www.sensorsportal.com

Sensors & Transducers
Published by IFSA Publishing, S. L., 2024
http://www.sensorsportal.com

# Adaptive System to Enhance Operator Engagement during Smart Manufacturing Work

[1, *] Loïc COUTURE, [2] Mario PASSALACQUA, [3] Laurent JOBLOT, [4] Florian MAGNANI, [2] Robert PELLERIN and [1] Pierre-Majorique LÉGER
[1] HEC Montréal, Tech3Lab, 3000 Chemin de la Côte-Sainte-Catherine, H3T 2A7, Canada
[2] Polytechnique Montréal, 2500 Chemin de Polytechnique, H3T 1J4, Canada
[3] Arts et Métiers Institute of Technology, LISPEN / HESAM University / UBFC, France
[4] Université Aix-Marseille, CERGAM, Centrale Méditerranée, 38, rue Frédéric Joliot-Curie, 13013, France
[1] Tel.: +15149981557
[*] E-mail: loic.couture@hec.ca

*Received: 6 April 2024 / Accepted: 14 May 2024 / Published: 30 May 2024*

**Abstract:** Sustaining optimal task engagement is becoming vital in smart factories, where manufacturing operators' roles are increasingly shifting from hands-on machinery tasks to supervising complex automated systems. However, because monitoring tasks are inherently less engaging than manual operation tasks, operators may have a growing difficulty in keeping the optimal levels of engagement required to detect system errors in highly automated environments. Addressing this issue, we created an adaptive task engagement feedback system designed to enhance manufacturing operators' engagement while working with highly automated systems. Utilizing real-time acceleration, heart rate, and respiration rate data, our system provides an intuitive visual representation of an operator's engagement level through a color gradient, ensuring operators can stay informed of their engagement levels in real-time and make prompt adjustments if required. This article elaborates on the six-step process that guided the development of this adaptive feedback system. We developed a task engagement index by leveraging the physiological distinctions between more and less engaging manufacturing scenarios and using automation to induce lower engagement. This index demonstrates a prediction accuracy rate of 80.95 % for engagement levels, as demonstrated by a logistic regression model employing leave-one-out cross-validation. The implications of deploying this adaptive system include enhanced operator engagement, higher productivity and improved safety measures.

**Keywords:** Engagement, Adaptive system, Manufacturing, Industry 5.0, Human-machine interaction, Design science.

## 1. Introduction

Recent advances in manufacturing technologies have significantly expanded the capabilities of automation, enabling even traditionally human-centric tasks to be automated. When automating such tasks, we frequently see the role of human operators transitioning from manual labor to supervisory roles, which can have negative implications for operators' engagement in their work [1]. In the context of Industry 5.0, which emphasizes the importance of workers' interests and well-being, ensuring that operators remain engaged and stimulated in their roles becomes crucial [2, 3]. This approach is not only fundamental to their personal development but is also imperative for enhancing their decision-making skills, especially in increasingly complex work environments [3-5]. By prioritizing the engagement and stimulation of operators, organizations can navigate the challenges of modern manufacturing landscapes more effectively, ensuring that technological advancements contribute positively to the work experience of human operators

[6]. Engagement varies in definition across the literature. It is often used either as “task engagement” or "operator engagement" to describe the effective allocation of attentional resources towards the task objectives [7-9]. This definition focuses on the cognitive aspects of the worker experience. Alternatively, terms like “work engagement” and “employee engagement” are used to characterize a positive, fulfilling psychological state related to work [10-13] that encompasses the cognitive, behavioral, and emotional aspects of the work experience. Given the prevalent focus on the cognitive dimension of engagement in existing research on human-machine interaction, we employ ‘task engagement’ to represent the cognitive aspect of the work experience. This decision acknowledges the broader spectrum of engagement but aligns our focus with the extensive body of research emphasizing cognitive engagement in the context of human-machine interaction.

To effectively oversee automated systems, an operator must remain alert to various signals, referred to as arousal, and must stay focused on the task at hand, known as task engagement [9]. There is a sweet spot of arousal and task engagement that operators need to sustain to ensure adequate and optimal monitoring [9]. If an operator is disengaged, they risk becoming distracted with mind-wandering [14], whereas being overly engaged can lead to tunnel vision, preventing the operator from staying alert to external signals [15]. Similarly, if an operator has too high or too low arousal, it might affect their cognitive capabilities [9]. However, it can be challenging for operators to maintain an optimal level of engagement in monitoring tasks, mainly because monitoring tasks are generally less engaging than manual operation tasks [1]. Consequently, an under-stimulated operator is much more likely to be distracted, which reduces their ability to detect system errors [16, 17]. This monitoring difficulty increases as the level of automation rises [18]. Therefore, in increasingly intelligent factories, there may be a growing difficulty in detecting errors in automated systems.

To limit these performance declines, one method is to ensure that operators can maintain optimal levels of task engagement during their work [19]. The work of Karran *et al.* [20] is particularly promising in this regard. Their research demonstrated the potential of using real-time engagement level feedback to improve users’ attentiveness during a passive monitoring task. In their paper, they used an adaptive system developed by Demazure *et al.* [21] that continuously informed operators of their level of engagement in the task through a color gradient, using electroencephalography (EEG) measurements to infer engagement. While this solution has shown promising results, it faces significant limitations in a manufacturing context, primarily due to the high sensitivity of EEG to movement. Therefore, our study seeks to adapt this approach for manufacturing, aiming to develop a tool designed to help manufacturing operators maintain optimal engagement levels when working with highly automated systems. The primary aim of this adaptation is to leverage engagement metrics collectible from mobile operators. The research question that guided the system’s design is: How can the engagement feedback system proposed by Demazure *et al.* [21] be effectively adapted and implemented in a manufacturing setting to monitor and enhance the engagement of mobile manufacturing operators?

The structure of this article is outlined as follows. Section 2 contains an overview of the current solutions to enhance operator engagement, why we hypothesize that adaptive feedback systems represent a good solution, and which methods are used to measure task engagement in the literature. Section 3 contains the research objectives that guided our design. In Section 4, we detail the six-step process that led to developing this new innovative feedback system. The results we obtained during the design process are detailed in Section 5. In Section 6, we discuss the results, and in Section 7, we provide our concluding remarks along with limitations of the system and insights into future developments.

## 2. Background

### 2.1. Solutions to Enhance Task Engagement During Monitoring Tasks

Solutions to keep operators cognitively engaged during monitoring tasks can be categorized into multi-tasking and adaptive interfaces. Multi-tasking involves engaging the operator with non-task-related tasks when they experience disengagement. Naujoks *et al.* [22] showed that engaging in secondary tasks reduced the reaction time of drivers when they needed to regain control, and Atchley *et al.* [23] showed that talking while driving could improve driving performance. However, one limitation of multi-tasking is that it requires the operator to divert some of their attention to a secondary task, which diminishes the total level of engagement they can apply to the primary task [24]. For this reason, adaptive systems appear to be a better alternative for keeping operators engaged when monitoring systems. Adaptive systems infer the cognitive state of human operators and use this information to adapt in real time. According to Hinss *et al.* [25], there are two main types of adaptive systems: adaptive automation and adaptive interfaces. Adaptive automation allows for the dynamic adjustment of task allocation based on the cognitive state of operators [26, 27]. The purpose of these systems is to reduce the level of automation of automated systems when decreases in engagement are detected. This decrease in automation necessitates that the operator takes on more stimulating tasks, thereby potentially restoring their engagement to a level considered adequate for environments characterized by higher levels of automation.

