UX assessment strategy to identify potential stressful conditions for workers

Published Version:
Khamaisi R. K., Brunzini A., Grandi F., Peruzzini M., Pellicciari M. (2022). UX assessment strategy to identify potential stressful conditions for workers. ROBOTICS AND COMPUTER-INTEGRATED MANUFACTURING, 78, 1-11 [10.1016/j.rcim.2022.102403].

Availability:
This version is available at: https://hdl.handle.net/11585/949349 since: 2024-02-05

Published:
DOI: http://doi.org/10.1016/j.rcim.2022.102403

UX Assessment Strategy To Identify Potential Stressful Conditions For Workers

Riccardo Karim Khamaisi 1[0000-0002-2584-4117], Agnese Brunzini 2[0000-0003-0450-2510], Fabio Grandi 1[0000-0001-5465-0349], Margherita Peruzzini 1[0000-0003-2260-0392], and Marcello Pellicciari 3[0000-0003-2578-4123]

[1] Department of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, Modena, Italy
riccardokarim.khamaisi@unimore.it, fabio.grandi@unimore.it, margherita.peruzzini@unimore.it

[2] Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche, Ancona, Italy
a.brunzini@staff.univpm.it

[3] Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Reggio Emilia, Italy
marcello.pellicciari@unimore.it

Abstract. The European Commission defined the new concept of Industry 5.0 meaning a more human-centric, resilient, and sustainable approach for the design of industrial systems and operations. A deep understanding of the work environment and organization is important to start analysing the working conditions and the resulting User eXperience (UX) of the operators. Also, the knowledge about users' needs and ergonomics is fundamental to optimize the workers' wellbeing, working conditions, and industrial results. In this context, the paper presents a strategy to effectively assess the UX of workers to promote human-centric vision of manufacturing sites, enhancing the overall sustainability of the modern factories. A set of non-invasive wearable devices is used to monitor human activities and collect physiological parameters, as well as questionnaires to gather subjective self-assessment. This set-up was applied to virtual reality (VR) simulation, replicating heavy duty work sequence tasks that took place in an oil and gas pipes manufacturing site. This approach allowed the identification of possible stressful conditions for the operator, from physical and mental perspectives, which may compromise the performance.
This research was funded by the European Community's HORIZON 2020 programme under grant agreement No. 958303 (PENELOPE).

Keywords: User Experience; Human-Centred Design; Industry 5.0; Virtual Reality; Cognitive Ergonomics.

1 Introduction

The last decade was characterized by the advent of the fourth industrial revolution (Industry 4.0) [1], which brought a radical change in modern factories and manufacturing sites. At the base of this concept there was the introduction of a set of enabling technologies (e.g., IoT, Cloud computing, Big Data, Augmented Reality, 3D-printing) both to empower the production and to speed up the time to market. Consequently, the introduction of new advanced technologies changed drastically the role of workers, increasing their cognitive load and changing the ratio between physical and cognitive effort [2]. The greater cognitive demand increased the need for research within human factors; this discipline has been introduced in engineering to consider the physical, psychological, social, and cultural needs of human beings, during the product/system design, development, and assessment processes (ISO 9241-210, 2010) [3]. Up to now, research about Industry 4.0 focused attention mainly on machines and systems, barely considering the role of human beings in the design of the modern factory [4]. For this reason, European Commission has recently promoted a complementary new approach, called Industry 5.0, where "the wellbeing of the worker is placed at the core of the production process and uses new technologies to provide prosperity beyond jobs and growth while respecting the production limits of the planet". Differently from Industry 4.0, that is considered to be technology-driven, Industry 5.0 is value-driven [5]: the interconnected core values of Industry 5.0 concept are sustainability, resilience and human-centricity. In this vision, human workers still have a central role in controlling the production processes and are the main responsible for factory productivity and high product quality, despite the increased level of automation in modern factories [6]. In this context, as suggested by [7], the wellbeing of industry workers should be placed at the centre of manufacturing processes, by developing transparent, trustworthy, and quantifiable technologies that provide a rewarding working environment driven by real-world needs. To guarantee the right workplace and right time to the operator, job rotation schedules considering workers' qualifications, the workplace's ergonomic exposure, and the most recent allocations of each worker should be provided [8]. Above all, it becomes indispensable to understand which is the User eXperience (UX) of operators during the work shift, in order to design products and processes that help them in achieving their goals without physical and cognitive overload.

This research work aims to develop a strategy to assess the operators' UX in order to identify potential stressful conditions, focusing on their postural and physiological data. From the latter, it is possible to understand which is the physical workload, mental workload, and stress of the operators during their work tasks, helping engineers in the identification of possible design issues and in the ergonomic optimization of the manufacturing processes. The evaluation strategy of workers' user experience was applied to real industrial case concerning the manufacturing of pipes for oil and gas industry. Also due to the COVID-19 pandemic restrictions, the assessment was performed through an immersive virtual reality (VR) simulation, ad-hoc developed for this case study, replicating the most representative tasks carried out by real operators at the shopfloor. The experimental data confirms how the proposed strategy is able to evaluate the workers' UX to support engineers to optimize the manufacturing processes. The paper is structured as follows: section 2 refers to the research background and deepens the importance of adopting UX-based design methods; section 3 presents the methodological approach; section 4 describes the experimental testing on the industrial case study; section 5 contains the results and discussion; finally, section 6 shows the conclusions and future works.

2 Research Background

With the advent of automation and a simultaneous radical change in the work organization, cognitive and physical ergonomics play a decisive role in the definition of the overall productivity of modern industrial systems, throughout several industrial sectors, from process manufacturing up to energy [9]. It is widely accepted that the optimization of physical and mental workload, comfort, and perceived effort is mandatory to prevent disorders and stressful situations, while assuring the best human performances and working conditions [10]. Moreover, the advent of human-robot collaboration (HRC) systems merges the strength and accuracy of robots with the high-level cognition and flexibility of humans to increase productivity and to support and empower operators physically and intellectually [11]. The natural outcome of such considerations is the research of systems and environments optimal design, product and process quality improvement and the reduction of industrial costs.

Recent studies started proposing platforms for human data collection, elaboration, and correlation in an integrated way, supporting factory ergonomics by monitoring of operators' activities, data analysis, and implementation of corrective actions to make the workplace socially sustainable [12]. As a matter of fact, Peruzzini et al. [13] defined a set of tools, such as wearable devices, to be applied in the modern manufacturing context to enhance the workers' wellbeing, safety and satisfaction and, at the same time, the overall factory performance. The operators' monitoring could be performed also in VR environments, applying the same approach to anticipate issues at the shop floor [14].

In particular, as for the physical evaluation, common analysed indicators are the frequency of movements and their duration, the use and the type of tools adopted, if awkward postures are observed, the postural loading and effect of vibration [15]. Several methodologies have been proposed, up to now, to quantitatively assess the factors related to physical risk exposure (e.g., Rapid Upper Limb Assessment [16], Rapid Entire Body Assessment [17], Occupational Repetitive Analysis [16]). Also, wearable sensor systems and cognitive architecture to track and analyse human physical ergonomics in real time on a shopfloor has been used [18]. Anthropometric measurements of operators were considered in literature even to calculate the target functions to optimize the ergonomic job rotation and prevent work-related musculoskeletal disorders (WMSDs) [19].

