FR	34	58.50	10.53
	HS	35	76.54	11.26
	RR	33	76.13	11.18
3-Back	FR	34	55.13	8.08
	HS	35	76.69	10.26
	RR	33	77.19	9.99
Emotional Stroop Blocks	FR	34	58.62	6.55
	HS	35	74.30	9.98
	RR	35	74.84	8.88

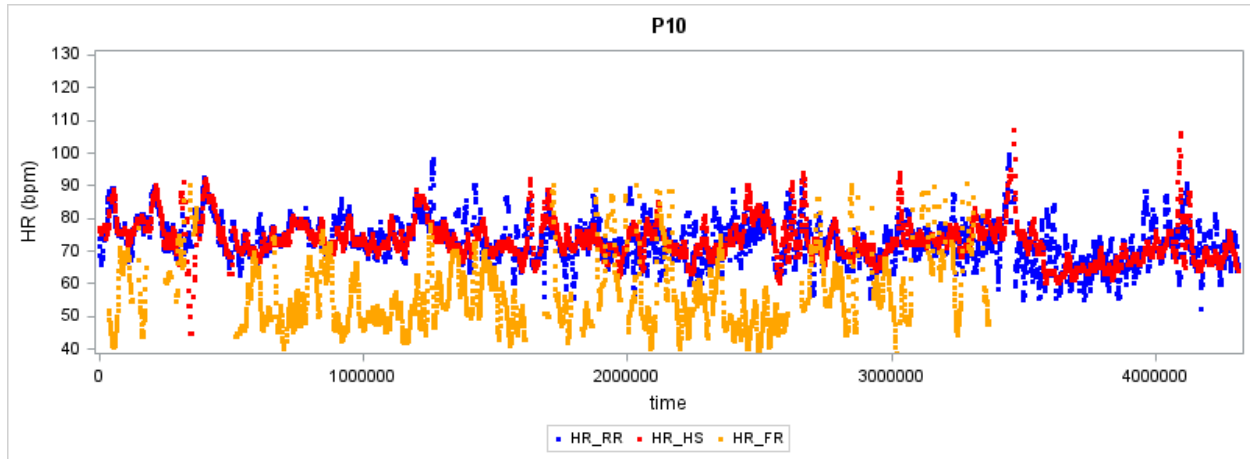
Note. FR = FaceReader (Noldus Technology, Netherlands); HS = Hexoskin Smart Shirt (Carré Technologies Inc., Canada); RR = Research Ring (Biopac Systems Inc., USA). FaceReader did not record signal during the walking condition due to technical constraints, as participants could not walk while maintaining webcam alignment required for remote photoplethysmography (rPPG).

3.2 Visual representation of signal quality across all task conditions

Prior to statistical analyses, heart rate (HR) time series were visually inspected at the participant level to assess synchronization and signal quality across devices. For each participant, time-aligned plots were generated to verify temporal correspondence, detect potential outliers, and observe amplitude consistency between the Hexoskin, Research Ring, and FaceReader.

Visual inspection of individual recordings (n=35) showed stable temporal overlap between Research Ring and Hexoskin across participants, whereas FaceReader exhibited frequent dropouts and unstable readings (Fig. 4 illustrates a representative case).

Figure 4. *HR signal quality and temporal alignment of Hexoskin, Research Ring, and FaceReader across all tasks for one participant (representative case)*



3.3 Convergent validity between Research Ring and Hexoskin

3.3.1 Regression analysis between Hexoskin and Research Ring using HR data from all physical, cognitive, emotional, and baseline conditions

To assess whether HR amplitude differed between Hexoskin (Carré Technologies Inc., Canada) and the Research Ring (Biopac Systems Inc., USA) across the full set of experimental tasks, a linear regression model with random intercepts was conducted to compare least-squares means of HR amplitude (HR_m) between devices. The analysis included all HR_m observations from every task condition, including baseline, sitting, standing, walking, heavy-machinery operation, 1-back blocks, 3-back blocks, and all Emotional Stroop blocks. Device was entered as the fixed effect, and p-values were adjusted using Holm-Bonferroni method. Across the full set of task conditions, difference in mean HR between Hexoskin (Carré Technologies Inc., Canada) and Research Ring (Biopac Systems Inc., USA) was not statistically significant (estimate = -0.086 bpm; $t(842) = -0.15$; adjusted $p = 0.8834$) indicating that the difference in mean HR (HR_m) amplitude between Hexoskin (Carré Technologies Inc., Canada) and Research Ring (Biopac Systems Inc., USA) was not statistically detectable under the tested conditions. This result reflects convergence in mean

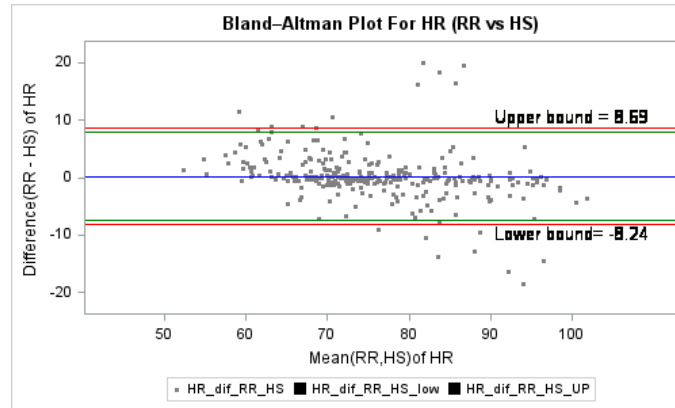
HR output across physical, cognitive, and emotional tasks and supports the Research Ring's ability to reproduce the mean HR (HR_m) amplitude measured by Hexoskin.

3.3.2 Bland-Altman analysis between Hexoskin and Research Ring using HR data from all task conditions

In line with these results, Bland-Altman analysis confirmed strong agreement between the two systems. The Bland-Altman (1986) is a method used to compare two measurements of the same variable through visual representation (Fig. 5) , quantifying agreement between two quantitative measurements by studying the mean difference and constructing limits of agreement (van Lier et al., 2020). Specifically, 95% confidence limits were computed as the mean difference $\pm 1.96 \times SD$.

The mean bias was negligible (≈ 0 bpm), with 95 % LOA ranging from -8.24 to 8.69 bpm, indicating that 95 % of paired measurements fell within this narrow range. These values lie well within the accepted tolerance for HR monitoring in psychophysiological research (van Lier et al., 2020). Residuals showed no systematic trend, suggesting that device agreement was stable across HR magnitudes and physical conditions (Fig. 5). Collectively, these results demonstrate a high degree of convergent validity between Hexoskin and Research Ring for HR amplitude in a controlled laboratory environment designed to reproduce key elements of a dynamic workplace.

Figure 5. Bland-Altman plot demonstrating HR agreement between Hexoskin and Research Ring

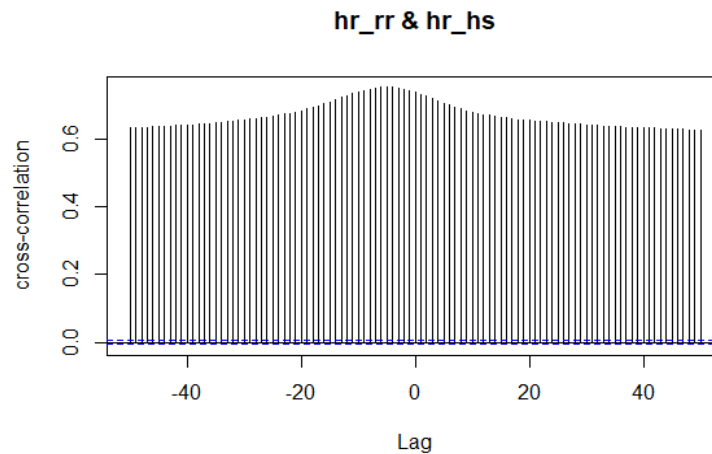


3.3.3 Temporal synchronization analysis between Hexoskin and Research Ring using HR data from all task conditions

To complement analyses, cross-correlation functions (CCF) were computed to quantify the temporal synchrony between HR time series recorded by Hexoskin and Research Ring. Cross-

correlation measures the similarity of two physiological signals as a function of time lag, that is, the degree to which fluctuations in one signal predict those in the other when shifted backward or forward in time. Each interval, which is called lag, is proportional to the frequency of the signal. High positive correlations indicate that the two signals rise and fall together, while a lag close to zero denotes near-perfect temporal alignment.