The second type of adaptive system consists of adaptive interfaces. Feigh *et al.* [27] developed a taxonomy for adaptive interfaces, proposing four modalities of adaptation, including task allocation, which refers to adaptive automation. The three other modalities are the following. When operators are in a state of cognitive disengagement adaptive interfaces can adjust task prioritization, for example, by requesting the operator to perform tasks that are either more stimulating or require less engagement. They can also adapt the interaction between the operator and the system, for instance, by changing the layout of components or the mode of interaction (e.g., from haptic to vocal). Lastly, the content of the interface can be adapted, for example, by increasing the amount of information displayed when the operator is engaged and reducing it when they are less engaged. Fig. 1 summarizes the different solutions to keep operators engaged based on the works of Feigh *et al.* [27] and Hinss *et al.* [25].

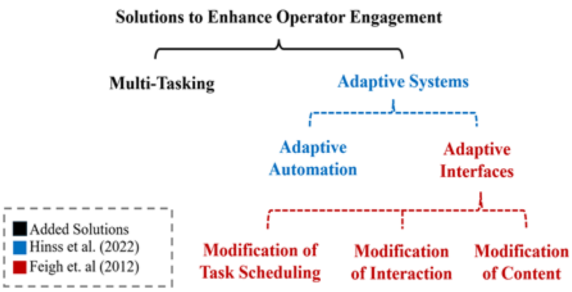


**Fig. 1.** Categorization of current solutions to enhance operator engagement during surveillance work.

Despite the introduction of adaptive automation and adaptive interfaces years ago, very few adaptive systems with generalizable applications have been developed [28]. The reasons for this include the absence of comprehensive multimodal models to infer operators' cognitive states [28], the dependence of adaptive systems applications on the specific working environment in which they are developed, and constraints regarding the physiological data collection across different work environments. However, feedback systems and alarm systems distinguish themselves from other adaptive systems solutions by providing a passive solution that does not need to interface with various systems and introduces minimal distraction, making it relatively easy to apply across different contexts. These systems would fall into the adaptive interfaces category by modifying the content of the interface (e.g., the presence or absence of visual or auditory cues). The key distinction between feedback and alert systems lies in the way they present countermeasures. Adaptive feedback systems usually give continuous feedback to operators on their cognitive and emotional states, whereas alert systems typically wait for specific thresholds to be reached before notifying the operator. Karran *et al.* [20] compared these two approaches, revealing that the continuous display of mental state had a greater impact on operator engagement compared to displaying the mental state after a disengagement threshold was reached. Therefore, we opted for the development of an adaptive feedback system.

The feedback system developed by Demazure *et al.* [21] offers a promising approach to enhancing task engagement during monitoring tasks by providing operators with real-time feedback on their level of engagement. This innovation keeps operators continuously aware of their engagement, promoting immediate adjustments as needed. Currently, the manufacturing sector lacks such adaptive feedback systems specifically designed for engagement levels. Despite this, the broad applicability and proven effectiveness of engagement-level feedback systems underscore their potential value in maintaining operator engagement, particularly in environments requiring high task engagement.

## 2.2. Measuring Task Engagement

Task engagement, or the cognitive aspect of work engagement, is most commonly measured using questionnaires or observational metrics. The most commonly used questionnaire to measure engagement is the Utrecht Work Engagement Scale (UWES) [29]. The UWES questionnaire facilitates the measurement of the multi-dimensional concept of engagement, defined earlier. It encompasses three dimensions: cognitive engagement in work (related to the absorption dimension of the questionnaire), behavioral engagement at work (related to the vigor dimension of the questionnaire), and emotional engagement (related to the dedication dimension of the questionnaire). Given that the absorption dimension encapsulates the cognitive aspect of engagement, it can be used as a measure of task engagement. Regarding observational metrics, task performance metrics are the most commonly utilized measures. Performance-based measures of task engagement include, for example, error detection performance [1-30], sampling time [31], and reaction time [32]. While performance-based and subjective metrics effectively identify instances of lower operator engagement when monitoring automated systems, both these measures have their limitations. Questionnaires, which depend on post-task subjective assessments, are prone to biases such as recall bias [33]. Performance metrics, while serving as useful engagement proxies, do not directly measure engagement and can be influenced by various extraneous factors. A solution to complement the limitations of questionnaires and performance metrics is the use of physiological measures, which allow for the continuous measurement of the participant's state throughout the task, without interference, thus limiting biases [34]. Consequently, recent research has increasingly focused on physiological metrics to understand the impact of operators' mental states on performance, providing deeper insights into engagement dynamics [33, 35, 36]. The physiological measures used to infer task engagement include eye-tracking, electroencephalography (EEG), heart rate variability (HRV), respiration rate (RR), electrodermal activity (EDA), and functional near-infrared spectroscopy (fNIRS). These various modalities and their application in measuring task engagement are presented in this section.

### 2.2.1. Eye-tracking

Eye-tracking tools detect where an operator's gaze lands and measure pupil diameter, both indicators of task engagement [37]. An operator is considered engaged when their gaze is on the main points of interest of a task and disengaged when their gaze deviates from them. When analyzing the gaze of operators, we typically distinguish between fixations and saccades. Fixations refer to the moments when the eyes are relatively stationary and are focused on a specific point for a period of time, generally lasting between 180 and 330 milliseconds [38], while saccades are rapid eye movements between fixations. One way to interpret task engagement using gaze data is to use the position, frequency, and duration of fixations [39]. Pupil diameter is used as an indicator of task engagement and cognitive fatigue [40] because this physiological mechanism is controlled by the locus ceruleus norepinephrine system (LC-NE) region of the central nervous system, which is also responsible for regulating attention [41].

### 2.2.2. Electroencephalography (EEG)

Several EEG metrics are used as measures of task engagement [42]. The most common task engagement metric is the Engagement Index, corresponding to the ratio between beta and the addition of alpha and theta wave power [7].

$$Engagement\ Index = \beta/(\alpha + \theta) \qquad (1)$$

Additionally, since the beta signal power is associated with a state of alertness and cognitive engagement and the alpha signal power with a state of relaxation, the ratio of beta to alpha is used to reflect arousal levels [43]. P3 event-related amplitudes are also often used to measure task engagement because of their close link with motivation and attention [40-44].

### 2.2.3 Heart Rate Variability (HRV)

HRV is defined as the variation of time intervals between consecutive heartbeats [45] and is mainly used as a measure of the activation of the autonomous nervous system (ANS) [46]. Many metrics can be derived from HRV, but the most commonly used are the power of the high-frequency band of HRV (0.15–0.4 Hz) (HF-HRV), the power of the low-frequency band of HRV (0.04–0.15 Hz) (LF-HRV), the ratio of LF-to-HF power, the standard deviation of normal sinus beats (SDNN) and the root mean square of successive RR interval differences (RMSSD). More details on all HRV metrics can be found in [47]. To accurately interpret the various measures of Heart Rate Variability (HRV), a brief overview of the Autonomic Nervous System (ANS) is essential. The ANS is governed by two primary components: the parasympathetic and sympathetic systems. The parasympathetic system orchestrates the body's "rest and digest" responses, promoting relaxation and energy conservation. Conversely, the sympathetic system triggers the "fight or flight" responses, preparing the body for action and mobilizing energy resources. Higher activation of the parasympathetic system is usually associated with better cognitive performance [48] and a better capacity for cognitive engagement [49].

HF-HRV reflects parasympathetic activation, where higher HF-HRV is associated with greater activation of the parasympathetic system [47]. Since HF-HRV reflects parasympathetic activation, higher HF-HRV is associated with a higher capacity for cognitive engagement.