On the other side, the analysis of the mental workload remains one of the most widely studied topics, concerning "the interaction between the requirements of a task, the circumstances under which it is performed, and the skills, behaviours, and perceptions of the operator" [20]: measuring the cognitive demand helps to quantify the mental "cost" of performing a task and to predict future performances [21]. As specified in [22], "cognitive assessment involves the analysis of psychological processes such as awareness, understanding, human information elaboration, reasoning, and the use of knowledge as it concerns human interaction with other system components"; this inevitably implies that each person presents in some grade a relatively limited cognitive capacity receptivity to the specific task. Moreover, user responsiveness to the same stimuli differs in relation to personal capabilities and habits. It must be underlined that the operator's cognitive engagement can positively or negatively affect human performances by impacting on the mental "cost" of performing a task [21]. The identification of errors' sources and the understanding of the task complexity perceived by workers thus remains central in the overall UX computation. Since the intricacy of each person cognitive processes, the combination of different methodologies is fundamental to have a clearer image of the perceived mental workload. So, end-users should be involved in the entire design process using a set of UX techniques, from the first research stage until the final evaluation [24]. Often self-assessment measures involve a self-subjective evaluation of the perceived workload needed to accomplish a task based on the personal experience of the interaction with the system, using questionnaires or psychometric scales, such as the multidimensional NASA Task Load Index (NASA-TLX) [25]. Physiological measures instead consider physiological signals of the human body which are thought to be correlated with mental workload [26]. Indeed, heart rate (HR), heart rate variability (HRV) [27], eye activity (like pupil diameter, blink

rate etc.) [28], brain activity (EEG) [29], breathing rate (BR), galvanic skin response or electrodermal activity (GSR or EDA) [30], can indirectly retrieve an on-time reliable picture of the mental workload [31]. Despite this wide set of parameters, they are barely used in industrial context to evaluate operators' UX due to the lack of guidance in their practical use. To overcome this issue, Peruzzini et al. [32] proposed the creation of a unique parameter in order to easily interpret the physiological parameters variations.
Nevertheless, all the above-mentioned physiological parameters are extremely affected by external factors such as the working environment, physical conditions, and psychological elements (such as the operator emotional involvement) that are not strictly related to the analysed activity. In this context, the presented strategy aims to identify the factors affecting the operator's user experience, following: i) a comprehensive analysis of the working environment and organization, ii) the identification of optimization objectives, technological set-up, and proper algorithm, and iii) the subject-specific assessment of mental and physical conditions.

3 Methodological Approach

The methodological approach presented in the study aims at creating a gold standard for the UX assessment to support the design of every kind of industrial process or product and more generally of working environment. It is based on a UX mapping strategy, as a sequence of activities able to analyse the operators' UX and propose suggestions for its optimization. Thus, it must be specified that such approach is traversal to the specific application chosen and could be adopted in several application areas, and in different environments (simulated as well as on the field). The strength of the proposed methodology relies on the following key points:

- The opportunity to sound out, item by item, the work organization and operators' perceptions to identify potential risky situations from several perspectives (environmental, physical, cognitive and social);
- The simultaneity of several analyses about the operators' UX (e.g., stress assessment, analysis of physical and cognitive workloads, user' interaction analysis) in performing daily tasks during work shifts;
- The overall consistency of the suggested procedure which aspires to concretely implement the *Human-In-the-Loop* (HITL) concept;
- The promotion of a human-centric vision, according to the Industry 5.0 paradigm, thanks to the proposed holistic UX assessment.

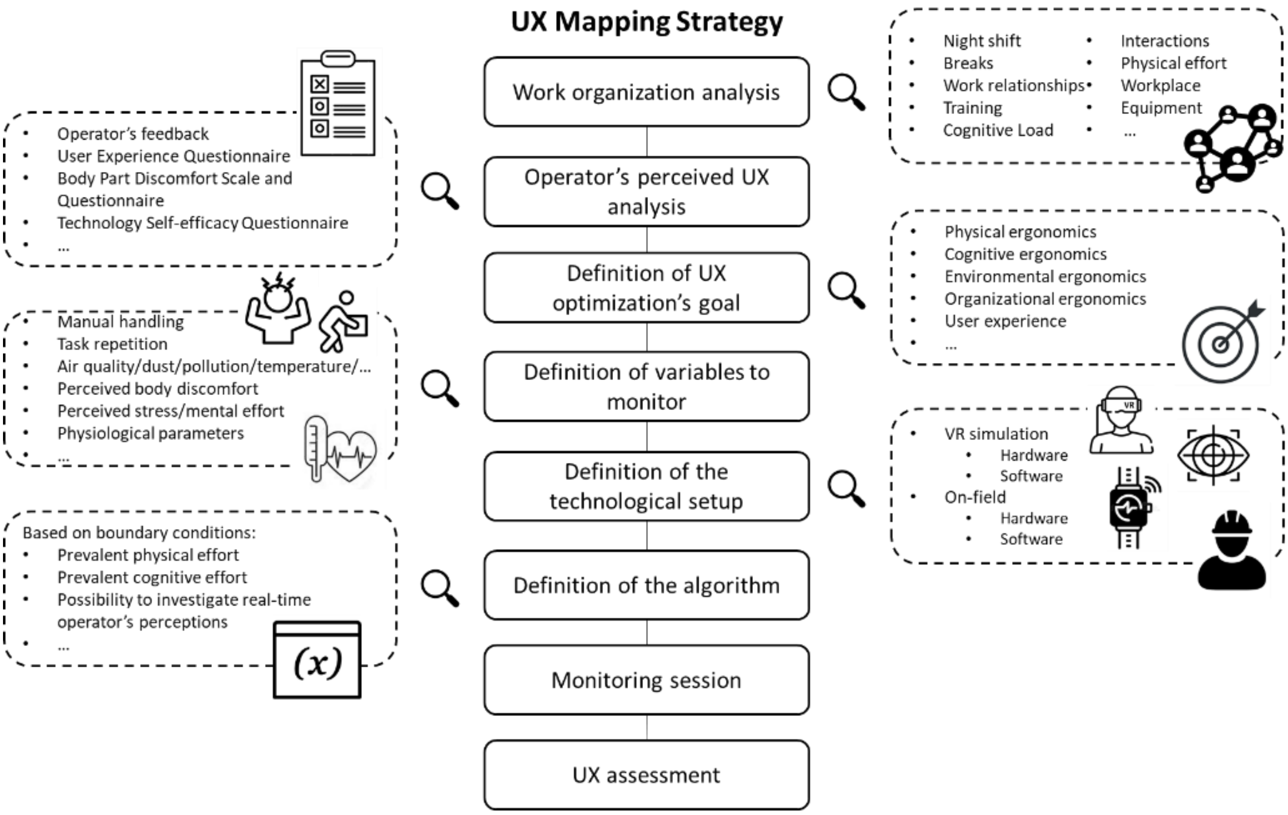

Fig. 1: UX mapping strategy workflow

Fig. *1* shows the methodology workflow as proposed. The first step (“*Work Organization Analysis”*) involves a thorough investigation concerning the targeted context and, specifically, its main features. Such an analyses could be carried out by interviews or by guided surveys to company process experts. Fig.1 described all the aspects that need to be investigated in this phase. Among them, night shifts, breaks planning, work relationships, training sessions could be considered because of their impact on the users’ UX. Also, for each workstation and each task, the following elements must be carefully analysed:

- Task duration;
- Frequently committed or predictable errors;
- Required cognitive load (e.g., knowledge-based tasks, availability of instructions or supports, decision-making tasks, time pressure.);
- Kinds of interactions (e.g., with devices or tools, with other operators, with robots or other supporting technologies);
- Required physical effort (e.g., handled loads, awkward postures, etc.)
- Workplace configuration (e.g., area and height to be covered, presence of dust, noise, pollution, etc.)
- Used equipment (e.g., gloves, glasses, noise headset).

Such an analysis is necessary to identify the features that could make each task stressful, physically risky, mentally overloading, or dangerous for the operator’s wellbeing. Subsequently, a UX research (“*Operator’s perceived UX Analysis*”) is brought forward to expand operator perceived workload both from a cognitive and physical point of view. In particular, subjective self-assessment questionnaires are provided to the operators involved in the different tasks. After task execution, they are asked to fill in specific surveys to collect demographic data and their feedback about their roles at work, the perceived workload, the user experience, the comfort, and discomfort felt in the different body parts or their familiarity with technology and innovative devices that could be adopted to support the process (such as, tools for extended reality experiences or wearable devices). The survey must be filled in immediately after the activity (e.g., for manufacturing task it could be completed at the end of the working shift too), at least by 5 users for each task, to avoid excessive personal bias.