Figure 6. *Cross correlation between RR & HS*



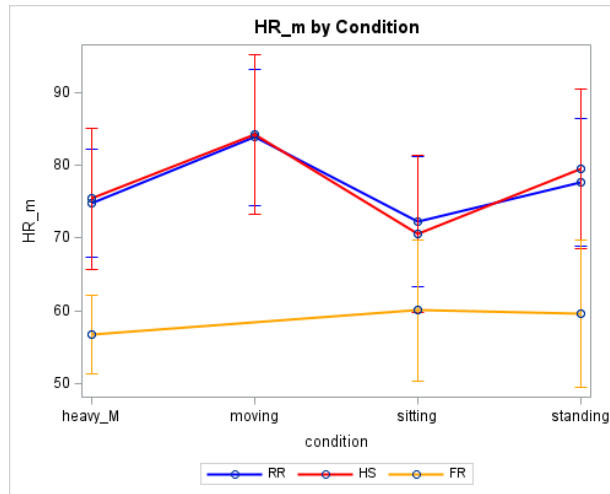
Results revealed a strong Pearson correlation coefficient between Research Ring and Hexoskin ($r = 0.75$) with a lag of -5 samples, demonstrating a high degree of temporal concordance between HR signals measured by both devices (Fig. 6).

3.3.4 Agreement in HR amplitude between Hexoskin and Research Ring during physical task conditions only (sitting, standing, walking, operating a heavy-machinery simulator)

Figure 7 displays a similar pattern visually, showing near-identical HR trajectories and overlapping confidence intervals for Hexoskin and the Research Ring. Both devices exhibited parallel modulation of HR across conditions, reflecting high amplitude consistency.

A linear mixed-effects regression confirmed a significant main effect of condition ($p < .05$) for both devices, indicating sensitivity to task-related changes. Pairwise contrasts followed the same direction and relative magnitude across sensors, supporting strong convergent validity between Hexoskin and the Research Ring for HR amplitude under dynamic conditions.

Figure 7. Mean HR across physical conditions for Hexoskin, Research Ring, and FaceReader



3.3.5 Agreement in HR amplitude between Hexoskin and Research Ring during cognitive and emotionally demanding task conditions (1-Back, 3-Back, Emotional Stroop Blocks)

After establishing that the Research Ring accurately reproduced Hexoskin's HR amplitude during physical activity, we conducted exploratory analyses on cognitively and emotionally demanding tasks to evaluate whether the device could also capture subtler HR changes associated with mental effort and affective engagement. Therefore, linear mixed-effects regressions with random intercepts were conducted separately for Hexoskin (HS) and Research Ring (RR) data to compare mean HR (HR_m) between baseline and task conditions.

No significant differences were observed between baseline and block segments for either device (all adj $p > .40$), suggesting comparable HR_m levels during these periods. In contrast, when comparing n-back tasks to baseline, significant increases in HR_m were detected for both the 1-back and 3-back conditions using HS data (adj $p = .0177$ and $.0167$, respectively), and for the 3-back condition using RR data (adj $p = .0201$). These results indicate that HR_m increased with task-related cognitive load, with both devices showing sensitivity to changes in mental workload intensity.

Figure 8 illustrates mean HR_m across the cognitive conditions (baseline, 1-back, 3-back) for Hexoskin and Research Ring. Both devices show a clear upward trend in mean HR_m from baseline to the n-back tasks, with the highest values observed during the 3-back condition. This pattern reflects an increase in cardiovascular activity as task difficulty and cognitive load rise. The

inferential results confirmed that these differences were statistically significant for both 1-back and 3-back using Hexoskin data, and for 3-back using Research Ring data, indicating that both systems captured the expected modulation of HR_m by cognitive workload.

Figure 8. *Mean HR across cognitive tasks for Hexoskin, Research Ring, and FaceReader*

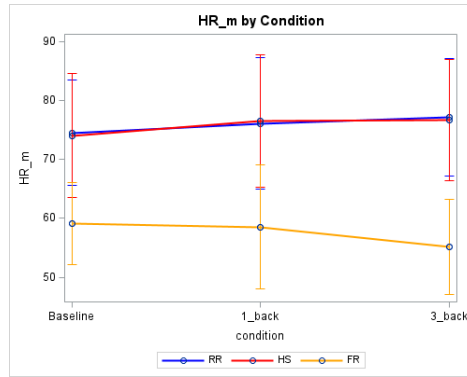
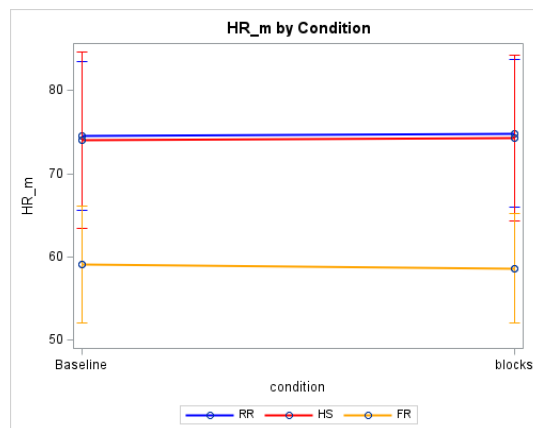


Figure 9 displays mean HR_m during the 6 blocks of Emotional Stroop Task compared to baseline. For both devices, HR_m values remained relatively stable across conditions, with no apparent upward or downward trend. This visual pattern is consistent with the regression results, which showed no significant differences between block and baseline segments.

Together, these figures illustrate that while both devices detected robust HR_m changes in response to cognitive task demands, HR_m remained steady during the emotional conditions, suggesting that physiological activation was more strongly driven by cognitive load than by emotional stimuli in this dataset.

Figure 9. *Mean HR across emotionally demanding tasks for Hexoskin, Research Ring, and FaceReader*

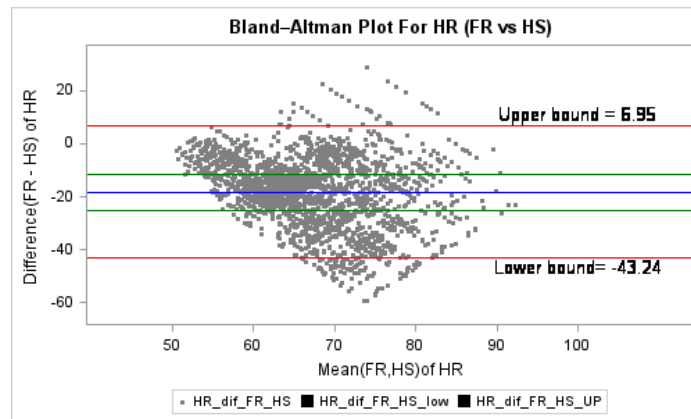


3.4 Convergent validity between Hexoskin and FaceReader across all task conditions

Linear mixed-effects regressions with random intercepts were conducted to compare mean heart rate (HR_m) between FaceReader and Hexoskin across identical task segments. Holm adjustments were applied to control for multiple comparisons. Agreement between systems was further examined through Bland–Altman analysis (amplitude-level agreement) and cross-correlation functions (temporal synchronization).

The regression analyses revealed no significant correspondence between FaceReader and Hexoskin HR_m values across conditions (all adj $p > .40$). Mean HR estimates derived from FaceReader systematically deviated from those recorded by Hexoskin, showing reduced amplitude and irregular fluctuations.

Figure 10. Bland-Altman plot showing HR disagreement between Hexoskin and FaceReader across all tasks

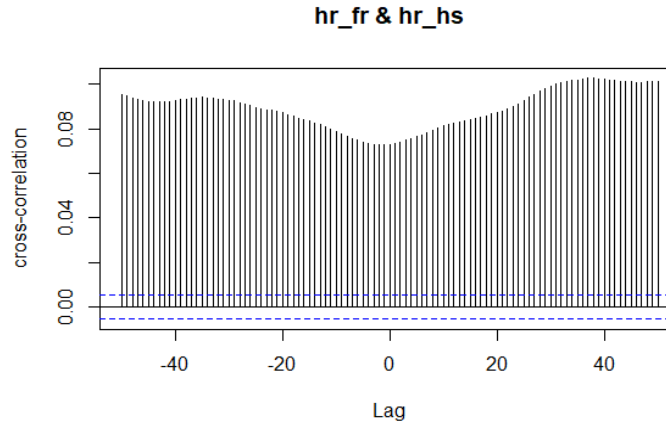


Bland–Altman analysis confirmed this lack of agreement, with wide limits of agreement ($> \pm 15$ bpm) and a consistent negative bias, indicating that FaceReader tended to underestimate HR relative to Hexoskin (Fig. 10). These results reflect substantial amplitude-level discrepancies between the two systems.