There is a certain ambiguity concerning the mechanisms underlying the LF-HRV. It may be produced by the sympathetic nervous system, parasympathetic nervous system, or baroreceptors [47]. Because of this ambiguity, there is no apparent interpretation of the LF-HRV in the literature. However, because of the potential link between LF-HRV and the sympathetic nervous system, the ratio of LF/HF has been used to reflect the ratio of parasympathetic to sympathetic activation [47]. Since parasympathetic activation is linked to better cognitive performance [48-49], lower values of LF/HF can be associated with higher capacity for task engagement.

RMSSD reflects the beat-to-beat variance in heart rate and is used to assess short-term heart rate variability. For ultra-short recordings of HRV (under 5 minutes), the RMSSD is correlated with HF-HRV and is usually the primary time domain metric used to estimate the vagally-mediated changes reflected In HRV [46]. Higher RMSSD is typically associated with higher parasympathetic activation, leading to a better cognitive engagement capacity. RMSSD has also been shown to decrease with task difficulty [50].

SDNN represents the overall variability in heart rate and is usually used to assess global heart rate variability in longer-term HRV measurements. Higher overall variability is associated with a better capacity for cognitive engagement.

### 2.2.4. Functional Near-infrared Spectroscopy (fNIRS)

Functional near-infrared spectroscopy measures the change in blood oxygenation in the brain's cortex and is often used in combination with EEG [20, 51].

The interest in using fNIRS with EEG is to leverage the spatial resolution of fNIRS with the temporal resolution of EEG [20]. Verdière *et al.* [52] have shown that fNIRS signals could be used to detect higher or lower task engagement during a piloting task. They also showed that connectivity measures lead to better classification performance than oxygenation measures.

#### 2.2.5. Electrodermal Activity (EDA) and Respiration Rate (RR)

Electrodermal activity reflects the skin's conductivity level and is used to indicate a state of arousal or stress [42, 53]. As for respiration, respiration rate has been found to have a significant positive relationship with task engagement [54].

Although eye-tracking and EEG methods are well-established in the literature for assessing task engagement [7, 42, 55], their practical application in manufacturing faces significant challenges, primarily due to the movement and dynamic environment in which manufacturing operators must operate. Due to these limitations, it is proposed to use measures of alternative metrics, such as respiration rate and HRV. Despite these metrics being less frequently utilized and explored in the literature on human-machine interaction, they are more easily applicable in a manufacturing setting due to their low cost, low intrusiveness, and low sensitivity to movement [56, 57].

### 2.3. Proposed Approach

The constraints inherent to the manufacturing sector make Moray and Inagaki's [58] approach particularly appealing. Their method evaluates monitoring performance by contrasting actual operator performance to an optimal standard. From this perspective, for any specific task, it seems feasible to establish a performance metric by initially recording the responses of an operator in a high-performance scenario and comparing it to a low-performance scenario. Therefore, in the case of operator engagement, a similar approach would be to establish an engagement metric by comparing physiological responses recorded in highly engaging scenarios with those from a minimally engaging scenario, using contrast to construct a reliable measure of engagement for this task [19]. To create high and low engagement scenarios, we can use the approach of Verdière *et al.* [52], who manipulated the level of automation to create more and less engaging piloting tasks. This approach is consistent with the findings that showed that higher automation can reduce operator engagement [1].

Hence, to maintain optimal engagement levels of manufacturing operators within their dynamic work environments, our proposal involves developing a new adaptive engagement feedback system inspired by the research of Demazure *et al.* [21] but tailored to the manufacturing context. Rather than depending on exact real-time engagement metrics and measurements, our system follows a methodology inspired by the work of Moray and Inagaki [58], leveraging physiological indicators that differentiate between optimal and suboptimal engagement states. A significant advantage of this approach is its adaptability to complex settings like manufacturing, where constraints exist concerning the feasibility of specific physiological measurements, such as eye-tracking and EEG.

## 3. Objectives

To guide the design process, we established three research objectives: (i) Identify the most appropriate physiological tools for measuring task engagement in a manufacturing context; (ii) Identify and characterize the physiological differences between "high" and "low" engagement manufacturing scenarios; and (iii) Develop an interface that intuitively translates the identified physiological markers into a color gradient, offering immediate feedback on engagement levels. While developing our system, we encountered two significant design challenges that needed careful consideration. The first challenge concerned the optimal display modality for the color gradient, which is crucial for providing clear and understandable feedback on engagement levels. The second challenge involved devising an engagement index scaling method that accurately reflects engagement levels, ensuring that the system's feedback is both intuitive and effective. To address these challenges, we introduced two additional objectives: (iv) Determine the most effective visual representation of engagement through a comparative analysis of a continuous color gradient with 100 shades versus a discrete color gradient with three distinct colors, and (v) Identify the optimal method for scaling the engagement index that accurately reflects perceived engagement, facilitating easier interpretation of physiological markers of engagement by users. With the addition of these two objectives, we were able to make informed design choices that significantly enhanced the usability and effectiveness of our system.

## 4. Methods

We used a design science methodology to develop a task engagement feedback system involving a six-step process that included three studies (see Table 1). We first selected non-intrusive physiological tools to measure task engagement in a manufacturing assembly context. Then, we collected physiological, performance, and subjective data during "high" and "low" engagement manufacturing scenarios. We identified the physiological differences between the "high" and "low" engagement scenarios and used these markers to design a task-specific "engagement index". Using this formula, we developed the first version of the feedback system. We then evaluated the best display modality and the best scaling method for our engagement index, which were critical aspects of our feature selection process.

**Table 1.** Methodology Employed to Design the Feedback System.

| Step | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 |
|---|---|---|---|---|---|---|
| **Title** | ***Select Physiological Tools*** | ***Collect Data*** | ***Identify Markers*** | ***System Design*** | ***Display Validation*** | ***Scaling Validation*** |
| **Description** | Comparative analysis of task engagement measures collected with EEG, fNIRS, ECG, eye-tracking and breathing bands | Study 1: Collection of Physiological Data in Scenarios with Varied Engagement Levels | Identify physiological markers of engagement | – | *Study 2:* Validating multiple display modalities of engagement | *Study 3:* Validating multiple index scaling methods |
| **Experimental design** | – | Within–subject | – | – | Within–subject | Between subject |
| **Conditions** | – | No automation<br>Automation | – | – | Discrete color gradient (3 shades of color)<br>Continuous color gradient (100 shades between green and red) | Min/Max since the beginning of the task<br>Min/Max of training data<br>Min = 25th & Max = 75th percentiles since the beginning of the task |
| **Experimental manipulation** | – | Manufacturing Q&A and assembly tasks using snowshoes | Feature extraction using a logistic regression model<br>Validation with LOOCV | – | Fully automated manufacturing Q&A and assembly tasks using images of snowshoes | Fully automated manufacturing Q&A and assembly tasks using images of snowshoes |
| **Data** | Literature review | Collected physiological data (bpm, breath rate, motion) and perceived work engagement (UWES) | Task 1 & Task 2 data from step 1 | – | 10 minutes semi-directed interviews | Three-item questionnaire |
| **Participants** | – | 22 participants | – | – | 3 participants | 10 participants |

## 4.1. Step 1 - Choosing Physiological Tools Suitable for a Manufacturing Environment

A thorough methodological reflection was necessary to select the tools and measurements most suited for a manufacturing environment. Our selection criteria dictated that (i) the tool must be non-intrusive for a manufacturing assembly context, (ii) ensure easy data collection, and (iii) provide reliable measurements. Since EEG and fNIRS require wearing a headset, these technologies were deemed too intrusive and potentially distracting for operators in a manufacturing context. Additionally, these technologies are highly sensitive to movement, which is not ideal in a manufacturing setting where the operator must perform physical work. Moreover, electrodermal activity is typically collected on the palm, which would have restricted operators in their assembly tasks. For this reason, EDA was also dismissed for intrusiveness. Given that manufacturing operators often need to interact with a 3D environment, static eye-tracking devices were ruled out. We conducted a pilot test with Tobii Pro glasses (Tobii Technologies, Danderyd, Sweden) that allow the collection of eye-tracking data for users in movement. However, we concluded that using these glasses would overly complicate data collection due to the low battery life and the lack of available tools for analyzing operators' attention when interacting with a 3D environment. This resulted in a preference for electrocardiography and respiration measurements. The Hexoskin vest (Carré Technologies, Montreal, Canada) was found to be a non-intrusive tool that allowed for the simultaneous measurement of these two parameters, as well as accelerometry data. Moreover, heart rate and respiration rate measurements obtained from the Hexoskin vest show little variation compared to laboratory-grade electrocardiograms and metabolic carts, as evidenced by [59]. Given its accuracy and non-intrusiveness, the Hexoskin vest was selected for the development of our system.