This twofold task-based and operator-based assessment results fundamental for the “*Definition of UX optimization’s goal*”. Indeed, the final aim of the UX assessment is the identification and consequent resolution of the raised issues per task. The UX optimization aims at improving physical, cognitive, environmental, organizational ergonomics, or a blend of some of them. This inevitably influences the choice of the final testing population to subsequently monitor and on which making inferences: in general, to reinforce the resulting considerations, the involved subjects should at least reflect the analysed population, both quantitatively and qualitatively. After the objective definition, the variables to monitor have to be determined (“*Definition of variables to monitor*”). Based on the UX assessment objectives, several configurations of variables to be monitored can arise. For example, if the aim is to analyse the UX in terms of physical and cognitive workload, a set of measures, comprehending subjective and objective variables, should be established. For instance, for the physical effort analysis, biometric and biomechanical data should be combined with subjective questionnaires such as the perceived body discomfort scale. Similarly, for the mental effort or stress analysis, physiological parameters should be combined with self-assessment questionnaires such as the NASA-TLX, the State Trait Anxiety Inventory, or the Numerical Analogue Scale. Then, based on the selected variables and on the workplace configuration and constrains, the hardware and software to be used for the UX assessment can be selected (“*Definition of the technological setup”).* Indeed, several devices can be available for the monitoring of the same variable, but specific used equipment or environmental conditions or work organization constrains, can favour and force the use of certain devices rather than others. Also, based on boundary conditions (e.g., prevalent physical effort rather than the cognitive one, or the possibility to investigate real-time operator’s perceptions), the algorithm for the UX assessment must be defined (“*Definition of the algorithm*”), as proposed in [33]. For example, if the analysis of the physical and cognitive workload is simultaneous and the physical effort is foreseen to be higher than the mental one, some variables in the algorithm should have different weights.

When all these steps are defined, the “*Monitoring session”* can take place and the “*UX assessment*” can be accomplished. The monitoring session consists of the adoption of the selected set-up to collect human data during task execution, on different users, in order to apply the proposed UX mapping strategy. Subsequently, the UX assessment is based on the data post-processing and data interpretation.

4 Use Case

The use case was provided by one of the European project Penelope's partners, operating in the Oil&Gas sector. The use case focused on a set of manual operations for pressure vessels manufacturing. The adoption of the proposed methodology allowed a more efficient and human-centric processes, promoting the operators' wellbeing according to ergonomics principles. A full report regarding the *Work Organization Analysis* was initially filled in with the company supervision, specifying for each task: user requirements, duration, errors, configuration, and coordination. In particular, among all the working procedures described, four different tasks were selected regarding the more stressful operations, considering both physical and cognitive efforts:

- Fitting-up of consecutive sections of pressure cans;
- Tracing and placement of auxiliary elements on the can;
- Welding of auxiliary elements on the can;
- Grinding and polishing of the can with the operator on the ground and on the scaffolding.

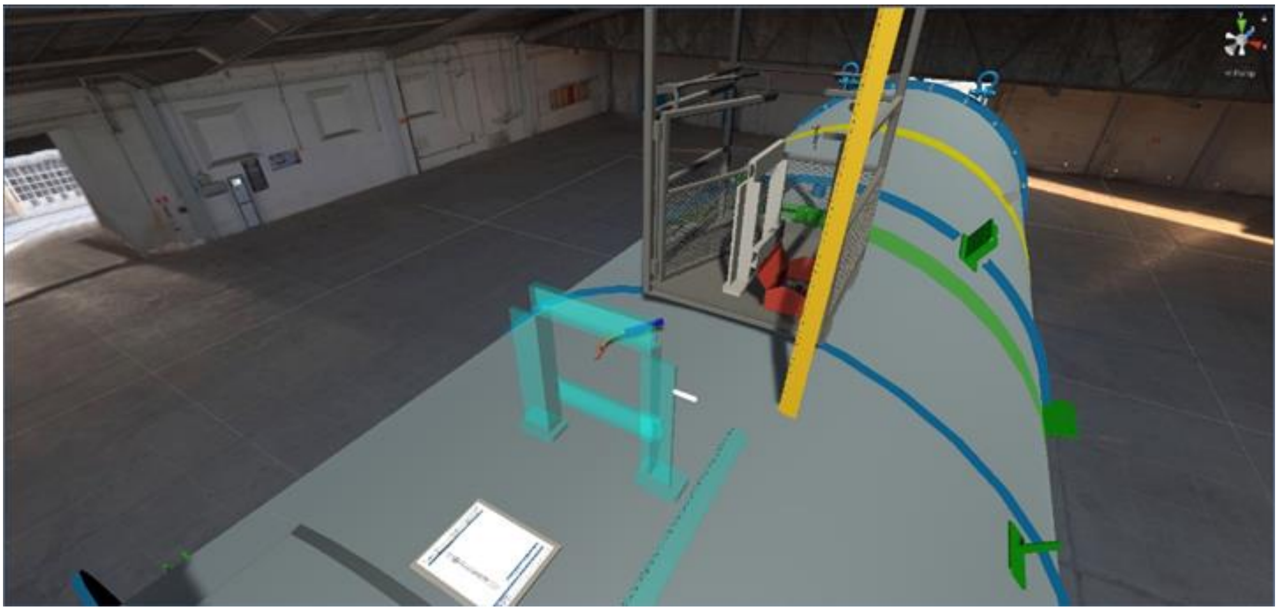

Fig. 2: VR simulation of tracing of auxiliary elements to be subsequently welded on the cans.

For each task, the on-field data collection made through paper questionnaires from the company operators (*Operator's perceived UX Analysis*) allowed to define the predominant aspect between the mental and physical effort. These surveys provided data on:

- Perceived UX;
- Perceived workload (NASA-TLX);
- Body discomfort.

This investigation led to the identification of the most relevant and problematic tasks and of the most important UX parameters to optimize according to the general working environment. The defined *UX optimization's goal* aims at defining which UX component (among stress, mental workload or postural overload) is more relevant in each task: according to the abovementioned framework, a set of related physiological parameters were thus delineated (*Definition of variables to monitor*). These include cardiac signals (heart rate, HR and inter-beat intervals, RR), electrodermal activity (EDA), pupil diameter (PD) and human joint angles (RULA). The choice of such variables was made in accordance with [33] where a similar use case regarding a manufacturing task and specifically an on-site maintenance operation was investigated.

Therefore, the most suitable non-invasive set of COTS (Commercial-Off-The-Shelf) wearable sensors, providing a sufficient level of detail and usable in the manufacturing environment, according to partner indications, has been selected (*Definition of the technological setup*). The current approach started from a set of sensors already available and tested [3] in other domains, without preventing a further technological selection on the market, according to each specific industrial requirement. The chosen setup integrate:

- HTC Vive Tracker suite, to track the human body angles;
- Empatica E4 wristband, to collect a set of physiological data such as HR and EDA signals;
- Zephyr Bioharness 3 thoracic band, to monitor the on-going cardiac activity (RR);

- HTC Vive Pro Eye headset supplied with Tobii eye tracking system, to collect relevant data (PD) regarding the cognitive and visual effort of the user during immersive virtual simulations.

The physiological parameters collected from the different devices can be analysed through the following software platforms:

- iMotions platform for the signals collected through Empatica E4 and Tobii systems;
- OmniSense Analysis for the signals collected through the Zephyr Bioharness chestband;
- XErgo software for the on-time postural load assessment according to RULA ergonomic index (Rapid Upper Limb Assessment).