Cross-correlation analysis further revealed a weak temporal association between the two HR signals ($r = 0.10$, lag = +37 samples), indicating poor synchronization and delayed responsiveness of the FaceReader-derived HR signal relative to Hexoskin (Fig. 11)

Visual inspection of participant-level traces corroborated these findings, showing frequent signal dropouts and unstable trajectories during movement or subtle head motion.

Figure 11. *Cross-correlation between Hexoskin and FaceReader across all tasks*



Collectively, these results demonstrate poor convergent validity between FaceReader and Hexoskin. The remote photoplethysmography (rPPG) signal extracted from facial video by FaceReader did not reproduce the temporal dynamics or amplitude variability of HR signals recorded from contact-based wearable sensors. This discrepancy likely reflects the high sensitivity of rPPG to lighting conditions, motion artifacts, and facial occlusion. Consequently, FaceReader's rPPG-derived HR estimates cannot be considered valid for psychophysiological measurement in a dynamic environment.

3.5 Summary of device agreement and convergent validity results across all tasks

Across all analyses, results consistently demonstrated strong agreement between Hexoskin and the Research Ring, both in terms of HR amplitude and temporal synchronization, confirming that the two wearable systems provide comparable measures of heart rate across physical, cognitive, and emotional contexts. In contrast, the comparison between FaceReader and Hexoskin revealed substantial discrepancies in HR amplitude, greater variability, and poor temporal stability. Together, these findings indicate that while the Research Ring yield convergent and reliable HR signals under dynamic experimental conditions, FaceReader's rPPG does not reach the level of accuracy required for psychophysiological research applications.

4. Discussion

4.1 Convergent validity of wearable and camera-based HR sensors

The present study examined the convergent validity of two novel HR measurement devices, the Research Ring (BIOPAC Systems Inc., USA) and FaceReader's remote photoplethysmography (rPPG) (Noldus Technology, Netherlands), against a validated ECG reference system, the Hexoskin Smart Shirt (Carré Technologies Inc., Canada), under simulated workplace physical, cognitive, and emotional tasks. The Research Ring demonstrated strong convergent validity with the Hexoskin Smart Shirt across regression, Bland-Altman, and cross-correlation analyses, showing minimal bias, narrow limits of agreement, and strong temporal alignment. In contrast, FaceReader's rPPG showed poor convergence characterized by large amplitude discrepancies, weak correlations, and temporal delays, confirming known rPPG limitations related to motion sensitivity and environmental factors (Poh et al., 2010; Yu et al., 2011; Yang et al., 2016). Both the Hexoskin Smart Shirt and Research Ring successfully captured expected cardiovascular responses to task demands, with HR increasing during physical and cognitive tasks (particularly the n-back paradigm) relative to baseline, consistent with sympathetic activation under elevated workload (Krantz & Manuck, 1984; Obrist, 2012). Interestingly, while the Emotional Stroop Task failed to elicit significant autonomic changes, the devices detected this lack of change as well. These findings support the Research Ring (BIOPAC Systems Inc., USA) as a valid alternative to chest-based ECG for HR monitoring in dynamic research settings, whereas current camera-based rPPG technology remains unsuitable for ecologically valid applications involving movement.

4.2 Sensitivity of HR to physical, cognitive, and emotional demands

Across all devices, HR increased during physical and cognitively demanding conditions compared to baseline, consistent with the expected activation of the sympathetic nervous system under elevated workload (Krantz & Manuck, 1984; Obrist, 2012). The strongest and most consistent HR modulation was observed during the n-back task, where HR rose with increasing cognitive load, replicating classic findings in cardiovascular psychophysiology (Hao et al., 2022). In contrast, HR remained stable across Emotional Stroop blocks, suggesting that the affective component of this paradigm was insufficient to elicit strong autonomic changes beyond baseline cognitive control demands. These results confirm that both Hexoskin Smart Shirt (Carré Technologies Inc., Canada) and the Research Ring (BIOPAC Systems Inc., USA) are sensitive to task-induced cardiac modulation across varying cognitive and physical intensities.

4.3 Methodological contributions

From a methodological perspective, this research contributes to applied workplace psychophysiology by identifying which physiological measurement technologies maintain convergent validity under ecologically realistic occupational demands. By comparing wearable and contactless heart rate monitoring devices against an ECG-based reference under dynamic physical, cognitive, and emotional task conditions, this study addresses a critical methodological gap between controlled laboratory validation and the requirements of field-oriented employee experience research. Rather than treating validation as a purely technical exercise, this work positions convergent validity as a prerequisite for the credible and safe use of physiological measurement in occupational contexts, thereby establishing a methodological foundation for future ambulatory psychophysiology studies in employee experience research.

More specifically, this study advances device validation research in three substantial ways. First, it extends PPG–ECG validity beyond traditional laboratory paradigms by incorporating physical (sitting, standing, walking, simulated heavy machinery operation), cognitive (n-back with varying loads), and emotional interference (Emotional Stroop) tasks that approximate real occupational contexts. Existing validation studies comparing PPG-derived HR with ECG-derived HR have focused predominantly on stationary conditions and predictable activities such as resting states, where electrode sensor contact remains stable, and movement is minimized (Volpes et al., 2016; Poh et al., 2010). This methodological choice reflects PPG's well-documented susceptibility to motion artifacts, which can severely compromise data reliability (Poh et al., 2010; Pinheiro et al., 2016; Volpes et al., 2024; Yang et al., 2016; Yu et al., 2011). By contrast, this study demonstrates that PPG-derived HR converges with ECG-derived HR across dynamic, multi-modal, ecologically relevant tasks.

Second, this work introduces a reproducible validation blueprint integrating three complementary workplace demand dimensions: dynamic physical movement, cognitive load variation, and emotional interference. While existing guidelines emphasize gold standard comparison (van Lier et al., 2020; Schoenmakers et al., 2025), they rarely specify which experimental tasks capture real-world occupational psychophysiological variability (Krantz & Manuck, 1984; Balapour & Riedl, 2025). Validation studies employing tasks that resemble activities workers encounter in occupational settings remain scarce, limiting ecological validity improvements (Balapour & Riedl,

2025). Our protocol addresses this gap by incorporating workplace-relevant stressors that span the physical demands faced by blue collar workers (standing, walking, machinery operation), the cognitive challenges common to both occupational groups (sustained attention, working memory load), and emotional interference tasks that simulate workplace pressures. By evaluating whether different HR measurement devices capture comparable autonomic responses across these diverse occupational demands, this approach provides a template for validating physiological sensors while standardizing experimental protocol, synchronizing data between devices, and signal quality inspection, elements inconsistently reported in prior validation work (van Lier et al., 2020; Schoenmakers et al., 2025; Giurgiu et al., 2011).

Third, this protocol establishes methodological foundations for ambulatory psychophysiology by bridging controlled laboratory validation with ecologically valid task demands. Unlike traditional validation approaches that rely on isolated resting states or single-activity protocols (Montes et al., 2018; Cherif et al., 2018), our hybrid framework preserves experimental rigor while systematically introducing workplace-relevant complexity. By incorporating concurrent physical movement, graduated cognitive load, emotional interference, and naturalistic motion artifacts within a controlled environment, this approach more closely approximates the psychophysiological dynamics encountered in occupational settings than in conventional laboratory (Balapour & Rield, 2025; Krantz & Manuck, 1984). This methodological advancement is particularly critical for ambulatory psychophysiology to optimize employee experience, where continuous physiological monitoring in actual work environments informs workplace safety interventions, human performance, health and wellness (Tindale et al., 2022; Maltseva et al., 2022). In high-risk industries, for instance, smart wearable sensors can monitor workers' physiological indices over extended periods, revealing fatigue trajectories under cognitive demands and enabling evidence-based schedule optimization to enhance performance (Hao et al, 2022). The research protocol's modular structure enables expansion to field deployments across diverse operational contexts, from logistics centers and manufacturing floors to transportation systems and office environments, where researchers can evaluate sustained signal quality, long-term wearability, and context-specific implementation barriers that exceed laboratory assessment capabilities. By establishing this translational pathway from controlled laboratory validation to field application, this methodology provides the foundation necessary for researchers to select optimal

psychophysiological measurement tools for future occupational studies, ensuring data quality while minimizing financial and operational risks associated with poorly validated devices.