### 4.2. Step 2 – Collect Data in More and Less Engaging Manufacturing Scenarios

In this step, we collected physiological and perceptual data from participants in more and less engaging manufacturing situations. We recruited 22 participants (age = 21.62±3.17; men = 14) for a within-subject experiment, in which they twice performed a quality control and assembly task on a simulated assembly line. All participants provided a signed consent in-line with the University ethics committee (project # 2023-5058) and were compensated with 40 euros. The task explained in more detail in [1], required participants to detect errors on partially assembled snowshoes and complete the assembly by fixing the binding to the base at its pivot point (see Fig. 2). In the "less engaging" condition, we automated the participants' decision-making, equipping them with a fully reliable error detection system that indicated to the operator whether a snowshoe had a defect. In the "more engaging" condition, participants had to manually detect errors before assembling the snowshoes. During each task, a total of 30 snowshoes had to be assembled by the participants, with six being defective. Participants realized the task once with automated support and once without automated support, with condition order being randomly assigned and counterbalanced. During the task, we collected physiological data using a Hexoskin vest, recording heart rate, respiration rate, and acceleration data. We also collected perceived cognitive absorption, vigor, and dedication using the Utrecht Work Engagement Scale (UWES) [29], which was collected post-task. Since our study specifically aims to modulate the cognitive aspect of work engagement, the absorption dimension is employed as a subjective measure of task engagement within our study. The raw physiological data from the Hexoskin was pre-processed and synchronized using the COBALT Photobooth software [60]. The list of physiological and self-reported data collected can be found in Table 2.

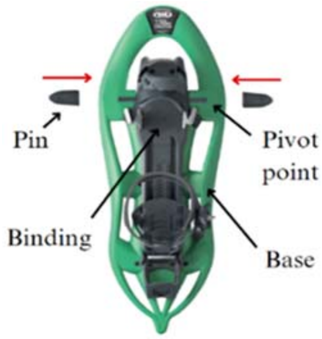


**Fig. 2**. Product Used in the Manufacturing Task.

**Table 2.** List of Collected Variables.

| Type of data | Measure | Description |
|---|---|---|
| Physiological data | Beats per minute | Number of beats per minute |
| | SDNN | Standard deviation of NN intervals |
| | LF | Power of the Low-frequency band (0.04-0.15 Hz) ($ms^2$) |
| | HF | Power of the High-frequency band (0.15-0.4 Hz) ($ms^2$) |
| | LF/HF | Ratio of Low-to-High frequency power |
| | Breathing Rate | Number of respirations per minute |
| | Minute Ventilation | Respiratory volume per minute (L/min) |
| | Cadence | Number of steps per minute |
| | Motion | Norm of the 3D acceleration vector (G) |
| Self-reported measures | Absorption score (UWES) | Perceived absorption (cognitive engagement) |

### 4.3. Step 3 – Identify Physiological Markers of Engagement and Create an Engagement Index

In this step, we began by validating our primary assumption that the condition with automation was less engaging than the manual condition. We compared the perceived absorption scores between automated and manual conditions using a one-sided Wilcoxon signed-rank test, which is suitable for evaluating non-parametric paired data. The analysis revealed a statistically significant difference in perceived absorption scores when comparing manual and automated conditions ($p$ = .0008), with the automated condition showing lower perceived absorption scores than the manual condition. This result aligns with our primary assumption that the automated condition was less engaging than the automated condition. Based on this result, we then categorized the data, assigning labels of "high" or "low" engagement to arrays of data, depending on the condition experienced by the participant. Data from the automated task was labeled as "low engagement", while data from the manual task was labeled as "high engagement". We then defined a task-specific engagement index (TS-EI) using the three physiological variables with the highest estimated weights in the logistic regression model used to predict the probability of a participant being more engaged in the task. The formula represents a weighted sum, where each coefficient corresponds to the respective variable's estimated power to predict if a participant is in a "high" or "low" state of engagement. The formula is based on 30-second data windows.

$$TS_{EI} = (435.7\,Motion_{std}) - \\ -(175.4\,Motion_{mean}) + \\ +(0.78\,RespirationRate_{std}) \quad (2)$$

Without a testing dataset, we validated Eq. (2) using the Leave-One-Out Cross-Validation (LOOCV) on the same dataset. We employed the LOOCV in a logistic regression model to predict if a participant's engagement during a task was "higher" or "lower". The results of this test demonstrated an average predictive accuracy of 80.95 % on the leave-out samples.

### 4.4. Step 4 – Design the Feedback System

In this step, we developed an initial version of the feedback system. To guide our development process, we established 5 main requirements: (i) The system must collect the user's respiration rate and acceleration data in real-time, (ii) communicate the user's task engagement in real-time using a color gradient, (iii) the displayed color must represent the operator's perceived engagement level, (iv) the system must be easy to interpret, and (v) it should not distract the operator during their task. An overview of the designed system can be found in Fig. 3.

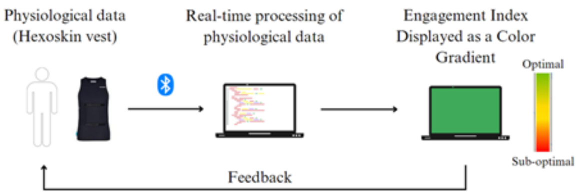


**Fig. 3.** Overview of the Adaptive Feedback System.

The system, developed in Python, receives respiration and acceleration data from the Hexoskin vest, which transmits data at a frequency of 1 Hz. Specifically, respiration rate data reflects the average number of respirations per minute based on the last seven breathing cycles, and acceleration data represents the average norm of the 3D acceleration vector over the last second. Our system received data encoded in UTF-8 through a Bluetooth Low Energy (BLE) connection directly established with the Hexoskin vest. It was possible to establish a direct connection using the UUID keys of the respiration and acceleration Bluetooth channels available in Hexoskin's documentation. The system included a Bluetooth reconnection protocol in case of connection failure. Eq. (2) was used by the system to calculate the task-specific Engagement Index based on 30-second data windows (or 30 data points, considering that the frequency of transmission of the Hexoskin is 1 Hz). In the first version of the system, the index was normalized using the minimum and maximum index values recorded since the beginning of the session and then scaled as an integer between 0 and 100. Based on the normalized index value, it was possible to select the color to be displayed. The color selection varied according to the display modality, mainly whether the color gradient was discrete (with 3 distinct colors) or continuous (with 100 shades of color). For the continuous gradient, we created a 1×100 matrix with a palette of 100 shades ranging from green to red and used the normalized index value to specify the color code to be fetched from the matrix. For a discrete gradient, only three colors were possible: green for normalized index values above 66 %, yellow for values between 33 % and 66 %, and red for values below 33 %. The color codes chosen were then sent via Wi-Fi for display at a frequency of 1 Hz. The system's architecture and the specific open-source Python Libraries used in the code are detailed in Fig. 4 and Table 3.