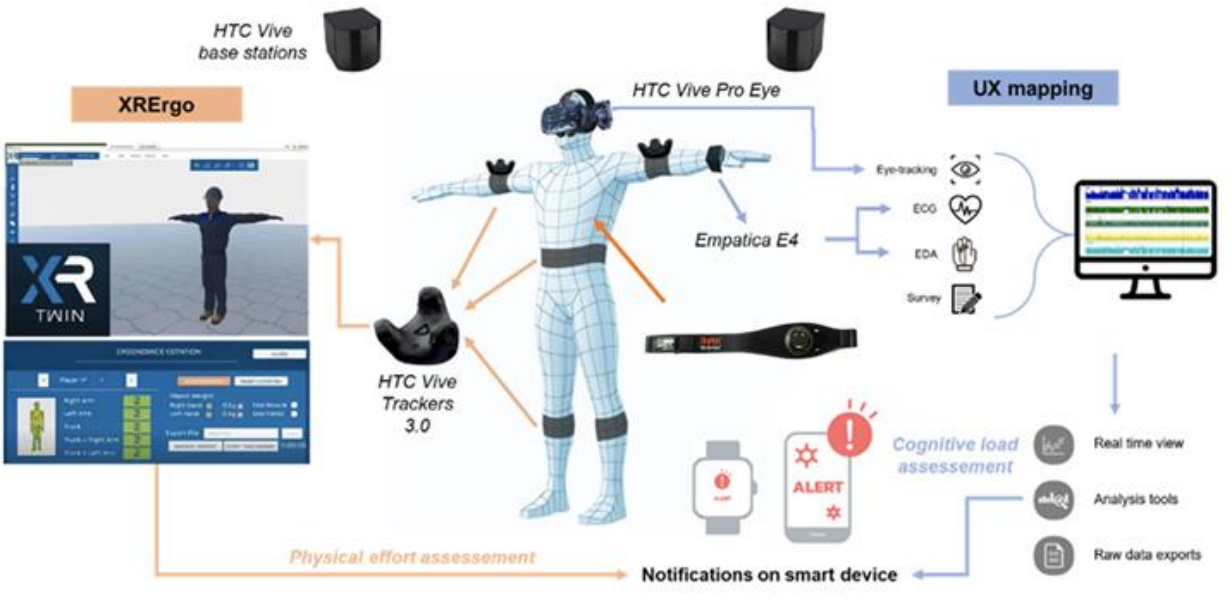

Fig. 3: The chosen technological setup for the monitoring session

A dedicated algorithm for the UX assessment has been defined for this use case for the assessment of physical workload, mental workload, and stress, as a modified version of the algorithm already presented by the same research group in [33]. Differently from the latter, the present algorithm does not consider the operator performance (i.e., time to accomplish the task) due to the lack of an expert reference time to compare the task execution. Moreover, the tasks are performed in a virtual scenario, so the execution timing may be different from those performed in a real work context: furthermore, the main intention was that of focusing on improving mental and physical engagement of the operator during the activity without any particular interest on the execution time. Regarding physiological signals, similarly to [33], Heart Activity (HA) parameter was calculated as in Eq. (1):

$$HA = \frac{HR\ mean - HR\ baseline}{HR\ max - HR\ baseline} \quad (1)$$

where *HR mean* is the mean value of the user's HR as collected during the task execution, *HR baseline* is the mean HR value as recorded during the user's baseline phase, and *HR max* is the maximum HR value reached during the task execution. In the same way, Pupil Activity (PA) and Electrodermal Activity (EA) parameters were calculated as shown in Eq. (2) and Eq. (3):

$$PA = \frac{PD\ mean - PD\ baseline}{PD\ max - PD\ baseline} \quad (2)$$

$$EA = \frac{EDA\ mean - EDA\ baseline}{EDA\ max - EDA\ baseline} \quad (3)$$

where *PD mean* is defined as the mean value of the user's PD as recorded during the task execution, *PD baseline* is the mean PD value as recorded during the user's baseline phase, and *PD max* is the maximum PD value as recorded during the task performance. With the same approach, *EDA mean*, and *EDA max* are calculated in the same manner as the previous parameters during task execution while *EDA baseline* is the mean EDA value as recorded during the user's baseline phase. These parameters are used to calculate the *Mental Workload (MW)* parameter for each task, according to Eq. (4):

$$MW = HA + PA + EA \tag{4}$$

Then, to compare the MW values for each task to the others, a percentage was calculated as in Eq. (5):

$$MW\% = \frac{MW}{\sum MW} * 100 \tag{5}$$

Differently, the *Stress (S)* parameter was calculated as in Eq. (6):

$$S = |\Delta EDA| + |\Delta RR| \tag{6}$$

were ΔEDA is the difference between the *EDA mean* value and *EDA baseline*, and similarly ΔRR is the difference between the *RR mean* value and *RR baseline.* The absolute values were calculated to solve the issues related to the different parameters' variation with stress fluctuation (when the stress increments, EDA increases and RR decreases). As for the MW parameter, a S percentage was calculated in order to easily compare the values between different tasks as shown in Eq. (7):

$$S\% = \frac{S}{\sum S} * 100 \tag{7}$$

Regarding the Physical Workload (PW), it was evaluated considering the mean value of RULA score during the entire task, as in Eq. (8). Moreover, the standard deviation was calculated in order to understand the data dispersion.

$$PW = mean\ (RULA) \tag{8}$$

Due to the recent pandemic impositions, the UX assessment campaign with operators at company premises has been postponed. Thus, a preliminary monitoring session has been conducted at UNIMORE premises involving one subject without previous experience on the task and without any knowledge on the specific industrial context. After a training period about the specific task and technologies to be used, each task was performed through an immersive virtual reality (VR) simulation, ad-hoc developed for this case study. A baseline monitoring session of 3 minutes in a relaxed sitting position has been recorded to acquire mean values of the physiological signals. The monitoring session workflow is shown in Fig.4.

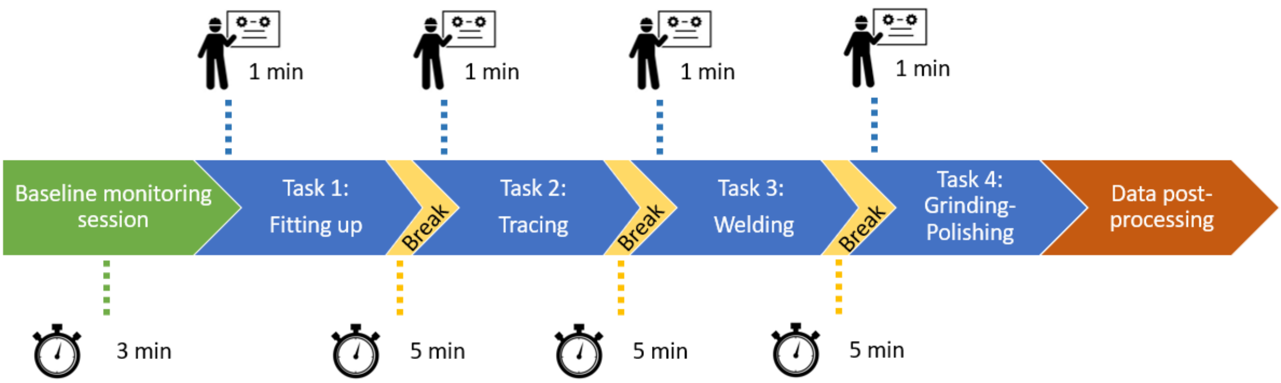

Fig. 4: Experimental monitoring session workflow.

5 Results and Discussion

Results concerning the three different analyses of the UX mapping strategy (i.e., work organization assessment, operators' perceived UX assessment, and UX assessment) are hereunder reported and discussed. It should be noted that the resulting considerations are strictly connected to the populations chosen and on the type of monitoring session prepared.

5.1 Work Organization Assessment

At the beginning of the study, tasks performed on site by the company operators have been studied to better understand the cognitive content and required mental demand, the human-machine interaction and eventual related stress, the required physical effort, the characteristics of the workplace and environment, and the used equipment. Fig. 5 shows the synthesis of the results on the four selected tasks.