4.4 Industry implications

Beyond its methodological contributions, this work has direct implications for how organizations design work in ways that sustain both human and organizational health. By establishing which heart rate measurement technologies maintain convergent validity under dynamic, ecologically realistic work conditions, this study enables organizations to access objective indicators of occupational stress as it unfolds during task execution, rather than inferring strain retrospectively from burnout, turnover, or safety incidents. These physiological signals make it possible to identify when work demands exceed human adaptive capacity and to redesign tasks, schedules, or environments accordingly. At an organizational level, designing work environments that respect human physiological limits supports healthier systems overall, reducing hidden costs associated with fatigue, disengagement, and error accumulation, while aligning employee experience with the long-term interests of organizations and their stakeholders. Moreover, by clearly delineating which technologies are valid under realistic conditions and which are not, these findings help organizations avoid premature deployment of tools that may compromise data quality and lead to misinformed interventions. In this sense, validated physiological measurement functions not as a monitoring mechanism, but as a design input for creating work that is sustainable for individuals and resilient for organizations.

4.5 Limitations and future directions

Several limitations should be acknowledged, each presenting opportunities for future research. First, the sample size ($n=35$) limits generalizability to broader populations. Future studies should replicate this validation protocol with larger, more diverse samples, including varied age groups, fitness levels, and occupational backgrounds, to establish population-specific validity parameters and normative data for workplace applications.

Second, HR provides a continuous, general index of cardiac activity but cannot disentangle sympathetic and parasympathetic contributions, as both autonomic branches directly influence cardiac output (Krantz & Manuck, 1984; Ikehara & Crosby, 2005). This limitation is particularly relevant when assessing stress responses, where distinguishing between activation (sympathetic)

and recovery (parasympathetic) processes is crucial (Kumar et al., 2007; Choi et al., 2012; Kop et al., 2011). Future research should integrate complementary psychophysiological measures such as HR variability (HRV) for autonomic balance assessment, particularly high-frequency components indexing vagal activity (Choi et al., 2012), along with electrodermal activity (EDA) for sympathetic arousal and respiratory sinus arrhythmia, creating comprehensive multi-modal validation frameworks that capture the full spectrum of autonomic responses to occupational stressors (Alberdi et al., 2016; Wu et al., 2019).

Third, the rPPG data were recorded under fixed indoor lighting conditions, limiting conclusions about performance in variable environments. Subsequent studies should systematically evaluate advanced video-based algorithms under diverse lighting conditions (natural daylight, fluorescent, and LED), varying camera angles, and different skin tones, while exploring machine learning approaches that could compensate for luminance fluctuations and movement artifacts in real-time.

Finally, the study focused on short-term validity under controlled conditions. Longitudinal field studies should deploy these devices in actual occupational settings over extended periods (weeks to months) to assess signal degradation, sensor drift, user compliance, and long-term reliability, while developing quality assurance protocols for continuous workplace monitoring systems that maintain measurement validity across time and environmental variations.

5. Conclusion

This study provides systematic evidence that the Research Ring, a wearable PPG-based device, achieves convergent validity with a validated ECG-based reference (Hexoskin) for mean heart rate across ecologically relevant laboratory tasks that simulate workplace demands. Across physical, cognitive, and emotional task conditions, the Research Ring showed negligible bias, narrow limits of agreement, and strong temporal alignment with the reference device. In contrast, FaceReader's camera-based rPPG did not reach acceptable validity, exhibiting substantial amplitude underestimation, wide limits of agreement, and poor temporal synchronization. Together, these results indicate that wearable PPG can reliably reproduce ECG-derived HR under dynamic, work-like conditions, whereas current camera-based rPPG remains unsuitable for use in settings involving movement and naturalistic variability.

From a methodological standpoint, this work advances applied workplace psychophysiology by demonstrating that convergent validity can be assessed under controlled yet ecologically realistic task demands, rather than restricted to static resting-state paradigms. By integrating a task-diverse protocol (physical movement, cognitive load variation, and emotional interference) and combining complementary agreement approaches (mixed-effects modeling, Bland-Altman analysis, and cross-correlation), this study provides a reproducible blueprint for validating physiological sensors intended for ambulatory and applied research. This contribution helps close a critical gap between laboratory validation conventions and the measurement requirements of workplace research, where motion artifacts, frequent task switching, and the concurrent presence of physical, cognitive, and emotional demands are the norm.

From an industry perspective, these findings support a responsible transition from relying solely on retrospective and subjective indicators of stress (e.g., surveys, interviews, downstream outcomes) toward integrating validated physiological signals as a complementary design input for employee experience and workplace safety. When physiological measurement is valid, it enables organizations to detect when work demands exceed human adaptive capacity as work unfolds, informing interventions that aim to create work environments that are sustainable for individuals and resilient for organizations. Notably, by delineating technological limits, particularly for camera-based rPPG, this study helps prevent premature adoption of tools that could generate misleading data and misdirect workplace decisions. Future research should extend these validations to longitudinal field deployments, examine additional autonomic indices (e.g., HRV, EDA), and develop multimodal measurement frameworks that integrate physiological data with behavioural and qualitative evidence to capture workplace experience with both rigor and context.

References

- Akbar, F., Mark, G., Pavlidis, I., & Gutierrez-Osuna, R. (2019). An Empirical Study Comparing Unobtrusive Physiological Sensors for Stress Detection in Computer Work. *Sensors (Basel, Switzerland)*, 19(17), 3766. <https://doi.org/10.3390/s19173766>.
- Alberdi, A., Aztiria, A., & Basarab, A. (2016). Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of Biomedical Informatics*, 59, 49–75. <https://doi.org/10.1016/j.jbi.2015.11.007>.
- Alian, A. A., & Shelley, K. H. (2014). Photoplethysmography. *Best Practice & Research. Clinical Anaesthesiology*, 28(4), 395–406. <https://doi.org/10.1016/j.bpa.2014.08.006>.
- Bañuelos-Lozoya, E., González-Serna, G., González-Franco, N., Fragoso-Díaz, O., & Castro-Sánchez, N. (2021). A Systematic Review for Cognitive State-Based QoE/UX Evaluation. *Sensors (Basel, Switzerland)*, 21(10), 3439. <https://doi.org/10.3390/s21103439>.
- Balapour, A., & Riedl, R. (2025). Ecological Validity in NeuroIS Research: Theory, Evidence, and a Roadmap for Future Studies. *Journal of the Association for Information Systems*, 26(1), 9–65. <https://doi.org/10.17705/1jais.00904>.
- Ben-Haim MS, Williams P, Howard Z, Mama Y, Eidels A, Algom D. The Emotional Stroop Task: Assessing Cognitive Performance under Exposure to Emotional Content. *J Vis Exp*. 2016 Jun 29;(112):53720. doi: 10.3791/53720.
- BIOPAC Systems, Inc. (2025). Research Ring product sheet [PDF]. <https://www.biopac.com/wp-content/uploads/RESRING-PACK.pdf>.

- Bland, J., & Altman, Douglas G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307–310.
[https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Cherif, N. H., Mezghani, N., Gaudreault, N., Ouakrim, Y., Mouzoune, I., & Boulay, P. (2018). Physiological Data Validation of the Hexoskin Smart Textile: Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies, 150–156. <https://doi.org/10.5220/0006588001500156>
- Definition of heart rate. (2012). NCI Dictionary of Cancer Terms. National Cancer Institute.
<https://www.cancer.gov/publications/dictionaries/cancer-terms/def/heart-rate>.
- Garg, A., & Kapellusch, J. M. (2009). Applications of biomechanics for prevention of work-related musculoskeletal disorders. *Ergonomics*, 52(1), 36–59.
<https://doi.org/10.1080/00140130802480794>.
- Giurgiu, M., Timm, I., Becker, M., Schmidt, S., Wunsch, K., Nissen, R., Davidovski, D., Bussmann, J. B. J., Nigg, C. R., Reichert, M., Ebner-Priemer, U. W., Woll, A., & Haaren-Mack, B. von. (2022). Quality Evaluation of Free-living Validation Studies for the Assessment of 24-Hour Physical Behavior in Adults via Wearables: Systematic Review. *JMIR mHealth and uHealth*, 10(6), e36377. <https://doi.org/10.2196/36377>.
- Hao, T., Zheng, X., Wang, H., Xu, K., & Chen, S. (2022). Linear and nonlinear analyses of heart rate variability signals under mental load. *Biomedical Signal Processing and Control*, 77, 103758. <https://doi.org/10.1016/j.bspc.2022.103758>.