**Table 3**. Python Open-Source Libraries used by the System.

| Library | Description | License | Link |
|---|---|---|---|
| Bleak | BLE communication | MIT | https://github.com/hbldh/bleak |
| Colour | Color code generation | BSD | https://pypi.org/project/colour/ |

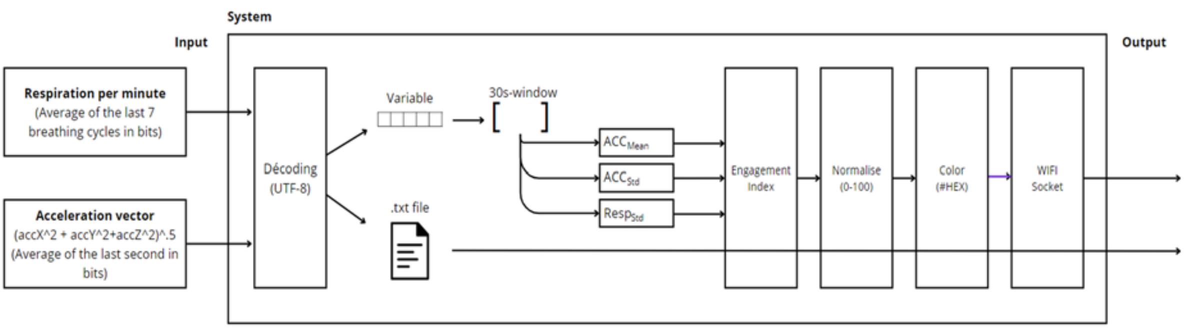


**Fig. 4.** Architecture of the Adaptive Feedback System.

### 4.5. Step 5 – Validation of the Display Modality

In this step, we assessed whether representing the index through a continuous color gradient (100 shades) or a discrete color gradient (3 colors) was more effective in conveying participants' engagement levels. We recruited three participants for a within-subjects pilot test. Each participant completed a low-fidelity version of the automated assembly task twice using printed images of snowshoes instead of authentic snowshoes, experiencing the feedback system in both formats. After completing each task, participants underwent a 5-minute semi-directed interview. During this interview, they were asked about the interpretability of the color gradient, the potential distractions caused by the system, and its effectiveness in representing their engagement levels. Positive and negative statements in each category were compiled and analyzed, revealing that the discrete color gradient was more distracting than the continuous color gradient. This led to the decision to retain the continuous color gradient.

### 4.6. Step 6 – Comparative Analysis of Three Scaling Methods

In the sixth step, we aimed to identify the most effective method for scaling the index. We tested three scaling methods: (i) dynamically adjusting the minimum and maximum values based on the minimum and maximum engagement index values recorded since the beginning of the task for this operator, (ii) using the minimum and maximum values of the training dataset, measured with Eq. (3) to exclude outliers, and (iii) dynamically setting the minimum and maximum values respectively to the 25$^{th}$ (Q1) and 75$^{th}$ (Q3) percentile of the engagement index data measured for this operator since the beginning of the task. A visual representation of each method can be found in Fig. 5.

$$MIN/MAX = TS\_EI_{mean} \pm 3 * TS\_EI_{std} \quad (3)$$

We performed a between-subjects experiment with 10 participants who each completed a low-fidelity version of the manufacturing task while receiving feedback from the system in one of its three possible formats. For this low-fidelity version of the manufacturing task, we asked users to identify errors on printed images of snowshoes instead of real snowshoes. After completing the task, participants were asked to rate the color display's representativeness, interpretability, and distractive nature on a scale from 0 to 100. Method (ii) emerged as the most representative of perceived engagement, leading to its selection for the final design. No differences were found in interpretability and distractive nature between the three methods.

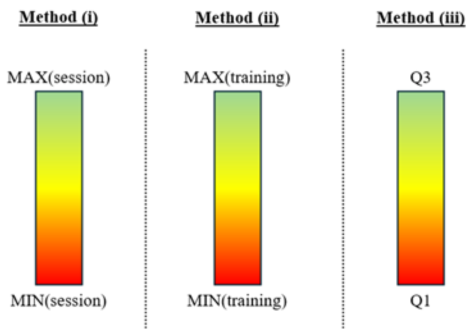


**Fig. 5.** Visual Representation of the Three Index Scaling Methods.

## 5. Results

The one-sided Wilcoxon signed-rank test applied in step three demonstrated a statistically significant difference in perceived absorption scores between automated and manual conditions ($p$ =.0008). This finding suggests that the distribution of the difference of absorption between automated and manual conditions, is not symmetric around zero, predominantly featuring negative values. This asymmetry suggests that perceived absorption scores are typically lower in the automated condition than in the manual condition, which supports the primary assumption that the automated condition was less engaging than the manual condition. Table 4 and Fig. 6 offer an overview of the distribution of reported absorption scores.

**Table 4.** Descriptive Analysis of UWES Absorption Scores between Manual and Automated Conditions.

| | Min | Q1 | Med | Q3 | Max | Mean | Std |
|---|---|---|---|---|---|---|---|
| **Manual** | 2.67 | 3.67 | 4.00 | 5.25 | 6.00 | 4.21 | 0.99 |
| **Automated** | 1.00 | 2.67 | 3.17 | 4.25 | 5.33 | 3.32 | 1.13 |

Using Formula 2 to predict if a participant was in a “high” or “low” state of engagement in a logistic regression model, we achieved an average of 81.31 % accuracy on the training set and 80.95 % on the testing set, as confirmed through leave-one-out cross-validation. For step five, where we assessed the display modality, we employed a qualitative labeling technique to categorize interview statements into three themes: effect on perceived engagement, distraction, and representativeness. The number of statements in each category was then compiled (see Table 5), showing that the discrete color gradient was more distracting (0 positive, six negative statements) than the continuous color gradient (2 positive, 0 negative statements).

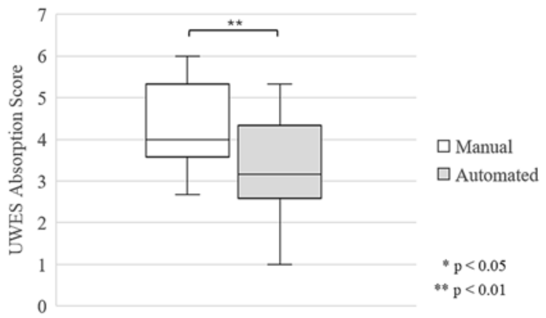


**Fig. 6.** UWES Absorption Scores Distributions between Manual and Automated Conditions.

**Table 5.** Compilation of Qualitative Statements on Continuous and Discrete Color Gradients.

| | Perceived effect on engagement | | Distraction | | Representa-tiveness | |
|---|---|---|---|---|---|---|
| | (+) | (-) | (+) | (-) | (+) | (-) |
| Continuous | 5 | 0 | 2 | 0 | 2 | 2 |
| Discrete | 2 | 1 | 0 | 6 | 0 | 3 |

In step six, the self-reported data from questionnaires revealed that all methods were equally easy to interpret and not distracting. However, the scaling method (ii) utilizing the minimum and maximum values from the training dataset proved to be more representative, with a mean score of 93.33 ±6.24 %. This was in contrast to the scaling method (i), which was based on the minimum and maximum values since the beginning of the task (mean = 57.33±12.28 %), and method (iii) which was based on percentiles (mean = 45.5±14.5 %), as illustrated in Fig. 7. Based on these analyses, we concluded that the continuous color gradient and scaling method, which utilized the minimum and maximum values of the training dataset, i.e., method (ii), was the preferred option.