		Task list			
		Fitting-up consecutive sections	Tracing & Placement of auxiliary elements	Welding of auxiliary elements	Grinding & Polishing
Duration		18h	2455h	2662h	2160h
COGNITIVE LOAD	Knowledge-based	x	x	x	
	Instruction / support		x		
	Amount of information				
	Decision making	x	x	x	x
	Real-time feedback				
	Time pressure				
INTERACTION	With devices / tools	x	x	x	x
	With technologies				
	With other operators	x	x	x	
	With robots				
	With HMI				
PHYSICAL EFFORT	Awkward postures	x	x	x	x
	Handled loads (kg)		1	20	5
	Repetitive tasks				
	Recovery				
WORKPLACE	Area to be covered	20x100	20x100	20x100	20x100
	Height to be covered	12	12	12	12
	Noise / Dust	x	x	x	x
	Lighting				
	Temperature	out	out	out	out
	Smell				
	Agent exposure	x			
EQUIPMENT	Glasses	x	x	x	x
	Helmet	x	x	x	x
	Gloves	x	x	x	x
	Noise headsets	x	x	x	x
	Special items			x	

Fig. 5: Work organization assessment

5.1.1 Cognitive Load

From the previous analysis, it emerged that all tasks involve decision-making skills, since the operator is firstly instructed on how to execute the task, and he/she autonomously decides how to perform the task during its execution. Real time feedbacks are not provided during task execution; where instructions or support material is available, they are always paper based. The perceived time pressure can be supposed equal for all the tasks; indeed, there is an estimation of time to perform each activity (i.e., task duration) that must be respected by each operator. In detail:

- **Fitting-up:** operators know and have experience on how to read and interpret plans. The most frequently committed errors concern joining cans whose numbering is not as indicated on the plan (errors rarely occur in this activity).

- **Tracing and placement:** Operators know how to read and interpret drawings and calculate deviations due to ovality. They consult the paper-based drawings before starting the activity, and during the activity they can consult them as many times as necessary or if any incident occurs. They have to translate coordinates given in drawings (elevation in mm and orientation in radial grades) on the surface of the component adapting “nominal shape” to the “real shape” of the vessel (out-of-roundness, real length, real circumference, …). The most frequently committed errors refer to bad track (for this reason it is checked three times) and in the location of auxiliary elements (bad coordinates).
- **Welding:** The operator makes all decisions, but all activities are supervised by a superior. The most frequently committed errors concern weld defects because of non-correct workmanship.
- **Grinding/polishing:** The operator makes all decisions, but all activities are supervised by a superior. The most frequently committed errors refer to a resulting bad surface.

5.1.2 Interaction

Concerning the interactions that the operators have during their shifts, the analysis highlighted that the tasks do not currently involve any interaction with specific supporting technologies (e.g., AR, scanners, wearable interfaces, etc.), robots or computers. However, some tasks require interaction with specific tools and/or with other operators, in particular:

- **Fitting-up:** Operators work in pairs and interact throughout the activity. They also interact with the turning rolls, and two assembly machines.
- **Tracing and placement:** Operators work in pairs and interact constantly. While one operator is measuring, the other one is marking. Several pairs of operators may work simultaneously on the same section. The used tools are the chalk, measuring, tape ruler, shield, and lifting platform.
- **Welding:** Operators normally work in pairs, working continuously until the assigned activity is completed. Several pairs of operators may work simultaneously on the same section. The used tool is the manual welding machine.
- **Grinding/polishing:** The operator works autonomously. The used tool are the grinding and polishing machines.

5.1.3 Physical Effort

Fig. shows that all tasks require to maintain awkward postures and/or handle loads. A recovery period is not foreseen for any task since they are not considered repetitive. It means that the physical effort is significant for all the selected tasks. Hereafter, the amin results obtained:

- **Fitting-up:** The main awkward posture is the back bending, that is kept at 70% of the activity, one week every two months. Operators rotate the activities; thus, it takes 2 months for an operator to repeat the activity.
- **Tracing and placement:** The main awkward postures are the back bending, crouch, on the knees, and outstretched arms. This activity is carried out according to the workload, therefore 15-20 days in a month can be taken as a reference. The handled load is about 1kg and refers to tools such as square and bevel, measuring ruler, chalk.
- **Welding:** The used tool is the manual welding machine. This activity is performed at height; therefore, the operator carries the machine and sits in a bad posture during the whole activity. The main awkward postures are the back bending, crouch, on the knees, and outstretched arms. The handled loads are about 20 kg.
- **Grinding/polishing:** Again, the main awkward postures are the back bending, crouch, on the knees, and outstretched arms. Throughout the activity the operator is on his/her knees, bent over and hunched back if the surface is on the ground; otherwise, if the area is at a height the operator must stretch up to the activity area. The handled loads are about 5 kg.

5.1.4 Workplace

The workplace is common for all the selected tasks. The analysis indicated that, at the shopfloor, the area and the height to be covered are respectively of 20x100 m and 12 m for every task. Similarly, during the tasks execution the noise level is around 85-135 dB, the temperature is variable with the external one, the light is artificial, the presence of dust is high, while the presence of smell is absent. Also, operators are exposed to metallic particles in every task, and to welding fumes during the welding activity.

5.1.5 Equipment

In all the tasks the operators must wear protective glasses, helmet, gloves, and noise headset. In the welding task they have also to wear welding masks and welder caps.

5.2 Operator’s Perceived UX Assessment

Results from the initial campaign with the company operators related to self-assessment questionnaires are hereunder presented. In particular, UX surveys have been administered to five operators for each task, on-field, immediately after the end of the shift, for a total number of twenty participants. After the explanation of the aim of the study, all participants signed the informed consent. Surveys allowed to collect data and impressions about the perceived UX, perceived workload (NASA-TLX), and body discomfort.

The perceived UX survey concerns the subjective assessment of the operators' opinions and feelings about their role and work activity. In detail, they have been asked to evaluate on a 5-points Likert scale:

- The amount of responsibility their role has;
- How much the activity is physical demanding;
- How much the activity is mentally stressful;
- How much attention the activity requires;
- If the activity involves spare time or interruptions.

The NASA-TLX provided information on how humans subjectively evaluate various aspects of workload for accomplishing a task. Indeed, it is a multidimensional assessment questionnaire that rates perceived workload under six different dimensions: mental, physical, and temporal demands, performance, effort, and frustration levels. This questionnaire was included in the proposed UX mapping strategy because it allows, at the same time, assessing both mental and physical perceived effort to perform the activity, together with emotional states related to stress such as perceived effort and frustration.

The body part discomfort scale is a subjective symptom survey tool that evaluates the respondent's direct experience of discomfort at different body parts [34]. The operator has to assess his/her level of discomfort for the body areas shown in a picture, on a 5-points scale (1 = not uncomfortable, 2 = barely uncomfortable, 3 = quite uncomfortable, 4 = very uncomfortable, 5 = extremely uncomfortable). This scale allows understanding the most physically stressed body parts for each task.

Table 1 shows the results about the perceived UX for the selected tasks (reporting average values over the five participants for task). Operators performing welding of auxiliary elements believed that their role has a great amount of responsibility, followed by the ones performing the grinding or polishing. However, for all the tasks the mean score is more than 4 (on the 5-points Likert scale). Also, the mean score is high for the perceived physical effort required; the less physically demanding task seems to be the tracing and placement of auxiliary elements. On the other hand, the perceived mental stress is low for all the tasks (mean values between 1.4 and 2), even if the operators do not think that the tasks are almost automatic requiring little or no attention. Lastly, operators in charge with the fitting-up of consecutive vessel's sections believe that they almost never have spare time and interruptions or overlap among activities are very frequent. This assumption reached lower mean scores for the other three tasks.