- Ikehara, C. S., & Crosby, M. E. (2005). Assessing Cognitive Load with Physiological Sensors. Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 295a–295a. 38th Annual Hawaii International Conference on System Sciences. <https://doi.org/10.1109/HICSS.2005.103>.
- Jaeggi, S. M., Buschkuhl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, 18(4), 394–412. <https://doi.org/10.1080/09658211003702171>.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4), 352–358. <https://doi.org/10.1037/h0043688>.
- Kop, W. J., Synowski, S. J., Newell, M. E., Schmidt, L. A., Waldstein, S. R., & Fox, N. A. (2011). Autonomic nervous system reactivity to positive and negative mood induction: The role of acute psychological responses and frontal electrocortical activity. *Biological Psychology*, 86(3), 230–238. <https://doi.org/10.1016/j.biopsycho.2010.12.003>.
- Krantz, D. S., & Manuck, S. B. (1984). Acute psychophysiologic reactivity and risk of cardiovascular disease: A review and methodologic critique. *Psychological Bulletin*, 96(3), 435–464. <https://doi.org/10.1037/0033-2909.96.3.435>.
- Kumar, M., Weippert, M., Vilbrandt, R., Kreuzfeld, S., & Stoll, R. (2007). Fuzzy Evaluation of Heart Rate Signals for Mental Stress Assessment. *IEEE Transactions on Fuzzy Systems*, 15(5), 791–808. *IEEE Transactions on Fuzzy Systems*. <https://doi.org/10.1109/TFUZZ.2006.889825>.

- Léger, P.-M., Courtemanche, F., Fredette, M., & Sénécal, S. (2019). A Cloud-Based Lab Management and Analytics Software for Triangulated Human-Centered Research. Lecture Notes in Information Systems and Organization, 93–99.
- Maltseva, K. (2020). Wearables in the workplace: The brave new world of employee engagement. *Business Horizons*, 63(4), 493–505.
<https://doi.org/10.1016/j.bushor.2020.03.007>.
- Marras, W. S., Cutlip, R. G., Burt, S. E., & Waters, T. R. (2009). National occupational research agenda (NORA) future directions in occupational musculoskeletal disorder health research. *Applied Ergonomics*, 40(1), 15–22. <https://doi.org/10.1016/j.apergo.2008.01.018>.
- Montes, J., Young, J. C., Tandy, R., & Navalta, J. W. (2018). Reliability and Validation of the Hexoskin Wearable Bio-Collection Device During Walking Conditions. *International Journal of Exercise Science*, 11(7), 806–816. <https://doi.org/10.70252/YPHF4748>.
- Obrist, P.A. *Cardiovascular Psychophysiology: A Perspective*; Springer Science & Business Media: New York, NY, USA, 2012. <https://doi.org/10.1017/S0033291700043671>.
- Pinheiro, N., Couceiro, R., Henriques, J., Muehlsteff, J., Quintal, I., Goncalves, L., & Carvalho, P. (2016). Can PPG be used for HRV analysis? Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2016, 2945–2949.
<https://doi.org/10.1109/EMBC.2016.7591347>.

- Poh, M.-Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10), 10762–10774. <https://doi.org/10.1364/OE.18.010762>.
- Poh, M.-Z., Swenson, N. C., & Picard, R. W. (2010). Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography. *IEEE Transactions on Information Technology in Biomedicine*, 14(3), 786–794. *IEEE Transactions on Information Technology in Biomedicine*. <https://doi.org/10.1109/TITB.2010.2042607>
- Riedl, R., Davis, F., & Hevner, A. (2014). Towards a NeuroIS Research Methodology: Intensifying the Discussion on Methods, Tools, and Measurement. *Journal of the Association for Information Systems*, 15(10). <https://doi.org/10.17705/1jais.00377>.
- Rowe, D. W., Sibert, J., & Irwin, D. (1998). Heart rate variability: Indicator of user state as an aid to human-computer interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 480–487. <https://doi.org/10.1145/274644.274709>.
- Sazuka, N., Katsumata, K., Komoriya, Y., Oba, T., & Ohira, H. (2024). Association of brain–autonomic activities and task accuracy under cognitive load: A pilot study using electroencephalogram, autonomic activity measurements, and arousal level estimated by machine learning. *Frontiers in Human Neuroscience*, 18. <https://doi.org/10.3389/fnhum.2024.1272121>.
- Schoenmakers, M., Saygin, M., Sikora, M., Vaessen, T., Noordzij, M., & de Geus, E. (2025). Stress in action wearables database: A database of noninvasive wearable monitors with

- systematic technical, reliability, validity, and usability information. *Behavior Research Methods*, 57(6), 171. <https://doi.org/10.3758/s13428-025-02685-4>
- Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the Association for Information Systems*, 13, 24. ProQuest Central. <https://doi.org/10.17705/1CAIS.01324>.
- Tindale, L. C., Chiu, D., Minielly, N., Hrincu, V., Talhouk, A., & Illes, J. (2022). Wearable Biosensors in the Workplace: Perceptions and Perspectives. *Frontiers in Digital Health*, 4. <https://doi.org/10.3389/fdgth.2022.800367>.
- van Lier, H. G., Pieterse, M. E., Garde, A., Postel, M. G., de Haan, H. A., Vollenbroek-Hutten, M. M. R., Schraagen, J. M., & Noordzij, M. L. (2020). A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. *Behavior Research Methods*, 52(2), 607–629. <https://doi.org/10.3758/s13428-019-01263-9>
- Villar, R., Beltrame, T., & Hughson, R. L. (2015). Validation of the Hexoskin wearable vest during lying, sitting, standing, and walking activities. *Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition Et Metabolisme*, 40(10), 1019–1024. <https://doi.org/10.1139/apnm-2015-0140>
- Volpes, G., Valenti, S., Genova, G., Barà, C., Parisi, A., Faes, L., Busacca, A., & Pernice, R. (2024). Wearable Ring-Shaped Biomedical Device for Physiological Monitoring through Finger-Based Acquisition of Electrocardiographic, Photoplethysmographic, and Galvanic

- Skin Response Signals: Design and Preliminary Measurements. *Biosensors*, 14(4), 205.
<https://doi.org/10.3390/bios14040205>.
- Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, 120(1), 3–24. <https://doi.org/10.1037/0033-2909.120.1.3>.
- Wright, R. A., & Gendolla, G. H. E. (Eds.). (2012). *How Motivation Affects Cardiovascular Response: Mechanisms and Applications*. American Psychological Association.
<https://www.jstor.org/stable/j.ctv1chs2jh>.
- Wu, W., Pirbhulal, S., Zhang, H., & Mukhopadhyay, S. C. (2019). Quantitative Assessment for Self-Tracking of Acute Stress Based on Triangulation Principle in a Wearable Sensor System. *IEEE Journal of Biomedical and Health Informatics*, 23(2), 703–713.
<https://doi.org/10.1109/JBHI.2018.2832069>.
- Yang, Y., Liu, C., Yu, H., Shao, D., Tsow, F., & Tao, N. (2016). Motion robust remote photoplethysmography in CIE Lab color space. *Journal of Biomedical Optics*, 21(11), 117001. <https://doi.org/10.1117/1.JBO.21.11.117001>.
- Yu, S., Hu, S., Azorin-Peris, V., Chambers, J. A., Zhu, Y., & Greenwald, S. E. (2011). Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise. *Journal of Biomedical Optics*, 16(7), 077010.
<https://doi.org/10.1117/1.3602852>.

Chapter 3: Conclusion

1. The Convergent Validity Requirement for Workplace Psychophysiology

This thesis identifies which physiological measurement devices maintain convergent validity under realistic workplace demands. The findings demonstrate that one novel device, the Research Ring (BIOPAC Systems Inc., USA), a wearable PPG-based sensor, achieved strong convergent validity with the Hexoskin Smart Shirt (Carré Technologies Inc., Canada), a validated ECG reference, across physical, cognitive, and emotional task conditions. In contrast, the other novel device, FaceReader's remote photoplethysmography (Noldus Technology, Netherlands), failed to achieve acceptable agreement, showing poor temporal synchronization and substantial amplitude discrepancies. These results directly answer the research question: under conditions simulating a dynamic work environment, wearable PPG technology can reliably reproduce ECG-derived heart rate, whereas current camera-based rPPG technology cannot. This validation is essential not only for technical accuracy but for establishing physiological sensing as a credible method in EX research.

Current approaches, such as surveys, interviews, and organizational metrics, rely on subjective self-report, making them vulnerable to recall bias, mood effects, and social-desirability distortion (Akbar et al., 2019). They capture only conscious, self-perceived aspects of stress, even though many stress responses occur outside conscious awareness.

Validated physiological monitoring offers a complementary solution by providing continuous, objective measurement that does not depend on perception, memory, or willingness to disclose. It therefore fills a critical methodological gap in understanding real stress as it unfolds in the workplace.

2. Objective Measurement in Previously Inaccessible Populations

For operational and frontline workers, this validation opens a critical possibility: continuous measurement of lived experience, captured in real time, independent of their ability or willingness to complete surveys. Currently, operational managers possess only backward-looking metrics. Validated wearable measurement changes this. By continuously capturing heart rate data during

actual work, managers gain access to granular, real-time evidence. For instance, a worker's resting heart rate is elevated across Tuesday-Wednesday shifts, indicating inadequate recovery, and therefore, revealing late-shift fatigue accumulation. This team's stress markers drop measurably post-intervention. These observations enable evidence-based, adaptive EX design targeted at populations whose experiences have not been extensively measured.

3. Practical, Methodological, and Ethical Constraints

However, workplace deployment of physiological measurement requires careful governance. Wearable sensors raise legitimate concerns about privacy, data misuse, and surveillance (Tindale et al., 2022). Devices must be thoughtfully designed to not impede work or create unacceptable intrusiveness (Akbar et al., 2019). Therefore, the generalizability of findings may be limited by regional deployment and workforce characteristics.

All in all, physiological measurement should be understood as complementary to, not a replacement for, traditional EX measurement methods. It addresses specific blind spots in current practice, real-time objective measurement, access to operational workers, and detection of implicit processes, while introducing new measurement challenges. The optimal approach integrates multiple modalities: objective physiological data combined with qualitative data, performance metrics, and employee voice (Alberdi et al., 2016). This multimodal framework acknowledges both what physiological measurement reveals and what it cannot.

4. A Framework for Responsible Implementation

This recognition of practical limitations strengthens the case for rigorous validation. If organizations are to use physiological data to inform decisions affecting workers' lives, such as task allocation, schedule changes, and intervention timing, the data must be trustworthy. Poor validation would mean deploying measurement tools with unknown reliability, potentially leading to misguided interventions based on faulty data. Conversely, validated physiological measurements used thoughtfully and transparently can extend EX research into dimensions currently unmeasured. The empirical foundation provided by this research is necessary for organizations to deploy physiological monitoring with confidence, understanding both its capabilities and its boundaries. It establishes which devices maintain validity under realistic

conditions, enabling informed decisions about when physiological measurement is appropriate, where it adds genuine insight, and where it must be complemented by other research methods.

5. Pathway Forward: From Laboratory Validation to Field Research

This thesis establishes which physiological sensors maintain validity under realistic workplace demands, creating the empirical foundation for a broader research agenda in ambulatory psychophysiology and employee experience. By establishing this foundation, this thesis enables a fundamental methodological transition: from measuring only what employees report about their experience to measuring what employees' physiology reveals about their experience.

Several research priorities emerge from this work. First, field validation studies should examine whether the convergent validity demonstrated here holds across extended deployment periods in actual occupational settings such as logistics centers, manufacturing floors, transportation systems, where environmental variability, sensor drift, and user compliance introduce challenges that laboratory protocols cannot fully replicate. Second, future research should expand the physiological indices under investigation, integrating heart rate variability (HRV), electrodermal activity (EDA), and respiratory metrics to capture the full spectrum of autonomic responses to occupational stressors. Third, comparative studies across diverse workforce populations, varying in age, ethnicity, fitness level, occupational role, and regional context, are needed to establish generalizability and identify population-specific validity parameters.

Beyond validation, a parallel research stream should examine how physiological data can be meaningfully integrated with qualitative methods, behavioral observation, and organizational metrics to build multimodal EX measurement frameworks. This integration is particularly consequential for operational workers whose burnout rates are highest, yet whose lived experience remains most invisible to current measurement systems.

This research progression, from laboratory validation to field validation to multimodal integration, represents the pathway through which physiological measurement can responsibly transition from methodological promise to evidence-based practice. This thesis provides the groundwork for that progression; in service of workers whose experience has remained unmeasured and whose needs demand to be seen.

Bibliography

- Abramson, A. (2022, January 1). *Burnout and stress are everywhere*. <https://www.apa.org/monitor/2022/01/special-burnout-stress>.
- Alberdi, A., Aztiria, A., & Basarab, A. (2016). Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of Biomedical Informatics*, 59, 49–75. <https://doi.org/10.1016/j.jbi.2015.11.007>.
- Alian, A. A., & Shelley, K. H. (2014). Photoplethysmography. *Best Practice & Research. Clinical Anaesthesiology*, 28(4), 395–406. <https://doi.org/10.1016/j.bpa.2014.08.006>.
- Akbar, F., Mark, G., Pavlidis, I., & Gutierrez-Osuna, R. (2019). An Empirical Study Comparing Unobtrusive Physiological Sensors for Stress Detection in Computer Work. *Sensors (Basel, Switzerland)*, 19(17), 3766. <https://doi.org/10.3390/s19173766>.
- Bañuelos-Lozoya, E., González-Serna, G., González-Franco, N., Fragoso-Díaz, O., & Castro-Sánchez, N. (2021). A Systematic Review for Cognitive State-Based QoE/UX Evaluation. *Sensors (Basel, Switzerland)*, 21(10), 3439. <https://doi.org/10.3390/s21103439>.
- Balapour, A., & Riedl, R. (2025). Ecological Validity in NeuroIS Research: Theory, Evidence, and a Roadmap for Future Studies. *Journal of the Association for Information Systems*, 26(1), 9–65. <https://doi.org/10.17705/1jais.00904>.
- Ben-Haim MS, Williams P, Howard Z, Mama Y, Eidels A, Algom D. The Emotional Stroop Task: Assessing Cognitive Performance under Exposure to Emotional Content. *J Vis Exp*. 2016 Jun 29;(112):53720. doi: 10.3791/53720.

BIOPAC Systems, Inc. (2025). Research Ring product sheet [PDF].

<https://www.biopac.com/wp-content/uploads/RESRING-PACK.pdf>.

Bland, J., & Altman, Douglas G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307–310.

[https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8).

Brassey, J., Coe, E., Dewhurst, M., & Enomoto, K. (2022, May 27). Addressing employee burnout-Are you solving the right problem. McKinsey.

https://www.mckinsey.com/mhi/our-insights/addressing-employee-burnout-are-you-solving-the-right-problem#/.

Cambridge Dictionary. (2025). Blue-collar.

<https://dictionary.cambridge.org/dictionary/english/blue-collar>.

Cambridge Dictionary. (2025, October 3). White-collar.

<https://dictionary.cambridge.org/dictionary/english/white-collar>.

Cherif, N. H., Mezghani, N., Gaudreault, N., Ouakrim, Y., Mouzoune, I., & Boulay, P. (2018).

Physiological Data Validation of the Hexoskin Smart Textile: Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies, 150–156. <https://doi.org/10.5220/0006588001500156>.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress.

Journal of Health and Social Behavior, 24(4), 385–396. <https://doi.org/10.2307/2136404>

Colligan, T. W., & Higgins, E. M. (2006). Workplace Stress: Etiology and Consequences.

Journal of Workplace Behavioral Health, 21(2), 89–97.

https://doi.org/10.1300/J490v21n02_07.

Contreras-Cruz, A., Kirbac, A., Dennett, C., & Daim, T. U. (2023). Human-centered design as a tool to improve employee experience: The case of a US plant-based food manufacturer.

Technology in Society, 73, 102248. <https://doi.org/10.1016/j.techsoc.2023.102248>.

Definition of heart rate. (2012). NCI Dictionary of Cancer Terms. National Cancer Institute.

<https://www.cancer.gov/publications/dictionaries/cancer-terms/def/heart-rate>.

de Oliveira, C., Saka, M., Bone, L., & Jacobs, R. (2023). The Role of Mental Health on Workplace Productivity: A Critical Review of the Literature. *Applied Health Economics and Health Policy*, 21(2), 167–193. <https://doi.org/10.1007/s40258-022-00761-w>.

Dėdelė, A., Miškinytė, A., Andrušaitytė, S., & Bartkutė, Ž. (2019). Perceived Stress Among Different Occupational Groups and the Interaction with Sedentary Behaviour. *International Journal of Environmental Research and Public Health*, 16(23), 4595.

<https://doi.org/10.3390/ijerph16234595>.

Dyerly, R. (2025, January 21). The Myth of Replaceability: Preparing for the Loss of Key

Employees. SHRM. <https://www.shrm.org/executive-network/insights/myth-replaceability-preparing-loss-key-employees>.

Garg, A., & Kapellusch, J. M. (2009). Applications of biomechanics for prevention of work-related musculoskeletal disorders. *Ergonomics*, 52(1), 36–59.

<https://doi.org/10.1080/00140130802480794>.

- Giurgiu, M., Timm, I., Becker, M., Schmidt, S., Wunsch, K., Nissen, R., Davidovski, D., Bussmann, J. B. J., Nigg, C. R., Reichert, M., Ebner-Priemer, U. W., Woll, A., & Haaren-Mack, B. von. (2022). Quality Evaluation of Free-living Validation Studies for the Assessment of 24-Hour Physical Behavior in Adults via Wearables: Systematic Review. *JMIR mHealth and uHealth*, 10(6), e36377. <https://doi.org/10.2196/36377>.
- Hao, T., Zheng, X., Wang, H., Xu, K., & Chen, S. (2022). Linear and nonlinear analyses of heart rate variability signals under mental load. *Biomedical Signal Processing and Control*, 77, 103758. <https://doi.org/10.1016/j.bspc.2022.103758>.
- Ikehara, C. S., & Crosby, M. E. (2005). Assessing Cognitive Load with Physiological Sensors. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 295a–295a. 38th Annual Hawaii International Conference on System Sciences. <https://doi.org/10.1109/HICSS.2005.103>
- Jaeggi, S. M., Buschkuhl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, 18(4), 394–412. <https://doi.org/10.1080/09658211003702171>.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4), 352–358. <https://doi.org/10.1037/h0043688>.
- Kop, W. J., Synowski, S. J., Newell, M. E., Schmidt, L. A., Waldstein, S. R., & Fox, N. A. (2011). Autonomic nervous system reactivity to positive and negative mood induction: The role of acute psychological responses and frontal electrocortical activity. *Biological Psychology*, 86(3), 230–238. <https://doi.org/10.1016/j.biopsycho.2010.12.003>.

- Krantz, D. S., & Manuck, S. B. (1984). Acute psychophysiologic reactivity and risk of cardiovascular disease: A review and methodologic critique. *Psychological Bulletin*, 96(3), 435–464. <https://doi.org/10.1037/0033-2909.96.3.435>.
- Kumar, M., Weippert, M., Vilbrandt, R., Kreuzfeld, S., & Stoll, R. (2007). Fuzzy Evaluation of Heart Rate Signals for Mental Stress Assessment. *IEEE Transactions on Fuzzy Systems*, 15(5), 791–808. *IEEE Transactions on Fuzzy Systems*. <https://doi.org/10.1109/TFUZZ.2006.889825>.
- Léger, P.-M., Courtemanche, F., Fredette, M., & Sénécal, S. (2019). A Cloud-Based Lab Management and Analytics Software for Triangulated Human-Centered Research. *Lecture Notes in Information Systems and Organization*, 93–99.
- Levine, J. A., Pavlidis, I. T., MacBride, L., Zhu, Z., & Tsiamyrtzis, P. (2009). Description and clinical studies of a device for the instantaneous detection of office-place stress. *WORK*, 34(3), 359–364. <https://doi.org/10.3233/WOR-2009-0934>
- Maltseva, K. (2020). Wearables in the workplace: The brave new world of employee engagement. *Business Horizons*, 63(4), 493–505. <https://doi.org/10.1016/j.bushor.2020.03.007>.
- Marras, W. S., Cutlip, R. G., Burt, S. E., & Waters, T. R. (2009). National occupational research agenda (NORA) future directions in occupational musculoskeletal disorder health research. *Applied Ergonomics*, 40(1), 15–22. <https://doi.org/10.1016/j.apergo.2008.01.018>.
- Maslach, C. (1993). Burnout: A multidimensional perspective. In *Professional burnout: Recent developments in theory and research* (pp. 19–32). Taylor & Francis.

- Maslach, C., Jackson, S. E., and Leiter, M. P. (1996). MBI: Maslach Burnout Inventory. Sunnyvale, CA: CPP, Incorporated.
- Montes, J., Young, J. C., Tandy, R., & Navalta, J. W. (2018). Reliability and Validation of the Hexoskin Wearable Bio-Collection Device During Walking Conditions. *International Journal of Exercise Science*, 11(7), 806–816. <https://doi.org/10.70252/YPHF4748>.
- Novacek, G., Lovich, D., & May, G. (2024, June 6). Four Keys to Boosting Inclusion and Beating Burnout. BCG Global. <https://www.bcg.com/publications/2024/four-keys-to-boosting-inclusion-and-beating-burnout>
- Obrist, P.A. Cardiovascular Psychophysiology: A Perspective; Springer Science & Business Media: New York, NY, USA, 2012. <https://doi.org/10.1017/S0033291700043671>.
- Öztürk Eyimaya, A., & Tezel, A. (2021). Evaluating Occupational Stress Levels of the Railway Workers. *Florence Nightingale Journal of Nursing*, 29(1), 40–55. <https://doi.org/10.5152/FNJJN.2021.19082>
- Pinheiro, N., Couceiro, R., Henriques, J., Muehlsteff, J., Quintal, I., Goncalves, L., & Carvalho, P. (2016). Can PPG be used for HRV analysis? Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2016, 2945–2949. <https://doi.org/10.1109/EMBC.2016.7591347>.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies.

The Journal of Applied Psychology, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>

Poh, M.-Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10), 10762–10774. <https://doi.org/10.1364/OE.18.010762>.

Poh, M.-Z., Swenson, N. C., & Picard, R. W. (2010). Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography. *IEEE Transactions on Information Technology in Biomedicine*, 14(3), 786–794. *IEEE Transactions on Information Technology in Biomedicine*. <https://doi.org/10.1109/TITB.2010.2042607>

Riedl, R., Davis, F., & Hevner, A. (2014). Towards a NeuroIS Research Methodology: Intensifying the Discussion on Methods, Tools, and Measurement. *Journal of the Association for Information Systems*, 15(10). <https://doi.org/10.17705/1jais.00377>.

Rowe, D. W., Sibert, J., & Irwin, D. (1998). Heart rate variability: Indicator of user state as an aid to human-computer interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 480–487. <https://doi.org/10.1145/274644.274709>.

Sazuka, N., Katsumata, K., Komoriya, Y., Oba, T., & Ohira, H. (2024). Association of brain–autonomic activities and task accuracy under cognitive load: A pilot study using electroencephalogram, autonomic activity measurements, and arousal level estimated by machine learning. *Frontiers in Human Neuroscience*, 18. <https://doi.org/10.3389/fnhum.2024.1272121>.

- Schaufeli, W. B., & Buunk, A. (2003). Burnout: An overview of 25 years of research and theorizing. In M. J. Schabracq, J. A. M. Winnubst, & C. L. Cooper (Eds.), *The handbook of work and health psychology*. (pp. 383–425). Wiley.
- Schoenmakers, M., Saygin, M., Sikora, M., Vaessen, T., Noordzij, M., & de Geus, E. (2025). Stress in action wearables database: A database of noninvasive wearable monitors with systematic technical, reliability, validity, and usability information. *Behavior Research Methods*, 57(6), 171. <https://doi.org/10.3758/s13428-025-02685-4>
- Shaw, J. D. (2011). Turnover rates and organizational performance: Review, critique, and research agenda. *Organizational Psychology Review*, 1(3), 187–213. <https://doi.org/10.1177/2041386610382152>
- Shirom, A. (2003). Job-related burnout. In J. C. Quick & L. E. Tetrick (Eds.), *Handbook of occupational health psychology* (pp. 245–265). Washington, DC: American Psychological Association.
- Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the Association for Information Systems*, 13, 24. ProQuest Central. <https://doi.org/10.17705/1CAIS.01324>.
- Tindale, L. C., Chiu, D., Minielly, N., Hrincu, V., Talhouk, A., & Illes, J. (2022). Wearable Biosensors in the Workplace: Perceptions and Perspectives. *Frontiers in Digital Health*, 4. <https://doi.org/10.3389/fdgth.2022.800367>.
- van Lier, H. G., Pieterse, M. E., Garde, A., Postel, M. G., de Haan, H. A., Vollenbroek-Hutten, M. M. R., Schraagen, J. M., & Noordzij, M. L. (2020). A standardized validity assessment

protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. *Behavior Research Methods*, 52(2), 607–629.

<https://doi.org/10.3758/s13428-019-01263-9>.

Villar, R., Beltrame, T., & Hughson, R. L. (2015). Validation of the Hexoskin wearable vest during lying, sitting, standing, and walking activities. *Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition Et Metabolisme*, 40(10), 1019–1024.

<https://doi.org/10.1139/apnm-2015-0140>.

Volpes, G., Valenti, S., Genova, G., Barà, C., Parisi, A., Faes, L., Busacca, A., & Pernice, R. (2024). Wearable Ring-Shaped Biomedical Device for Physiological Monitoring through Finger-Based Acquisition of Electrocardiographic, Photoplethysmographic, and Galvanic Skin Response Signals: Design and Preliminary Measurements. *Biosensors*, 14(4), 205.

<https://doi.org/10.3390/bios14040205>.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, 120(1), 3–24. <https://doi.org/10.1037/0033-2909.120.1.3>.

Wright, R. A., & Gendolla, G. H. E. (Eds.). (2012). *How Motivation Affects Cardiovascular Response: Mechanisms and Applications*. American Psychological Association.

<https://www.jstor.org/stable/j.ctv1chs2jh>.

Wu, W., Pirbhulal, S., Zhang, H., & Mukhopadhyay, S. C. (2019). Quantitative Assessment for Self-Tracking of Acute Stress Based on Triangulation Principle in a Wearable Sensor

System. IEEE Journal of Biomedical and Health Informatics, 23(2), 703–713.

<https://doi.org/10.1109/JBHI.2018.2832069>.

Yang, Y., Liu, C., Yu, H., Shao, D., Tsow, F., & Tao, N. (2016). Motion robust remote photoplethysmography in CIELab color space. *Journal of Biomedical Optics*, 21(11), 117001. <https://doi.org/10.1117/1.JBO.21.11.117001>.

Yu, S., Hu, S., Azorin-Peris, V., Chambers, J. A., Zhu, Y., & Greenwald, S. E. (2011). Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise. *Journal of Biomedical Optics*, 16(7), 077010. <https://doi.org/10.1117/1.3602852>.

Appendix

Appendix A. Ethics Approval



Comité d'éthique de la recherche

Le 03 juin 2025

À l'attention de : Sylvain Sénécal, HEC Montréal

Objet : Approbation éthique de votre projet de recherche

Projet : 2025-6408

Titre du projet de recherche : Étude sur la validité d'un nouvel outil de mesure psychophysiologique non-obstrusif dans un contexte dynamique de travail

Source de financement : CRSNG/PROMPT/Partenaires industriels (R2870B, R2870C, R2870)

Titre de la subvention : Improving Human Machine Interface Learnability : Cognitive Readiness and IT Learning Experience

Bonjour Sylvain Sénécal,

Votre projet de recherche a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains par le CER de HEC Montréal.

Un certificat d'approbation éthique qui atteste de la conformité de votre projet de recherche à la *Politique relative à l'éthique de la recherche avec des êtres humains* de HEC Montréal est émis en date du 03 juin 2025. Prenez note que ce certificat est **valable jusqu'au 03 juin 2026**.

Vous devrez obtenir le renouvellement de votre approbation éthique avant l'expiration de ce certificat à l'aide du formulaire *F7 - Renouvellement annuel*. Un rappel automatique vous sera envoyé par courriel quelques semaines avant l'échéance de votre certificat.

Si des modifications sont apportées à votre projet, vous devrez remplir le formulaire *F8 - Modification de projet* et obtenir l'approbation du CER avant de mettre en oeuvre ces modifications.

Notez qu'en vertu de la *Politique relative à l'éthique de la recherche avec des êtres humains* de HEC Montréal, il est de la responsabilité des chercheurs d'assurer que leurs projets de recherche conservent une approbation éthique pour toute la durée des travaux de recherche et d'informer le CER de la fin de ceux-ci. De plus, toutes modifications significatives du projet doivent être transmises au CER avant leurs applications.

Lorsque votre projet est terminé, vous devrez remplir le formulaire *F9 - Fin de projet (ou F9a - Fin de projet étudiant sous l'égide d'un autre chercheur)*, selon le cas. **Les étudiants doivent remplir un formulaire F9 afin de recevoir l'attestation d'approbation éthique nécessaire au dépôt de leur thèse/mémoire/projet supervisé.**

Vous pouvez dès maintenant procéder à la collecte de données pour laquelle vous avez obtenu ce certificat.

Nous vous souhaitons bon succès dans la réalisation de votre recherche.

Le CER de HEC Montréal

CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains et qu'il satisfait aux exigences de notre politique en cette matière.

Projet # : 2025-6408

Titre du projet de recherche : Étude sur la validité d'un nouvel outil de mesure psychophysiologique non-obstrusif dans un contexte dynamique de travail

Chercheur principal : Sylvain Sénécal, Professeur titulaire, Marketing, HEC Montréal

Cochercheurs : Théophile Demazure; Mehdi Benbouda; Camila Burga Villarreal; Pierre-Majorique Léger; Constantinos K. Coursaris; Marc Fredette; Alexander John Karran; Thaddé Rolon-Mérette; Alejandra Ruiz Segura; Frédérique Bouvier; Luis Carlos Castiblanco; Juan Fernandez Shaw; Andrada Toma; Lucas Dietz; Rayane Benhenni; David Briegne; Salima Tazi; Xavier Côté; Élise Imbeault; François Courtemanche; Shang Lin Chen

Date d'approbation du projet : 03 juin 2025

Date d'entrée en vigueur du certificat : 03 juin 2025

Date d'échéance du certificat : 03 juin 2026



Maurice Lemelin
Président
CER de HEC Montréal

Signé le 2025-06-03 à 10:41

Appendix B. *Participant inclusion status and randomization of task conditions*

Participant	Status	RR Placement	N-back order	T3 condition	T4 condition	T5 condition
P01	Done	0	0	2	0	1
P02	Done	1	0	0	1	2
P03	Done	1	0	1	0	2
P04	Done	0	0	2	0	1
P05	Done	0	0	1	0	2
P06	Done	1	0	0	1	2
P07	Data Unusable	N/A	N/A	N/A	N/A	N/A
P08	Done	1	0	1	2	0
P09	Done	0	0	1	2	0
P10	Done	1	0	0	1	2
P11	Cancelled	N/A	N/A	N/A	N/A	N/A
P12	Cancelled	N/A	N/A	N/A	N/A	N/A
P13	Done	0	0	0	2	1
P14	Cancelled	N/A	N/A	N/A	N/A	N/A
P15	Included	1	0	1	2	0
P16	Included	0	0	2	1	0
P17	Included	1	0	2	1	0
P18	Included	0	0	1	2	0
P19	Included	1	0	0	1	2
P20	Included	0	0	2	1	0
P21	Included	1	0	2	1	0
P22	Included	0	0	0	2	1
P23	Included	0	0	0	1	2
P24	Included	0	0	2	0	1
P25	Included	1	0	2	0	1
P26	Included	1	1	1	0	2
P27	Included	0	1	1	0	2
P28	Included	1	1	1	2	0
P29	Included	1	1	0	2	1
P30	Data Unusable	N/A	N/A	N/A	N/A	N/A
P31	Included	1	1	2	1	0
P32	Included	0	1	1	0	2
P33	Included	1	1	0	2	1
P34	Included	0	1	0	1	2
P35	Included	1	1	1	2	0
P36	Included	0	1	2	0	1
P37	Included	0	1	0	2	1

P38	Included	1	1	2	1	0
P39	Included	1	1	1	0	2
P40	Included	0	1	0	1	2

Note. “Included” refers to participants whose data was successfully collected and used in the analysis. “Cancelled” refers to participants who cancelled their participation attendance. “Data Unesable” refers to participants who completed the session, but whose data could not be used due to technical issues. Therefore, 35 participants were included in the study.