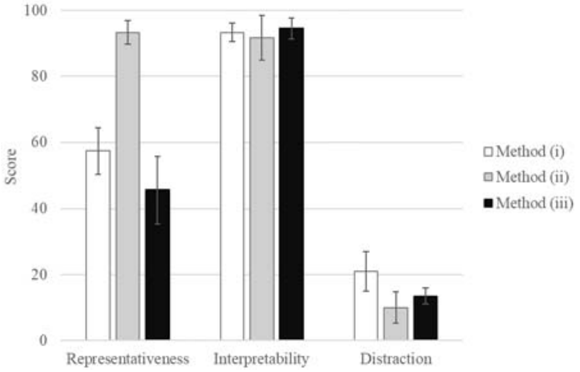


**Fig. 7.** Scaling Method Comparison: Evaluating Representativeness, Interpretability, and Distraction.

## 6. Discussion

The objectives of this study were to (i) identify the most suitable physiological tools for measuring task engagement in a manufacturing context, (ii) discern physiological differences between more and less engaging manufacturing situations, (iii) develop an adaptive feedback system that translates these physiological differences into a color gradient for immediate feedback on task engagement, (iv) determine the best mode of displaying engagement between a discrete and a continuous color gradient, and finally (v) find the most representative normalization method for the engagement index as perceived by operators.

For objective (i), we compared eye tracking, EEG, fNIRS, EDA, electrocardiography (ECG), and respiratory rate monitoring tools against three criteria: (a) data collection tools must not distract or disturb the operator during their work, (b) they must allow for easy real-time data collection, and (c) they must provide reliable measurements. EEG, fNIRS, and EDA systems were deemed unsuitable for manufacturing due to their intrusiveness and limitations in dynamic settings. Similarly, static eye-tracking systems failed in 3D environments, and eye-tracking glasses faced battery and analytical challenges. In contrast, ECG and respiratory rate monitoring, conducted via a Hexoskin vest, provided non-intrusive, reliable data collection of engagement metrics, proving effective for manufacturing environments. ECG and respiration metrics are less frequently utilized in the human-machine interaction literature. However, HRV (an ECG metric) has been shown to correlate with well-established engagement metrics such as EEG and eye-tracking, as documented in aviation scenarios by Roy *et al.* [62]. Additionally, the study by Venables and Fairclough [54] illustrates that, within their research context, respiration exhibited a stronger correlation with engagement than EEG metrics. While additional validation of respiration as a metric of engagement is required, these findings support the potential utility of ECG metrics and respiration in measuring task engagement.

For objective (ii), we simulated a manufacturing environment and subjected participants to varying engagement levels, using automation to reduce engagement. Participants in the automated condition reported lower absorption scores in the UWES questionnaire, indicating lower perceived cognitive engagement during the automated manufacturing task. This result aligns with previous findings that showed that higher levels of automation can lead to lower task engagement [1].

Based on these findings, we analyzed physiological differences between automated and manual conditions to identify physiological features that could be used to construct a task engagement index. Our observations indicated that participants in the manual condition (condition of higher cognitive engagement) had, on average, lower acceleration means and greater acceleration variability. Without the aid of an error detection tool, participants in the manual condition had to take the time to analyze each product for longer periods than in the automated condition. This contributed to a lower acceleration mean for the manual condition, while the acceleration when fetching a new product increased variability. Considering that a manufacturing operator might not be as focused when moving around as they are when stationary at their worktable, these results suggest the potential use of acceleration mean and acceleration variability as indicators of task engagement. Despite the absence of observable differences in respiration rates across conditions, it was noted that participants engaged in the manual condition (a scenario characterized by greater engagement) exhibited a more consistent respiration rate on average compared to those in the automated condition. This observation is supported by Wientjes [63], who suggests that rapid shallow breathing is often associated with higher mental workloads and enhanced sustained attention. Consequently, the observed lower variability in respiration rates could be indicative of heightened cognitive effort among participants in the more engaging manual condition. Moreover, [Soni, 2019] reports a positive correlation between respiration rate variability and heart rate variability, which indicates the potential relationship of this measure with mechanisms underlying task engagement.

For objective (iii), we developed a task-specific engagement index based on the physiological differences explained above. Eq. (2) demonstrated good predictive ability on the samples used to create the formula (80.95 % predictive capacity).

For objective (iv), displaying the engagement level with a discrete gradient proved less distracting and less representative than using a continuous gradient. This is likely due to the lower sensitivity of the discrete gradient, which affects the operators' sense of control over the system. Additionally, a color oscillation can occur when the measured engagement level approaches a threshold of the discrete gradient, further distracting operators. Conversely, the higher sensitivity of the continuous gradient enhanced the operators' sense of control. It also prevented oscillations between distinct colors, making this method a better alternative for displaying the engagement level.

We compared three methods of normalizing the engagement index. Two of these methods featured dynamic thresholds that were adapted based on data collected since the start of the task, while the other method employed fixed thresholds based on the maximum and minimum values from the training dataset. Results show that the three scaling methods were equally easy to interpret and were not distracting the operators. However, the static threshold method (method ii) was significantly more representative than the two dynamic methods. One possible explanation for this is that the two methods with dynamic thresholds encountered a similar issue where the thresholds diverged as the task progressed, making it increasingly challenging for operators to return to an optimal ("green") engagement level, especially at the end of the task. Therefore, we opted for utilizing the static threshold option for this iteration of the system.

In sum, the adaptive feedback system proposed in this article utilizes respiration and acceleration data to provide engagement level feedback to manufacturing operators, using a continuous color gradient calibrated using the minimum and maximum engagement values recorded in the training dataset. This system aims to assist manufacturing operators in maintaining optimal engagement levels when interacting with highly automated systems. Providing operators with real-time feedback on their engagement levels ensures they stay informed of their mental state, allowing them to prevent drops in engagement that could adversely impact their performance and, more importantly, safety. The application of this system is particularly relevant in safety-critical manufacturing environments or roles demanding high cognitive engagement, where errors could have significant financial and safety repercussions. A significant benefit of this system is its wide-ranging applicability to various tasks, regardless of their specific characteristics. Additionally, the visual display of engagement can be implemented as an exogenous signal, meaning it does not interfere with the primary task at hand. This versatility underscores the potential of adaptive feedback systems to bolster cognitive engagement during monitoring tasks.

## 7. Conclusion

This study employed a design science methodology to create an adaptive task engagement feedback system designed to help manufacturing operators stay engaged in their evolving workplace. A comparative analysis was utilized to identify the most suitable tools for measuring task engagement in a manufacturing setting, emphasizing the ease of implementation using heart rate variability and respiration rate metrics. A task-specific engagement index was developed using the physiological differences between more and less engaging manufacturing scenarios (acceleration mean, acceleration variability, and respiration variability), achieving an average engagement state prediction accuracy of 80.95 % using the leave-one-out cross-validation method in a logistic regression model. We assessed two display modalities and three scaling methods to inform our design. The final design utilized a continuous color gradient calibrated based on the lowest and highest engagement index values recorded in the training set. A subsequent study was conducted to test this advancement on a broader scale, which will be discussed in forthcoming scientific publications.

By offering real-time monitoring and optimization of engagement, this system could help minimize errors and downtime, mitigate safety risks, and promote a healthier work environment. Thus, it represents a promising approach that could improve both the operational performance and the human experience within manufacturing settings. The theoretical contributions of our work introduce the potential of using measures such as respiration variability and acceleration to infer manufacturing operators' engagement while in motion, as well as the possibility of defining an engagement metric utilizing various physiological differences between optimal and suboptimal scenarios.