Table 1: Results (mean values) of the perceived UX survey

	Fitting-up	Tracing and Placement	Welding	Grinding/ Polishing
I believe my role has a great amount of responsibility	4,20	4,00	5,00	4,33
I believe this task is very physical-demanding	4,60	3,40	4,20	4,67
I feel mentally stressed while performing this task	2,00	1,40	1,40	2,00
This task is almost automatic requiring little or no attention	1,20	1,60	1,80	1,67
I almost never have spare time and interruptions or overlap among activities are very frequent	4,00	2,40	3,40	3,00

Concerning the perceived workload, average results (computed over the five participants for each task) of NASA-TLX are similar among the different tasks, with mean values of NASA-TLX total score equal to:

- 83,13 ± 10,46 for fitting-up consecutive sections
- 71,33 ± 8,29 for tracing and placing auxiliary elements
- 84,33 ± 10,45 for welding auxiliary elements
- 75,22 ± 9,97 for grinding/polishing

Therefore, the perceived workload results high for tracing and placing the auxiliary elements and for grinding and polishing, while it results very high for the fitting-up of consecutive sections and for the welding of auxiliary elements (score interpretation by [35]). However, several differences can be found in the six sub-scales (Fig. 6). The fitting-up of consecutive sections and the tracing and placement of the auxiliary elements are the most mentally demanding tasks. However, while the fitting-up is also the most physically demanding task, on the opposite, the tracing and placement is the one with the lowest perception of physical effort. The temporal demand is almost equal for all the tasks, except for the tracing and placement (where the perception of the time pressure is lower). The performance is very important for the operators involved in the tracing, placement, and welding of the auxiliary elements, where accurate results must be assured. However, the average perceived effort is quite high for all the operators in all the tasks. Finally, the feeling of frustration is absent or low for all the tasks, with a small increment in the grinding or polishing.

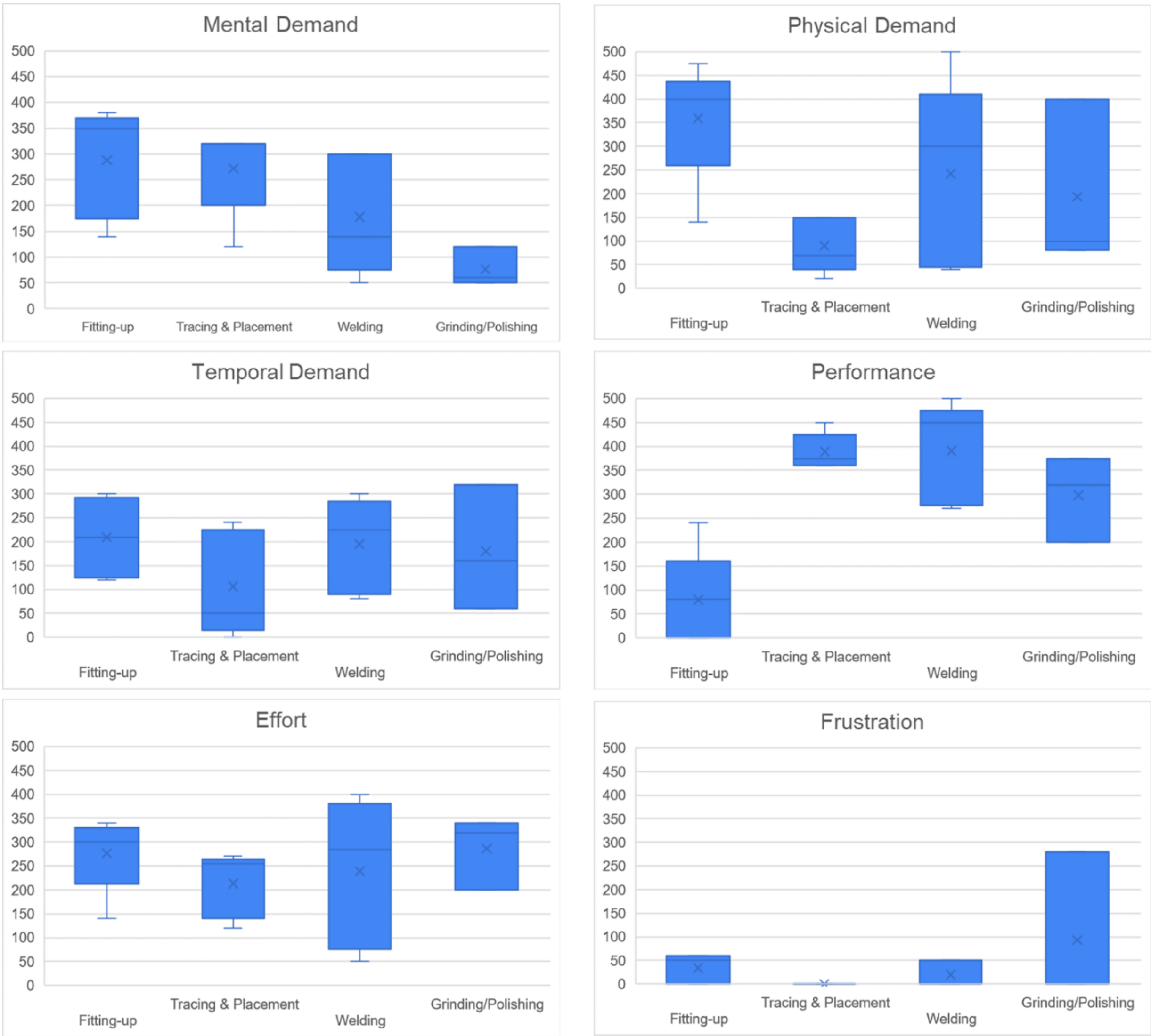

Fig. 6: Boxplots of the six weighted NASA-TLX subscales

Concerning the physical discomfort assessment, average results over the five participants per each task are shown in Table 2 Attention must be paid on grinding and polishing tasks because they resulted as the ones causing the greatest discomfort in all the body parts. Indeed, operators felt very or extremely uncomfortable on arms, head, neck, shoulders, middle and low back, knees, legs, and feet. Also, they felt quite uncomfortable on thigh and buttock. Also, the fitting-up of consecutive sections is quite or very uncomfortable, especially for the upper body parts and knees. The tracing and placement of auxiliary elements is the most comfortable task. Indeed, it is also the less physically demanding, as confirmed by the NASA-TLX results.

Table 2: Results (mean values) of the body parts discomfort scale (perceived physical discomfort)

	Fitting-up	Tracing and Placement	Welding	Grinding/Polishing
Head and neck	3,60	1,60	3,00	4,33
Shoulder	3,60	1,20	3,40	4,33
Arm	4,20	1,60	3,00	4,33
Middle back	3,40	2,40	3,00	4,67
Low back	3,80	3,00	3,40	4,33
Buttock	1,40	1,80	2,40	3,00

Thigh	1,60	1,60	2,00	3,67
Knee	3,80	2,60	3,80	4,00
Leg and foot	2,00	1,60	3,00	4,00

5.3 UX Assessment

The final step of the proposed methodology consists of the interpretation of the physiological signals collected during the monitoring sessions carried out in VR simulations. The results offer important considerations compared to the initial questionnaire outcomes: statistical evidence will be further necessary to allow for a systematic adoption. Considerations on the following results aim at offering interesting insight on the cause of discomfort in the chosen tasks, eventually defining which peculiar component of the overall UX needs revision to improve operator well-being and to optimize the working package.
The current section has been divided according to the different variables provided in the algorithm to exhaustively analyse data referred to specific biometrics. No further numerical corrections with adapted weights have been performed to each of variables involved in the equations presented: to avoid useless computations, all the collected data are rounded at the third decimal place.