It is essential to acknowledge certain limitations inherent in this system. First, our assessment of engagement relied solely on self-reported data. Ideally, employing real-time physiological monitoring tools, like EEG, would have enhanced the validation of the measured engagement levels but would have been more intrusive than the Hexoskin vest we used, potentially distracting operators. Additionally, it should be noted that while the leave-out samples were not employed in training the predictive models, they were utilized in creating Eq. (2). As a result, the model's effectiveness for new participants might not be as robust as measured in this study. It is also important to note that the formula used in this system strongly depends on the task and is specifically tailored to the context of our study. This means that Eq. (2) may not yield reliable results in different contexts and should not be applied to other scenarios without appropriate modifications and validation. Moreover, using a color gradient can make reading difficult for color-blind users, which affects approximately 8 % of the male population. Therefore, in future iterations, it would be important to integrate a color-blindness feature to adjust the displayed colors and improve contrast. Finally, the normalization methods explored in this study did not account for individual physiological differences or natural fatigue occurring during a monitoring task. Regarding individual physiological differences, our study applied a general formula across all participants without differentiation. While effective for establishing a baseline, this approach overlooks the nuances of individual responses and their impact on engagement metrics. Recognizing this limitation, we propose, in further iterations of our research, to refine our engagement threshold criteria by incorporating individual physiological differences into our analysis. This adjustment aligns with the methodology employed by Demazure *et al*. [21]. As for the fatigue consideration, in our tests, we managed to circumvent the fatigue challenge by conducting short tasks (~15 minutes) where fatigue effects could not realistically take hold. However, employing these methods would result in thresholds that fail to consider fatigue for longer tasks. Therefore, we suggest that future improvements consider the approach outlined by Demazure *et al*. [21] to incorporate fatigue considerations into establishing task engagement thresholds.

## References


[1]. M. Passalacqua, et al., Practice with less AI makes perfect: partially automated AI during training leads to better worker motivation, engagement, and skill acquisition, *International Journal of Human-Computer Interaction*, 2024, pp. 1-21.

[2]. Y. Lu, et al., Outlook on human-centric manufacturing towards Industry 5.0, *Journal of Manufacturing Systems*, Vol. 62, 2022, pp. 612-627.

[3]. A. Goujon, F. Rosin, F. Magnani, S. Lamouri, R. Pellerin, L. Joblot, Industry 5.0 use cases development framework, *International Journal of Production Research*, 2024, pp. 1-26.

[4]. F. Rosin, P. Forget, S. Lamouri, R. Pellerin, Impact of Industry 4.0 on decision-making in an operational context, *Advances in Production Engineering & Management*, Vol. 16, Issue 4, 2021.

[5]. F. Rosin, P. Forget, S. Lamouri, R. Pellerin, Enhancing the decision-making process through industry 4.0 technologies, *Sustainability*, Vol. 14, Issue 1, 2022, 461.

[6]. M. Passalacqua, G. Cabour, R. Pellerin, P.-M. Léger, P. Doyon-Poulin, Human-centered AI for Industry 5.0 (HUMAI5. 0): Design framework and case studies, in Human-centered AI, *Chapman and Hall/CRC*, 2024, pp. 260-274.

[7]. A. T. Pope, E. H. Bogart, D. S. Bartolome, Biocybernetic system evaluates indices of operator engagement in automated task, *Biological Psychology*, Vol. 40, Issues 1-2, 1995, pp. 187-195.

[8]. G. Matthews, et al., Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry, *Emotion*, Vol. 2, Issue 4, 2002, 315.

[9]. F. Dehais, A. Lafont, R. Roy, S. Fairclough, A neuroergonomics approach to mental workload, engagement and human performance, *Frontiers in Neuroscience*, Vol. 14, 2020, 268.

[10]. G. Mazzetti, E. Robledo, M. Vignoli, G. Topa, D. Guglielmi, W. B. Schaufeli, Work engagement: a meta-analysis using the job demands-resources model, *Psychological Reports*, Vol. 126, Issue 3, 2023, pp. 1069-1107.

[11]. A. B. Bakker, E. Demerouti, Towards a model of work engagement, *Career Development International*, Vol. 13, Issue 3, , 2008 pp. 209-223.

[12]. U. E. Hallberg, W. B. Schaufeli, "Same same" but different? Can work engagement be discriminated from job involvement and organizational commitment? *European Psychologist*, Vol. 11, Issue 2, 2006, pp. 119-127.

[13]. A. M. Saks, Antecedents and consequences of employee engagement, *Journal of Managerial Psychology*, Vol. 21, Issue 7, 2006, pp. 600-619.

[14]. J. Allan Cheyne, G. J. F. Solman, J. S. A. Carriere, D. Smilek, Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors, *Cognition*, Vol. 111, Issue 1, 2009, pp. 98-113.

[15]. S. Pooladvand, S. Hasanzadeh, Impacts of stress on workers' risk-taking behaviors: cognitive tunneling and impaired selective attention, *Journal of Construction Engineering and Management*, Vol. 149, Issue 8, 2023, 04023060.

[16]. D. R. Thomson, D. Besner, D. Smilek, A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms,

*Perspectives on Psychological Science*, Vol. 10, Issue 1, 2015, pp. 82-96.

[17]. J. Smallwood, J. W. Schooler, The restless mind, *Psychological Bulletin*, Vol. 132, Issue 6, 2006, pp. 946-958.

[18]. R. Parasuraman, Designing automation for human use: empirical studies and quantitative models, *Ergonomics*, Vol. 43, Issue 7, 2000, pp. 931-951.

[19]. L. Couture, M. Passalacqua, L. Joblot, F. Magnani, R. Pellerin, P.-M. Léger, Enhancing operator engagement during AI-assisted manufacturing work using optimal state deviation feedback system, in *Proceedings of the 4th IFSA Winter Conference on Automation, Robotics and Communications for Industry 4.0/5.0 (ARCI'24)*, 2024, pp. 232-237.

[20]. A. J. Karran, et al., Toward a hybrid passive BCI for the modulation of sustained attention using EEG and fNIRS, *Frontiers in Human Neuroscience*, Vol. 13, 2019.

[21]. T. Demazure et al., Enhancing sustained attention, *Business & Information Systems Engineering*, Vol. 63, Issue 6, 2021, pp. 653-668.

[22]. F. Naujoks, S. Höfling, C. Purucker, K. Zeeb, From partial and high automation to manual driving: Relationship between non-driving related tasks, drowsiness and take-over performance, *Accident Analysis & Prevention*, Vol. 121, 2018, pp. 28-42.

[23]. P. Atchley, J. Dressel, T. C. Jones, R. A. Burson, D. Marshall, Talking and driving: applications of crossmodal action reveal a special role for spatial language, *Psychological Research*, Vol. 75, 2011, pp. 525-534.

[24]. E. M. Argyle, A. Marinescu, M. L. Wilson, G. Lawson, S. Sharples, Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments, *International Journal of Human-Computer Studies*, Vol. 145, 2021, 102522.

[25]. M. F. Hinss, A. M. Brock, R. N. Roy, Cognitive effects of prolonged continuous human-machine interaction: The case for mental state-based adaptive interfaces, *Frontiers in Neuroergonomics*, Vol. 3, 2022.

[26]. M. Scerbo, Adaptive automation, in Neuroergonomics: The Brain at Work, *Oxford University Press*, 2006, pp. 239-252.

[27]. K. M. Feigh, M. C. Dorneich, C. C. Hayes, Toward a characterization of adaptive systems: a framework for researchers and system designers, *Human Factors*, Vol. 54, Issue 6, 2012, pp. 1008-1024.

[28]. M. Bernabei, F. Costantino, Adaptive automation: Status of research and future challenges, *Robotics and Computer-Integrated Manufacturing*, Vol. 88, 2024, 102724.

[29]. W. B. Schaufeli, A. B. Bakker, M. Salanova, Utrecht work engagement scale-9, Educational and Psychological Measurement, *Utrecht University*, 2003.

[30]. R. Parasuraman, R. Molloy, I. L. Singh, Performance consequences of automation-induced 'complacency', *The International Journal of Aviation Psychology*, Vol. 3, Issue 1, 1993, pp. 1-23.

[31]. N. Moray, T. Inagaki, Attention and complacency, *Theoretical Issues in Ergonomics Science*, Vol. 1, Issue 4, 2000, pp. 354-365.

[32]. M. Körber, A. Cingel, M. Zimmermann, K. Bengler, Vigilance decrement and passive fatigue caused by monotony in automated driving, *Procedia Manufacturing*, Vol. 3, 2015, pp. 2403-2409.

[33]. A. O. de Guinea, R. Titah, P.-M. Léger, Explicit and implicit antecedents of users' behavioral beliefs in information systems: A neuropsychological investigation, *Journal of Management Information Systems*, Vol. 30, Issue 4, 2014, pp. 179-210.

[34]. A. O. de Guinea, R. Titah, P.-M. Léger, Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research, *Computers in Human Behavior*, Vol. 29, Issue 3, 2013, pp. 833-844.

[35]. R. Riedl, T. Fischer, P.-M. Léger, F. D. Davis, A decade of NeuroIS research: progress, challenges, and future directions, *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, Vol. 51, Issue 3, 2020, pp. 13-54.

[36]. M. Passalacqua et al., Playing in the backstore: interface gamification increases warehousing workforce engagement, *Industrial Management & Data Systems*, Vol. 120, Issue 7, 2020, pp. 1309-1330.

[37]. A. Vasseur, M. Passalacqua, S. Sénécal, P.-M. Léger, The use of eye-tracking in information systems research: a literature review of the last decade, *AIS Transactions on Human-Computer Interaction*, Vol. 15, Issue 3, 2023, pp. 292-321.

[38]. B. T. Carter, S. G. Luke, Best practices in eye tracking research, *International Journal of Psychophysiology*, Vol. 155, 2020, pp. 49-62.

[39]. J. Gouraud, A. Delorme, B. Berberian, Out of the loop, in your bubble: mind wandering is independent from automation reliability, but influences task engagement, *Frontiers in Human Neuroscience*, Vol. 12, 2018.

[40]. J. F. Hopstaken, D. van der Linden, A. B. Bakker, M. A. J. Kompier, The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics, *Biological Psychology*, Vol. 110, 2015, pp. 100-106.

[41]. E. E. Benarroch, The locus ceruleus norepinephrine system: functional organization and potential clinical significance, *Neurology*, Vol. 73, Issue 20, 2009, pp. 1699-1704.

[42]. P.-M. Léger, F. D. Davis, T. P. Cronan, J. Perret, Neurophysiological correlates of cognitive absorption in an enactive training context, *Computers in Human Behavior*, Vol. 34, 2014, pp. 273-283.

[43]. A. Eldenfria, H. Al-Samarraie, Towards an online continuous adaptation mechanism (OCAM) for enhanced engagement: an EEG study, *International Journal of Human–Computer Interaction*, Vol. 35, Issue 20, 2019, pp. 1960-1974.

[44]. P. R. Murphy, I. H. Robertson, J. H. Balsters, G. O'Connell R, Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans, *Psychophysiology*, Vol. 48, Issue 11, Nov. 2011, pp. 1532-1543.

[45]. R. Mccraty, F. Shaffer, Heart Rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk, *Global Advances in Health and Medicine*, Vol. 4, Issue 1, 2015, pp. 46-61.

[46]. F. Shaffer, R. McCraty, C. L. Zerr, A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability, *Frontiers in Psychology*, Vol. 5, 2014, 1040.

[47]. F. Shaffer, J. P. Ginsberg, An overview of heart rate variability metrics and norms, *Frontiers in Public Health*, Vol. 5, 2017, 258.

[48]. A. L. Hansen, B. H. Johnsen, J. F. Thayer, Vagal influence on working memory and attention, *International Journal of Psychophysiology*, Vol. 48, Issue 3, 2003, pp. 263-274.

[49]. D. P. Williams, J. F. Thayer, J. Koenig, Resting cardiac vagal tone predicts intraindividual reaction time variability during an attention task in a sample of young and healthy adults, *Psychophysiology*, Vol. 53, Issue 12, 2016, pp. 1843-1851.

[50]. S. G. Hajra, P. Xi, A. Law, A comparison of ECG and EEG metrics for in-flight monitoring of helicopter pilot workload, in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC'20)*, 2020, pp. 4012-4019.

[51]. F. Dehais, et al., Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI, in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC'18)*, 7-10 October 2018, pp. 544-549.

[52]. K. J. Verdière, R. N. Roy, F. Dehais, Detecting pilot's engagement using fNIRS connectivity features in an automated vs. manual landing scenario, *Frontiers in Human Neuroscience*, Vol. 12, 2018.

[53]. I. A. Castiblanco Jimenez, J. S. Gomez Acevedo, F. Marcolin, E. Vezzetti, S. Moos, Towards an integrated framework to measure user engagement with interactive or physical products, *International Journal on Interactive Design and Manufacturing (IJIDeM)*, Vol. 17, Issue 1, 2023, pp. 45-67.

[54]. S. H. Fairclough, L. Venables, Psychophysiological predictors of task engagement and distress, in Human Factors in Design, Safety, and Management (D. de Waard, K. A. Brookhuis, R. van Egmond, Th. Boersema, Eds.), 2005, *Shaker Publishing*, pp. 349-362.

[55]. D. Vadeboncoeur, R. Pellerin, C. Danjou, Assessing the influence of human factors on overall labor effectiveness in manufacturing: a comprehensive literature review, in *Proceedings of the 4th IFSA Winter Conference on Automation, Robotics & Communications for Industry 4.0/5.0 (ARCI'2024)*, Innsbruck, Austria, February 7-9 2024, pp. 135-140.

[56]. T. Kundinger, N. Sofra, A. Riener, Assessment of the potential of wrist-worn wearable sensors for driver drowsiness detection, *Sensors*, Vol. 20, Issue 4, 2020, 1029.

[57]. D. He, Z. Wang, E. B. Khalil, B. Donmez, G. Qiao, S. Kumar, Classification of driver cognitive load: exploring the benefits of fusing eye-tracking and physiological measures, *Transportation Research Record*, Vol. 2676, Issue 10, 2022, pp. 670-681.

[58]. N. Moray, T. Inagaki, Attention and complacency, *Theoretical Issues in Ergonomics Science*, Vol. 1, Issue 4, 2000, pp. 354-365.

[59]. N. H. Cherif, N. Mezghani, N. Gaudreault, Y. Ouakrim, I. Mouzoune, P. Boulay, Physiological data validation of the hexoskin smart textile, in *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC'18)*, Vol. 1, 2018, pp. 150-156.

[60]. P.-M. Léger, F. Courtemanche, M. Fredette, S. Sénécal, A cloud-based lab management and analytics software for triangulated human-centered research, in Information Systems and Neuroscience: NeuroIS Retreat 2018, *Springer*, 2019, pp. 93-99.

[61]. R. N. Roy, A. Bovo, T. Gateau, F. Dehais, C. P. Carvalho Chanel, Operator engagement during prolonged simulated UAV operation, *IFAC-PapersOnLine*, Vol. 49, Issue 32, 2016, pp. 171-176.

[62]. C. J. Wientjes, Respiration in psychophysiology: Methods and applications, *Biological Psychology*, Vol. 34, Issues 2-3, 1992, pp. 179-203.


---