5.3.1 Physical workload – RULA score interpretation

Table 3 sums up the mean values of the RULA score for each task and the relative standard deviation for data dispersion's indications. RULA score was computed from the data recorded by the XRErgo software (with a sampling frequency of 18 Hz) which differentiates between the score for the left and right part of the human body: then, the maximum value among the two was considered and the mean obtained from all the recordings for the specific task. For grinding and polishing task, two scores were calculated referring to the posture of operator, on the ground and on the scaffolding. Being the postural data extremely dependent on the position of the Vive Trackers on the subject as well as on the software calibration, the same setup was kept throughout all the monitoring sessions: for future statistical campaign, precise positions for the trackers will be identified according to the software guidelines and tracking issues, in order to overcome possible anatomical differences. Alternatively, a reliable reconstruction of the human manikin through real measurements of the human segments' length will be provided to the software thanks to ad-hoc .xml files for each tester. Furthermore, laboratories conditions in terms of noise frequencies and illumination will be reproduced accordingly to maximize tracking stability.

Table 3: RULA mean values and standard deviation for each task.

Task	Fitting up	Tracing & Placement	Welding	Grinding & Polishing	
Mean	3.576	4.044	4.305	3.252	2.884
Standard Deviation	0.744	0.684	0.688	0.570	0.401

As emerged from Table 3 and Fig. 5, welding and fitting up tasks represent highly postural demanding tasks: as for data regarding the tracing operations, the critical value reported could be due to the inexperience of the operator performing the virtual task compared to the real ones. It must be remarked that all these values do not consider real working issues, like weights due to working tools and equipment or awkward positions caused by inaccessible or dangerous spaces: this limitation could be overcome by mitigating the effect of specific physical parameters by imposing adequate weights while computing the aforementioned scores.
By examining the standard deviation data, it could be inferred that no relevant differences in term of tracking stability and thus RULA numerical dispersion have been encountered during the monitoring sessions, allowing for a reliable interpretation of the collected data. In case of outlier data due to interferences or tracking instabilities, a punctual statistical analysis could be performed to identify critical values (e.g., Dixon test's, Grubbs test or similar) or to test the hypotheses of a normal distribution (i.e., Shapiro Wilk test).

5.3.2 Mental workload and stress

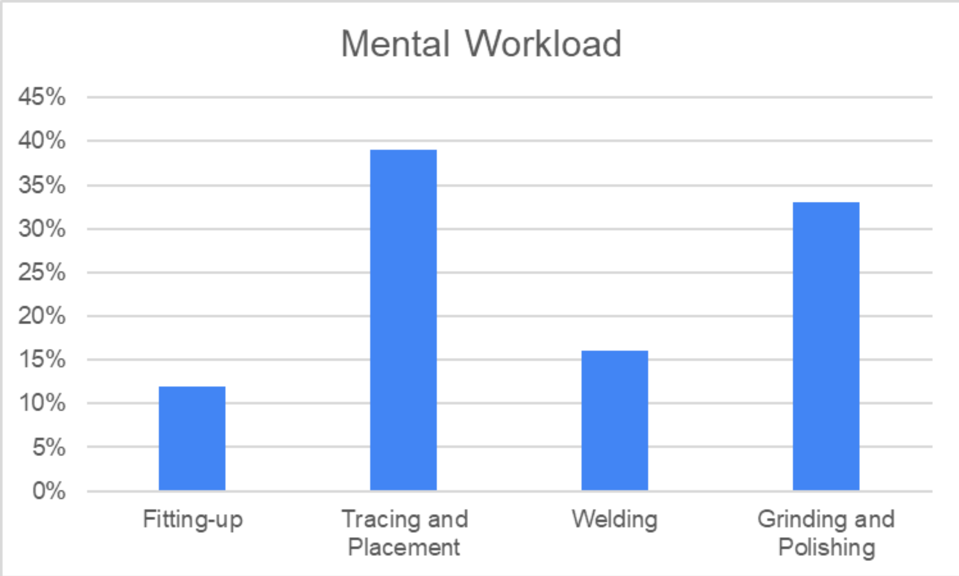

Fig. 7: Bar chart summarizing the percentage of each task's Mental Workload (MW) perceived by the operator in relation to the overall mental effort considering all the simulations.

Fig. 7 and Fig. 8 respectively summarize the percentage of Mental Workload and Stress perceived throughout the tests with reference to the algorithm chosen. As for the MW, the absolute value of the difference between the mean value for the PD, HR and EDA recorded during the baseline session and the one recorded during the specific task is computed for each simulation and summed with all the other to retrieve a normalized value and consequently the corresponding percentage: similarly, the Stress impact on the task was determined by analogous operations on RR and again EDA. However, Empatica E4 was considered unsuited to collect RR parameter being a peripheric and unstable acquisition of the cardiac activity, especially in extremely dynamic tasks.

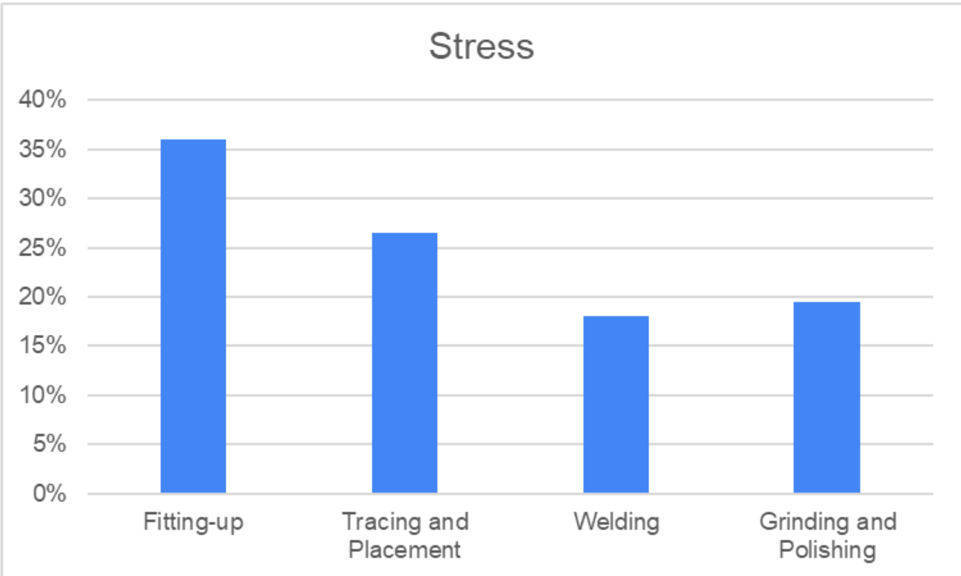

Fig. 8*:* Bar chart summarizing the percentage of each task's Stress condition (S) perceived by the operator in relation to the overall stress observed throughout all the tasks.

Comparing the physiological results with the starting questionnaire, no consistent correspondence was found: the MW recorded during the fitting up task does not seem to match the real operator assessment, probably due to the impossibility of reproducing the effort and pressure perceived while completing the task in harsh working environment. Task's responsibility could explain the discrepancy experienced too.

Concerning the Stress evaluation, from the interpretation of Fig. 8 it could be stated that no great variation among the four tasks is observed (the maximum difference is less than 20%): the fitting up task remains the most stressful followed by the tracing and the grinding ones.

6 Conclusion

The current research aimed at proposing a cost-effective and reliable strategy to evaluate the UX of operators' considering the diverse aspects impacting on the performance as well as on the operators' wellbeing. Despite a full statistical validation is foreseen in the period to come with a proper campaign on subjects belonging to different human percentiles and with disparate working experience, the study remarked that the proposed algorithm seems to be reliable and consistent in identifying specific criticalities during the execution of a task. This meaningfulness is strictly connected to the technological choice operated during the research: more tests should be performed with different device to test algorithm efficacy. Further investigations are needed to deepen the correlation existing between the Mental Workload, Physical Workload and Stress parameter and the physiological signals, according to the current literature research: specifically, the most incisive statistical variable with regard to the specific biometric and the relative interpretation must be examined in depth with greater samples. Moreover, algorithm improvements are foreseen to detach the analysis of the single task from a baseline reference, in order to compute scores which could be self-explicatory and independently meaningful, overcoming possible experimental discrepancies between the specific monitoring session and the baseline recording as well as subjective physiological state which could lead to incorrect interpretations.

As regards the postural workload, several ergonomics scores (REBA, OWAS etc.) are going to be tested and integrated within the presented methodology in order to understand and emphasize the peculiarities pertaining to each score and to prove the flexibility of the strategy: in fact, the aforementioned workflow should be modulated according to the desired ergonomic analysis and the distinct objectives.

Nevertheless, the presented approach includes a minimal set of physiological parameters for a meaningful and correct interpretation of the scores: further implementations will try to integrate the subjective analysis within the equations and expand the number of variables involved in the definition of each component.

Acknowledgments

This research was funded by the European Community under the HORIZON 2020 programme, grant agreement No. 958303 (PENELOPE).

References

1. Kagermann H, Wahlster W, Johannes H (2013). Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final Report of the Industrie 4.0, WG, (April), 82. https://doi.org/10.13140/RG.2.1.1205.8966.
2. Kong F (2019). Development of metric method and framework model of integrated complexity evaluations of production process for ergonomics workstations. International Journal of Production Research, 57(8): 2429–2445. https://doi.org/10.1080/00207543.2018.1519266
3. ISO 9241-210 Ergonomics of human system interaction - Part 210: Human-centered design for interactive systems, 2010.
4. Parrott A, Warshaw L (2017). Industry 4.0 and the digital twin – Manufacturing meets itsmatch. Deloitte University Press https://www2.deloitte.com/insights/us/en/focus/industry-4-0/digital-twin-technologysmart-factory.html.
5. Xu X, Lu Y, Vogel-Heuser B, Wang L (2021). Industry 4.0 and Industry 5.0—Inception, conception and perception. Journal of Manufacturing Systems, 61, 530-535.
6. Neumann WP, Kolus A, Wells RW (2016). Human factors in production system design and quality performance – A systematic review. In: Proc. 8th IFAC conference on manufacturing modelling, management and control MIM 2016, Troyes, France, 28–30June 2016: Vol. 49 (12): 1721–1724.
7. Lu Y, Zheng H, Chand S, Xia W, Liu Z, Xu X, Wang L, Qin Z, Bao J (2022). Outlook on human-centric manufacturing towards Industry 5.0. Journal of Manufacturing Systems, 62, 612-627, https://doi.org/10.1016/j.jmsy.2022.02.001.
8. Hochdörffer J, Hedler M, Lanza G (2018). Staff scheduling in job rotation environments considering ergonomic aspects and preservation of qualifications. Journal of Manufacturing Systems, 46, 103-114, https://doi.org/10.1016/j.jmsy.2017.11.005.
9. Dadashi N, Lawson G, Marshall M, Stokes G (2021). Cognitive and metabolic workload assessment techniques: A review in automotive manufacturing context. Hum. Factors Ergon. Manuf. 32: 20–34.
10. Pheasant S (1999). Body Space: Anthropometry, Ergonomics, and the Design of Work; Taylor & Francis: Abingdon, UK, 121–123.
11. Wang L. (2022). A futuristic perspective on human-centric assembly. Journal of Manufacturing Systems, 62, 199-201.
12. Ciccarelli M, Papetti A, Germani M, Leone A, Rescio G (2022). Human work sustainability tool. Journal of Manufacturing Systems, 62, 76-86, https://doi.org/10.1016/j.jmsy.2021.11.011

13. Peruzzini, M., Grandi, F., & Pellicciari, M. (2017). Benchmarking of tools for user experience analysis in industry 4.0. Procedia manufacturing, 11, 806-813.
14. Peruzzini, M., Grandi, F., & Pellicciari, M. (2018). How to analyse the workers' experience in integrated product-process design. Journal of Industrial Information Integration, 12, 31-46.
15. Joshi M, Deshpande V (2019) A systematic review of comparative studies on ergonomic assessment techniques. Int. J. Ind. Ergon. 74: 102865.
16. McAtamney L., Nigel Corlett E (1993). RULA: A Survey Method for the Investigation of Work-Related Upper Limb Disorders. Applied Ergonomics, 24 (2), 91–99. doi:10.1016/0003-6870(93)90080-S
17. McAtamney L., Hignett S (2004). Rapid Entire Body Assessment. Handbook of Human Factors and Ergonomics Methods, 31: 8@1@811. doi:10.1201/9780203489925.ch8
18. Colombini D, Occhipinti E. (2016) Risk analysis and management of repetitive actions: a guide for applying the OCRA system (occupational repetitive actions). CRC Press.
19. Oyekan J, Chen Y, Turner C, Tiwari A (2021). Applying a fusion of wearable sensors and a cognitive inspired architecture to real-time ergonomics analysis of manual assembly tasks. Journal of Manufacturing Systems, 61, 391-405, https://doi.org/10.1016/j.jmsy.2021.09.015
20. Huang SH, Pan YC (2014). Ergonomic job rotation strategy based on an automated RGB-D anthropometric measuring system, Journal of Manufacturing Systems, 33, 699-710. https://doi.org/10.1016/j.jmsy.2014.02.005
21. Hart SG (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 50: 904–908.
22. Cain B (2007). A Review of the Mental Workload Literature; RTO-TR-HFM-121-Part-II; Defence Research and Development Canada: Toronto, Canada.
23. Norman D. A. (1987). Cognitive engineering—cognitive science. In Interfacing thought: Cognitive aspects of human-computer interaction, 325-336.
24. Prati E, Peruzzini M, Pellicciari M, Raffaeli R (2021). How to include User eXperience in the design of Human-Robot Interaction, Robotics and Computer-Integrated Manufacturing, 68, 102072.
25. Hart SG, Staveland LE (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Advances in Psychology; Hancock, P.A., Meshkati, N., Eds.; Human Mental Workload: North-Holland, The Netherlands, 52:139–183.
26. Thorvald P, Lindblom J, Andreasson R (2019). On the development of a method for cognitive load assessment in manufacturing. Robotics and Computer-Integrated Manufacturing, 59, 252-266.
27. Akselrod S, Gordon D, Ubel FA, Shannon DC, Berger AC, Cohen RJ (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. Science, 213(4504), 220-222.
28. Ahlstrom U, Friedman-Berg FJ (2006). Using eye movement activity as a correlate of cognitive workload. International journal of industrial ergonomics, 36(7), 623-636.
29. Buerkle A, Eaton W, Lohse N, Bamber T, Ferreira P (2021). EEG based arm movement intention recognition towards enhanced safety in symbiotic Human-Robot Collaboration. Robotics and Computer-Integrated Manufacturing, 70, 102137.
30. Mehler B, Reimer B, Coughlin JF, Dusek JA (2009). Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. Transportation research record, 2138(1), 6-12.
31. Charles RL, Nixon J (2019). Measuring mental workload using physiological measures: A systematic review. Appl. Ergon. 74: 221–232.
32. Grandi, F., Peruzzini, M., Cavallaro, S., Prati, E., & Pellicciari, M. (2022). Creation of a UX index to design human tasks and workstations. International Journal of Computer Integrated Manufacturing, 35(1), 4-20.
33. Brunzini A, Peruzzini M, Grandi F, Khamaisi RK, Pellicciari M (2021). A Preliminary Experimental Study on the Workers' Workload Assessment to Design Industrial Products and Processes. Applied Sciences 11(24):12066. DOI: 10.3390/app112412066.
34. Corlett, E. N., Bishop, R. P. (1976). A technique for measuring postural discomfort. Ergonomics, 9, 175-182.
35. Sugarindra, M., Suryoputro, M. R., & Permana, A. I. (2017). Mental workload measurement in operator control room using NASA-TLX. In IOP Conference Series: Materials Science and Engineering (Vol. 277, No. 1, p. 012022). IOP Publishing.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: