



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

An integrated methodology for the assessment of stress and mental workload applied on virtual training

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Brunzini, A., Grandi, F., Peruzzini, M., Pellicciari, M. (2024). An integrated methodology for the assessment of stress and mental workload applied on virtual training. *INTERNATIONAL JOURNAL OF COMPUTER INTEGRATED MANUFACTURING*, 37(9), 1088-1106 [10.1080/0951192X.2023.2189311].

Availability:

This version is available at: <https://hdl.handle.net/11585/949299> since: 2024-02-06

Published:

DOI: <http://doi.org/10.1080/0951192X.2023.2189311>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Agnese Brunzini, Fabio Grandi, Margherita Peruzzini & Marcello Pellicciari (2023) An integrated methodology for the assessment of stress and mental workload applied on virtual training, International Journal of Computer Integrated Manufacturing.

The final published version is available online at:

<https://doi.org/10.1080/0951192X.2023.2189311>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

An integrated methodology for the assessment of stress and mental workload applied on virtual training

Agnese Brunzini^{a*}[0000-0003-0450-2510], Fabio Grandi^b[0000-0001-5465-0349],
Margherita Peruzzini^b[0000-0003-2260-0392], Marcello Pellicciari^c[0000-0003-2578-4123]

^a Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche, Ancona, Italy;

^b Department of Engineering “Enzo Ferrari”, University of Modena and Reggio Emilia, Modena, Italy;

^c Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Reggio Emilia, Italy.

* Corresponding author. E-mail address: a.brunzini@staff.univpm.it

Word count: 8250

An integrated methodology for the assessment of stress and mental workload applied on virtual training

The importance of training for operators in industrial contexts is widely highlighted in literature. Virtual Reality (VR) is considered an efficient solution for training, since it provides immersive, realistic, and interactive simulations environments promoting a learn-by-doing approach, far from the risks of the real field. Its efficacy has been demonstrated by several studies, but a proper assessment of the operator's cognitive response in terms of stress and cognitive load during the use of such technology is still lacking. This paper proposes an integrated methodology for the analysis of user's cognitive states, suitable for each kind of training in the industrial sector and beyond, fostering the human-centred design and manufacturing perspective. The methodology has been assessed using an industrial case study where virtual training is used for assembly of agricultural vehicles. Experimental results highlighted that, with VR additional supportive information, while operators' errors drastically decrease, the stress increases for complex tasks, due to the greater amount of information to manage. The proposed protocol allows understanding the operators' cognitive conditions in order to optimize the VR training application, avoiding operators' stress, mental overload, and improving performance.

Keywords: virtual reality; virtual assembly; virtual training; cognitive ergonomics; mental workload; stress.

1. Introduction

In recent years, industries are shifting towards Industry 4.0, where new intelligent machines, robots, and tools are added to the workforce. In this smarter environment, the role of the human operator (i.e., Operator 4.0 (Peruzzini et al. 2020)) remains fundamental. Thus, the need for adequate training of the operator arises, to guarantee the optimal integration between humans and the innovative advanced technological systems (Peruzzini et al. 2020).

In the context of Industry 4.0, digitalization is considered one of the most important drivers of innovation, useful not only to save time and cost, but also to optimize data and process management. In particular, digital manufacturing can be applied to different stages of the manufacturing process, such as design, prototyping, and assembly training (Abidi et al. 2019). Indeed, due to the importance of the assembly step in the manufacturing process, specific training should be provided to the operators, also to cope with the new technologies. In this context, the Operator 4.0 can be supported with different levels of cognitive automation, namely technical solutions helping the operator about how and what to assemble and to control the situation. Virtual Reality (VR) is categorized among these technological supports (Mattsson et al. 2020).

VR offers the opportunity of "learning-by-doing" instead of learning by observing or listening (Abidi et al. 2019). Moreover, VR allows to digitally simulate not only the industrial processes (from the product/system design to the prototyping, assembly, ergonomic analysis, and maintenance) but also the human-machine interaction in a risk-

free digital environment. Therefore, VR technology is considered an efficient solution for assembly training, since it provides immersive, realistic, and interactive simulations for helping and training the operator in the smart factory in the execution of complex tasks, far from the risks of the real operational environment (Abidi et al. 2019; Romero et al. 2016; Zolotová et al. 2020).

Indeed, smart interaction between Operator 4.0 and the advanced intelligent machines involves both the physical and cognitive dimensions. Cognitive interactions rely on the worker's cognitive skills and capabilities. VR technology can be used to supply the user with real-time relevant data that may reduce the dependency on the operator memory and decrease human errors. Moreover, wearable devices can be used to monitor workers' conditions under stressful or difficult situations, and proper warnings should be provided when needed (Zolotová et al. 2020).

In this context, the development and use of VR training applications can help the creation of the proper skills in a short time, also in delocalized sites. However, virtual training needs to be strongly human-centered in order to be effective and to fully exploit its great potential. As a consequence, human factors (HF) assessment assumes a critical importance in understanding whether and how the virtual training procedure is supporting the operators to leverage their skills effectively. It has been demonstrated that physical and cognitive ergonomics strongly impact manufacturing performance, and, consequently, factory productivity. For this reason, companies should necessarily deal not only with performance objectives (as cost, quality, speed, productivity, flexibility, adaptability) but also with human sustainability, in terms of health and safety, to enhance the operator's wellbeing and improve his/her skills (Peruzzini et al. 2020; Papetti et al. 2021). In addition, psychophysiological stress should be prevented and avoided, safeguarding mental wellbeing. Prolonged sensations of outrageous mental effort and stress may result in the user's burnout, lower performance, and reduced productivity (Etzi et al. 2019). For these reasons, systems should be designed based on the operators' cognitive and physical needs, to improve the quality of human-machine interaction and, finally, the workers' performances.

Several studies have tried to assess the effectiveness of virtual assembly training. Its utility and feasibility have been proven (Abidi et al. 2019; Etzi et al. 2019), as the efficacy of giving visible hints (Mattsson et al. 2020). However, a proper assessment of the operator's cognitive response in terms of stress and mental workload, during the use of such technology, is still lacking in literature (Etzi et al. 2019).

This paper wants to make a relevant contribution in the human-centred manufacturing field, by proposing a comprehensive methodology for the analysis of the users' cognitive states during virtual training sessions, trying to close the research gap. Indeed, new technologies (such as VR) are becoming key enablers for the digital transformation of industries, and their impact on operators should be carefully analysed to avoid adverse implications, especially in terms of stress and cognitive overload. The proposed method is suitable for each kind of training in the industrial sector and beyond, both traditional and virtual. An experimental testing referring to a case study about virtual assembly in the agricultural vehicle sector has been developed to validate the proposed method. A double VR application has been developed: the first one with basic aids to perform the tasks, the latter with additional support for the operator. The aim was to study

the effect of VR training (with different digital contents) on users' performance, mental workload, and stress, through a simplified algorithm that involves the computation of key performance indicators (KPIs), self-assessment surveys, and physiological parameters, differently than most literature contributions that use only one of the previous methods, and do not distinguish between stress and mental workload.

2. Research Background

Industry 4.0 introduced a set of new emerging technologies but there is a need to continuously qualify the human worker about new and changing technology trends since the human is the most flexible entity in the production system (Gorecky et al. 2015). In literature, several works about the effectiveness of VR training in the manufacturing context can be found. In fact, VR is a very helpful and valuable work tool for the simulation of manufacturing systems, and it can be used in both industrial and academic fields providing a low-cost, secure and fast analysis tool (Rubio et al., 2005). Authors used different qualitative and quantitative evaluation criteria, which can be summarized in cognitive skills, levels of trust/acceptance of extended reality tools, motivation in use, participants' attitude, previous experience, cybersickness, physiological reactions, level of presence and engagement, and technical aspects (Borsci et al. 2015). Usually, one or a few of these criteria are chosen for the VR training effectiveness assessment (Abich et al. 2021). Also, the analysis usually considers the demographic information of participants and performance-related variables such as time of performance for each task of the procedure, number of unsolved and recovered errors, time for error recovery, number of tasks without errors, etc. (Borsci et al. 2015). Performance measures (in terms of task completion time and error rate), associated with subjective measurement about system usability (through questionnaires), are used in most of the works related to assembly VR training (Otto et al. 2019; Hoedt et al. 2017; Khalid et al. 2021; Gavish et al. 2015). In some cases, even physical ergonomics has been assessed applying RULA and REBA protocols (Vosniakos et al. 2017).

However, only a few papers focus on the cognitive and psychophysiological conditions of operators in the smart manufacturing context. Among them, Grandi et al. in (Grandi et al. 2020) analysed the quality of human-machine interaction through the use of sensors for user experience analysis during virtual simulations. However, they did not use a structured protocol dedicated to the analysis of the cognitive conditions. Etzi et al. in (Etzi et al. 2019) used VR to simulate the collaboration between human and robot and evaluated not only the system usability and users' performance but also their mental and physical states. To assess the workers' cognitive conditions and eventual stressful episodes related to the tasks, the physiological parameters of heart rate (HR) and skin conductance level (SCL) have been analysed. The differences between slow and fast tasks sessions were computed: even if the users asserted to have a greater level of stress in the fast session, HR and SCL remained stable in the two different sessions. Nevertheless, only a small sample of users and a short temporal window (2 min) was tested (Etzi et al. 2019). In addition, Leone et al. (Leone et al. 2020) proposed a method to analyse the features extracted from the heart rate, electrodermal activity and electrooculography to

distinguish between stressful and relaxed conditions during manufacturing activities such as assembly and manual handling.

However, more attention should be paid to the discrimination between stress and mental load that could arise in a smart environment where the traditional human-machine interaction is subject to changes. Indeed, Operator 4.0, interacting with advanced technological systems (such as collaborative robots, extended reality technologies, etc.), needs to develop the proper skills necessary for the management of the intelligent factory (Romero et al. 2016, Kadir et al. 2020). The development of these new skills must be based on the user's cognitive needs and must guarantee low levels of stress and mental effort. In this scenario, the use of VR-based simulations can offer safe virtual space for testing and validation for design of human-centred production systems, especially in collaborative human-robot workstations (Malik et al. 2019, Ottogalli et al. 2021).

According to the ISO 10075-1, psychological stress is the effect of all conditions with a mental impact on a subject, either cognitive or emotional. It emerges when the perceived demands of the environment exceed a person's ability to cope with these demands (Lazarus and Folkman 1984). Stress is also defined as a "state of high general arousal and negatively tuned emotion, which appears as a consequence of stressors acting upon individuals" (Boucsein 2012). Commonly recognized stressors include technical complications, time pressure, distractions, interruptions, errors, and increased workload (Brunzini et al. 2019).

From a medical point of view, stress is usually described as two general types of response: anxiety or frustration, and the physiological response of the sympathetic nervous system which emerges after a challenge or threat. Concerning this second category, it has been demonstrated that stress causes reactions such as changes in skin conductance (sweating), heart rate (tachycardia), blood pressure (increase), and in the stress hormone cortisol (increase) that spreads to saliva within minutes, during and immediately after performing a stressful task (Sandroni et al. 2005).

The multimodal dimension of stress makes the research field very broad; however, according to ISO 10075-3, four main criteria can be distinguished in detecting stress: psychological, physiological, behavioural, and biochemical. The most common analysis typically includes the subjective assessment based on self-reports (e.g., the State-Trait Anxiety Inventory) and the physiological assessment based on electrocardiography (for heart rate monitoring) and skin conductivity (to measure sweat activity). Indeed, the electrodermal activity (EDA), or Galvanic Skin Response (GSR), reflects the surface changes in skin conductance due to the sympathetic nervous system and it is considered "one of the most sensitive psychophysiological indicators of stress" (Boucsein 2012). Even the heart rate variability (HRV) (i.e., the variability of the inter-beat interval (IBI) in ms) is under the control of the autonomous nervous system that commands our capability to react to external stimuli. For this reason, HRV is considered a reliable indirect means to monitor cognitive states. HRV fluctuations can be analysed using time domain, frequency domain, and non-linear domain methods. Four measures in the time domain (RR, SDRR, RMSSD, and pNN50) and one measure in the non-linear domain (D2) result significantly reduced during stressful events. The ratio LF/HF in frequency domain results instead significantly increased, suggesting a sympathetic activation and a parasympathetic withdrawal during acute stress (Castaldo et al. 2015). Moreover, it has

been shown that the extent of inter-beat variability decreases with increasing cognitive load (Luque-Casado et al. 2016).

The analysis of Cognitive Load (CL) is one of the most widely studied topics in cognitive ergonomics (CE) (Gualtieri et al., 2022, Atici-Ulusu et al. 2021). CE involves psychological processes and concerns humans interacting with other system components (Johnson and Proctor 2013). Some significant items include workload, decision-making, perception, attention, motor response, skill, memory, and learning (Parasuraman et al. 2008). It is oriented towards the optimization of human-machine interaction, according to three main criteria: characteristics of human cognitive processes, software science knowledge, and knowledge in diverse work domain technologies. As a logical consequence, the training topic is included in such perspective, since it can contribute to the enhancement of human performances and work conditions (Green and Hoc 1991).

The increase in professional activities that have a mental dimension has therefore encouraged the development of cognitive ergonomics, which thus results fundamental in the design and assessment of training activities. Indeed, its objective is to improve the performance of cognitive tasks in dynamic and technologically advanced environments, through the design of effective support, understanding the fundamental principles of human activities associated with the principles of engineering design and development.

Nevertheless, even if performance metrics are strongly used to evaluate the users' skills, the assessment of their cognitive state is more uncommon. The introduction of new technologies as virtual reality devices may help the user during the training and the practice but also may result in a potential risk of information overload. For this reason, the study of cognitive load, related to extended reality applications, merits further in-depth analysis.

Cognitive load “emerges from the interaction between the requirements of a task, the circumstances under which it is performed, and the skills, behaviours, and perceptions of the operator” (Hart and Staveland 1988). Since CL can positively or negatively affect human performances, the principal reason for measuring it is to quantify the mental cost of performing a task to predict the performances (Cain 2007).

Current studies mainly refer to three assessment methods: performance assessment method, self-assessment method, and physiological measures method (also according to the ISO 10075-3).

The class of task performance measures assumes that CL is relevant only if it affects performance and the most common measurement parameters are response, reaction time, accuracy, error rate, estimation time, objective speed, and signal detection (Karwowski 2006). However, it is demonstrated that performance errors are not necessarily related to a high mental load imposed by the main activity. For this reason, the secondary task method, in which the user is required to perform a secondary activity concurrently with the main activity, is more used.

The class of self-assessment/subjective measures is based on the personal perceived experience about the interaction with the system and is obtained from the direct estimation of task difficulty. The self-assessment provides information on how humans subjectively evaluate various aspects of workload for accomplishing a task, using questionnaires or psychometric scales.

The class of physiological measures considers physiological responses of the body that are believed to be correlated with the cognitive load. Indeed, changes in psychophysiological parameters, such as HR, HRV, EDA, breathing rate (BR), brain activity (EEG), muscular activity (EMG), eye activity (EOG, pupil diameter, gaze entropy, and velocity), can be indirect indicators of mental workload. The heart rate variations (i.e., the variations of the number of heartbeats per unit of time, typically expressed as beats per minute (BPM)), are proved to be directly related to the mental load (i.e., HR increases as CL increases) (de Waard and Brookhuis 1996). Even the changes in BR reflect variations in the mental effort. Indeed, for an increase in the mental demand, the respiratory rate increases, and the breathing depth decreases (Roscoe 1992). Also, the blinks and eye movements (from electrooculography EOG) have been correlated to cognitive aspects. Researchers demonstrated that blink rate decreases as cognitive load increases (Ledger 2013). Moreover, pupil diameter changes have been shown to be indicative of user engagement and cognitive load for various tasks, also in virtual environments using HMDs (John et al., 2018, John, 2019).

However, these physiological parameters are not selectively optimal indices for measuring mental workload, since they are sensitive to physical activity, strong emotional reactions, environment, and speech. Therefore, it is suggested to use multiple concurrent kinds of measurements to increase the validity of cognitive load assessment (Naismith and Cavalcanti 2015).

3. Methodology

In the training context, physical and mental workload, comfort, and perceived effort must be optimized not only to prevent disorders and stressful conditions, but also to guarantee proper human performances (Pheasant 1986). For this reason, especially when advanced systems are employed (e.g., head-mounted displays, extended reality applications), HF should be analysed to assure effective training.

This study proposes a methodology (Figure 1) for the comprehensive analysis of subjects' mental workload (MWL) and stress, which combines the use of the three main assessment methods found in the literature (i.e., the self-assessment method, the evaluation of the performance, and the assessment of the physiological parameters related to the functioning of the nervous system and the mental conditions). This methodology can be indiscriminately applied during training sessions or other kinds of activity that involve the cognitive domain and could generate stressful situations. Therefore, it can be applied also in the industrial setting during the operator training and while using systems for extended reality.

Concerning the performance assessment, some of the essential KPIs are the execution time, number of consultations, number of errors. Also, specific checklists should be prepared to distinguish among correct, incorrect, and not performed tasks.

For the self-assessment, several surveys could be provided to the subjects. The choice of the most proper questionnaires depends on the investigated activity and on the precise scope of the analysis. However, to assess the MWL and stress, two specific questionnaires have been included in the proposed methodology: the numerical analogue

scale (NAS), and the NASA-Task Load Index (NASA-TLX). The NAS is used to assess the perceived stress quickly and effectively on a 10-points scale (0 = no stress, 10 = very strong stress) (Lesage, Berjot, and Deschamps 2012). The NASA-TLX is used to assess the perceived workload in terms of mental, physical, temporal demands, effort, performance, and frustration (Hart and Staveland 1988). The total score indicates the level of the overall perceived workload (e.g., low, high, very high), while the six domains can be used to investigate the weight of the different elements involved in the workload. This questionnaire has been included in the proposed methodology because it allows to discriminate between the perceived mental demand needed to perform the activity, and other emotional states such as perceived effort, performance, and frustration that can be related both to stress and MWL.

Regarding the psychophysiological parameters, the proposed methodology includes the monitoring of the HR, HRV, EDA, and pupil diameter (PD), as the most relevant for the assessment of stress and MWL (EEG has been excluded because, during the execution of physical activities, the level of noise would compromise the correct signal analysis).

The main innovation of the methodology consists in the combination of the physiological parameters with data related to subjects' perceptions and performance. The needed assessment protocol and the resultant proposed algorithm are respectively described in Section 3.1 and Section 3.2.

3.1 Assessment Protocol

The assessment protocol has been defined to be less intrusive as possible for the operator. Figure 2 shows the general approach that must be used to apply the data elaboration algorithm described in Section 3.2, during training activities with XR technologies. Then, the protocol can be adjusted based on the specific case study.

First, operators are asked to wear the head-mounted display (HMD) for the execution of the training XR application and the non-invasive wearable devices (e.g., smart bands or bracelets, smart glasses) for the monitoring and collection of the physiological parameters (i.e., HR, HRV, EDA, and PD). They must wear them from the arrival in the training room until the end of the post-training self-assessment, to track the variations in the collected parameters while performing different activities. The NAS is administered at rest to record the basal level of perceived stress. Similarly, the NAS is administered again at rest, at the end of the procedure, to verify if the feeling of stress came back to the subject's baseline. Also, before the beginning of the training activity, two minutes of baseline must be recorded from all the wearable devices and for all the physiological signals. This is useful to understand the variation of the biometric parameters between the rest and the training activity. A couple of minutes of warm-up should be provided to let the operator become familiar with the use of the XR technology. During the training session, the operator's performance is evaluated through the recording of KPIs such as execution times, errors, and consultations. Then, after the training session, the NAS scale must be administered again with the NASA-TLX survey.

In this way, it is possible not only to discriminate the perceptions and parameters' variations among different stressful, restful, and mentally demanding situations, but also to combine different assessment methods in an overall, weighted algorithm for the

analysis of stress and MWL, considering all the variables affecting them in an integrated manner.

The training session should be video recorded since it could be useful for the data analysis to stopwatch and track events in relation to physiological variations and times.

3.2 Data Elaboration

First, performance and self-assessment data can be analysed as “stand-alone” results. Indeed, by investigating the committed errors, timing, and consultations as metrics of performance, it is possible to evaluate the quality and the effectiveness of the training. Also, the self-assessment questionnaires give hints on operators’ perceptions related to stress, mental workload, and emotional conditions (such as frustration and effort) which are to be considered when designing (and optimizing) a training path.

Then, in order to objectively analyse stress and mental workload through the physiological parameters, a specific algorithm is proposed. It has been studied to properly integrate the data gathered through the three different methods (performance, self-assessment, and physiological monitoring).

The physiological data collected during the various phases of the test are post-processed similarly to previous studies in this field, aiming in the user experience (UX) monitoring of operators in real (Brunzini et al. 2021) and virtual (Grandi et al. 2022) environments. The general approach is to compare the values of the operators’ physiological data recorded during the test with the values collected during an initial resting phase, called baseline phase. The previous methods have been enhanced in order to develop a more robust and complete algorithm, focusing on the cognitive assessment of operators during the manufacturing task execution in VR. Starting from the various collected data (HR, HRV, EDA, PD), the proposed approach is able to evaluate operators’ mental workload and stress, anticipating potentially dangerous situation. Before the calculation of mental workload and stress score, a set of parameters are necessary, such as *pupil activity (PA)*, *electrodermal activity (EA)*, *heart activity (HA)*, *heart rate variability activity (HVA or RR)*, and *user time (UT)* as defined in (Khamaisi et al. 2022; Brunzini et al. 2021). These parameters are calculated as in Eq. (1):

$$XA = \frac{X \text{ mean} - X \text{ baseline}}{X \text{ max} - X \text{ baseline}} \quad (1)$$

where X stands for the generic physiological parameter, *X mean* is the mean value of the specific user’s physiological parameter as collected during the task execution, *X baseline* is the mean X value as recorded during the user’s baseline phase, and *X max* is the maximum X value reached during the task execution for each user. Similarly, the *user time (UT)* parameter is calculated considering the time to complete the various phases for each user. The time to accomplish the task is clocked for each user and compared with the time employed by the user who took less time (T min) and the time employed by the user who took the longest time (T max).

Then, the subjective assessment is used to weight the calculated physiological parameters, using NAS and NASA-TLX questionnaires. NAS questionnaire investigates the stress score (structured on a 1 to 10 scale) and has been considered to weight the EDA

parameter. With the same approach, NASA-TLX assesses the perceived workload for each user, according to six domains (Mental, Physical, Temporal Demands, Performance, Effort, and Frustration), on a total score of 100 points. Each one of the six domains is assessed on a 500-points scale. In this study, the Mental Demand has been used to weight the PA parameter, the Frustration Level to weight the HVA parameter, the Effort Level to weight the HA parameter, and finally the Overall Performance to weight the UT parameter. Both for NAS and NASA-TLX, each domain is then normalized to a 5-point scale, to achieve the same range of value for all the five weights ($\omega_1, \omega_2, \omega_3, \omega_4, \omega_5$). Table 1 summarizes collected data, used tools, and parameters used in the algorithm.

The above-mentioned parameters, properly weighted as described, are used to generate two metrics: the Mental Workload (MWL) and Stress (S). The first one is computed as in Eq. (2):

$$MWL = \omega_1 * PA + \omega_3 * HA + \omega_5 * UT \quad (2)$$

The Stress is calculated as shown in Eq. (3):

$$S = \omega_2 * |HVA| + \omega_4 * |EA| \quad (3)$$

Combining the “stand-alone” data with the proposed algorithm’s results it is possible to identify eventual mutual correlations, for example between the worsening of performance and the increment in negative cognitive states (such as excessive stress, frustration, and/or mental effort).

4. Industrial Case Study

An interesting application of the proposed methodology is related to the training of operators about assembly tasks, by means of VR technology. Indeed, even in Industry 4.0, the manual assembly, done by the operator (and not by robots), is still widespread, especially in the final stages of assembly, in low-batches processes, and for customizations (Rossi et al. 2020). In order to reduce assembly times and, consequently, to increase the company’s productivity, the training of the operator becomes necessary. Therefore, the effect that the use of technological devices for VR training has on the cognitive conditions of the operator results essential to understand how to avoid cognitive stress and overload, and how to give the most effective learning support.

The case study has been developed in collaboration with CNH Industrial, a global manufacturer of agricultural machines, buses, and trucks. In particular, the collaboration was developed within the San Matteo plant, located in Modena, Italy and the Noida plant, located in India. The first one has one of the most relevant R&D unit in the field of tractors in Europe, using the most advanced technologies for design and engineering purposes such as VR, AR and simulations, while the latter has one of the biggest tractors manufacturing sites of Asia. The selected case study focuses on the exhaust system assembly on a large-size agricultural machine (i.e., the New Holland T7 tractor commercialized by CNH Industrial), which is currently a complex and manual procedure.

This case provides a valuable example of the application of the proposed approach and allows showing how potential stressful situations at the shop floor could be effectively predicted in advance.

4.1 VR Training Application

The VR simulation has been created with the aim to allow users to replicate (as accurately as possible) the assembly procedure to be analyzed. The process to create the VR simulation requires several steps, listed below:

- Step 1. Creation of the virtual scene in the VR platform by importing the 3D CAD models to recreate the assembly workstation layout;
- Step 2. Identification of movable and fixed objects and interaction paths to use in the VR scene by settings of the behaviour of the different objects and components in the scene through the use of scripts;
- Step 3. Choice and setup of the input device controller;
- Step 4. Tests and optimization of VR scene and interactions with virtual objects.

These virtual simulations require the use of several VR technologies to create immersive training experience, integrating motion capture and gesture recognition to replicate a faithful user experience within an assembly workstation layout at the shop floor (Figure 3). This allows creating reliable process simulations in which analyzing operators' experience through the use of human monitoring devices.

The VR simulation entails the use of the following hardware: HTC Vive Pro Eye, a HMD equipped with 32 infrared sensors for the 360-degree tracking, a gyroscope, an accelerometer, and a laser position sensor, that allow for 6 DOF tracking; Leap Motion, a hand-tracking device used to control the objects in the scene, based on gesture recognition.

The immersive training simulation set-up has been created by the following software architecture: Unity 3D, Leap Motion Controller, and Steam VR installed on the same workstation. Unity 3D is the main VR engine platform, for generation of the virtual workstation layout and interaction features, while Steam VR and Leap Motion Controller are required to connect and use the HMD and the hand-tracking device in VR. Leap Motion sensor was placed on the centre of HTC Vive HMD with a specific plastic support.

The virtual environment has been developed based on the tasks sequence of the exhaust system assembly which is detailed in Table 2. To shorten the execution time of the procedure in VR, the tasks sequence has been divided into two consecutive parts (the first one consists in 15 steps, and the second part in 19 steps). The complexity degree of the tasks to be executed is approximately the same in the two different parts. However, the latter can be more challenging, stressful, and mentally demanding because it involves

a greater number of small, similar elements (e.g., screws, nuts) and a wider space for the assembly activities.

Then, a second VR assembly application has been developed with the aim of helping the operator in the tasks' execution. In this case, the standard operating procedure (SOP) have been digitally installed in a virtual panel near the workstation and the part number can be visualized above the component to mount, as shown in Figure 4.

Moreover, the component that must be assembled and its final position on tractor chassis are also suggested in cyan. When the operator is in the correct position, it turns yellow, as shown in Figure 5, in which are reported a couple of frames showing tasks 1-3 of the first part of the assembly process.

4.2 Experimental testing on the industrial case study

After the development of the VR simulation, a preliminary laboratory test has been carried out to verify the feasibility of the protocol setup, focusing on the assessment of cognitive conditions during a training session for assembly. The experimental set-up for user biometric monitoring involved a set of hardware and software tools to collect the necessary data to properly apply the proposed approach:

- Empatica E4: wrist-wearable device that record a set of physiological data of the user, using the photoplethysmogram (PPG) sensor to measure the Blood Volume Pulse (BVP). From the latter, the HR and the IBI signals can be calculated. Differently, the EDA sensor measures the changes in skin conductance resulting from the sympathetic nervous system activity.
- Zephyr BioHarness 3.0: chest-band device that collect data about HR and HRV, able also to monitor user posture (back flexion) through the use of integrated accelerometers and gyroscopes;
- Eye-tracker integrated in HTC Vive Pro eye: hardware that record the eye movements and the pupil diameter with a frequency of 50 Hz.
- E4 realtime: an application for the real time streaming and management of data from Empatica to a smartphone or a tablet. This is used to control Empatica E4 calibration and data recording.
- Omnisense analysis: software for human monitoring data post-processing from BioHarness 3.0.
- iMotions: software that collect and export data about eye activity.

Figure 6 shows the laboratory technological setup, with the user wearing the VR headset and the smart devices for biometric monitoring.

The tests have been conducted in the XiLab laboratory of the Department of Engineering “Enzo Ferrari” of the University of Modena and Reggio Emilia and involved ten users with no previous experience with assembly in VR context, and a mean age of 26.7 years

old (SD = 2.908). Tests were executed involving one user at a time. Participation in the test was voluntary and no reward was given. All participants signed the informed consent after the detailed explanation of the study (from the authors) and the reading of the information sheet. Ethical review and approval were not required for this study since it was conducted in the university laboratory as preliminary investigation.

All participants presented normal vision and did not need corrective lenses, and none of the participants had heart conditions. Firstly, a brief pre-test questionnaire has been filled in by the user to collect demographic information (gender and age), using Google Form. Then, a warm-up session was conducted in order to become familiar with the VR technology. For each participant, a specific code has been assigned (ex. OpX) to keep the data anonymous. The participants demographic information is reported in **Errore. L'origine riferimento non è stata trovata.**

Subsequently, each participant was helped in wearing the devices for physiological signals monitoring defined in the technological set-up. NAS and NASA-TLX questionnaires data were collected using Google Form. Physiological signals recording was performed from the beginning until the end of the test, both during the tasks execution and while answering the questionnaires. Before the assembly task execution, three-minutes of signal recording from the wearable sensors were recorded with user in resting condition (upright, being still) in order to analyze the baseline and the supposed variations in physiological signals. As previously described, the training has been divided into two sessions: while the first session is provided without showing the standard operating procedure (SOP) in the virtual environment, the second session includes the visualization of a box with the explanation of the tasks to be executed. Following this procedure, it can be supposed that the levels of cognitive load and stress will be higher in the first session, both because the user has to become familiar with the use of VR and because instructions are not supplied. A lowering of MWL and stress is therefore expected in the second session, where the operator knows how to use the VR device and is helped by the SOPs.

Therefore, the test protocol has been adapted, based on the one previously described in section 3.1, dividing the entire tasks sequence into four macro-phases and questionnaires filling (Figure 7).

The NAS and NASA-TLX are administered many times primarily to analyse the impact that the two different virtual configurations have on the operator's cognitive conditions.

The overall duration of the test for each user is about one hour and a half, and the temporal length of the experimental procedure is similar to the one needed for the assembly activity at the shop floor. The temperature in the room was measured and was constant at 26 °C, thanks to the air conditioning system. The light sources were totally artificial, provided by neon lamps positioned on the lab ceiling. The tests have been video-recorded using two cameras: an external camera in a fixed position, and the eye-tracking camera mounted on eye-tracking glasses, providing the users' viewpoint.

5. Results and Discussion

The experimental testing on the industrial case study has confirmed the feasibility of the proposed assessment methodology for the analysis of operators' stress and cognitive

conditions during training activities with VR technology.

The experimental test has been divided into two separate sessions of different difficulty to make comparisons and consequently accomplish a threefold aim:

- To understand if it is useful, in terms of performance, to supply the operator with virtual support (i.e., to verify if the extra-aid allows a reduction of committed errors, consultations, and tasks execution time).
- To study the effect of furnishing VR support on operators' perceptions in terms of subjective stress, mental demand, and overall workload.
- To analyse the impact of giving VR support on operators' physiological response, simultaneously considering the weights of performance and self-assessment.

Accordingly, results are presented in three different sections: section 5.1 shows the results related to operators' performance, section 5.2 describes the subjective assessments, and section 5.3 analyses the results of the proposed comprehensive algorithm.

5.1 Performance

Performance have been assessed in terms of execution times, committed errors and consultations. Results are shown as boxplots (data distribution and mean values) to evaluate the differences between the training session with virtual support and the one without it (i.e., comparisons between the first and the second session). Also, comparison between the two different tasks' parts, for each session, is available. Figure 8 shows the boxplots with the execution times of the two different parts with and without virtual SOPs. Similarly, Figure 9 shows the boxplots with the committed errors, and Figure 10 shows the boxplots with the number of consultations for the two parts, for the first and second session.

It is possible to confirm the correctness of the trial design, preliminary assumptions, and expected results: the first session (1st and 2nd parts without support) resulted in higher execution times, greater numbers of errors and consultations than the second session (1st and 2nd parts with support). Therefore, the first session proved to be more difficult due to the operators' inexperience and lack of virtual support.

Thus, it is clear how providing a VR support helps the operator in the tasks' execution. Indeed, a decrement in the minutes employed to finish the tasks is evident in both the 1st and 2nd parts when giving virtual aids (Figure 8). Also, the number of committed errors decreases in both parts in the second session (Figure 9). In both parts the number of committed errors is almost halved when the operator can rely on the virtual SOPs and aids.

Even for the consultations, the decrement is visible in both parts in the second session (Figure 10). In particular, while the 1st part was easier and required less consultations, when executing the 2nd one without support, the operators asked for help more frequently. The provision of virtual support highly reduced this issue.

Therefore, supplying the operator with virtual support resulted useful and effective from the performance perspective. The reduction in execution times, errors, and consultations during the training would consequently positively affect also the quality of work and operators' satisfaction.

5.2 Self-Assessment of Stress and Mental Workload

In this section, results about the operators' self-assessment related to perceived stress (NAS) and workload (NASA-TLX) are reported.

Table 4 shows the mean values and standard deviation of NAS answered at rest (before and after the training activity), and after each training part (1st and 2nd parts with and without support).

As expected, an increment of perceived stress is visible from rest to the training activity. Also, the perceived stress comes back to the baseline level at the end of the procedure (end - rest). However, great differences in the perceived stress among the different training parts are not present. Mean values during the training activity are all around 4,7 (on the 10-points scale), indicating medium levels of perceived stress. The introduction of the virtual support leads, on average, to a slight increment in the 1st part and to a minimal decrement in the 2nd part. Thus, the perceived feeling of stress remains stable regardless of the VR content. Giving hints and helps through VR SOPs does not affect the operator's perceptions about fatigue and overload.

The NASA-TLX total score indicates the level of perceived workload (on a one-hundred based scale) considering the weighted domains of mental demand, physical demand, temporal demand, performance, effort, and frustration. In this case, with the virtual support provided to the operator, the total score, on average, slightly decreases in the 1st part and increases in the 2nd one (Table 5). This is probably related to the different difficulty of the two parts. Giving VR support for complicated tasks' sequences (2nd part) reduces the stress but, at the same time, increases the perceived workload. The operator has to deal both with the tasks and with the virtual content, and this may lead to increment in effort and perceived mental workload.

The same result is confirmed by the analysis of the six domains (Figure 11). In the 1st part, the mental demand, the temporal demand, and performance decrease with the VR support. Instead, the mental demand, the temporal demand, and the effort on average increase moving from the 2nd part without support to the 2nd part with support. However, in the meantime, performance and frustration slightly decrease; in the 2nd part of the second session, the pressure related to performance diminishes and the feeling of frustration is also reduced. The physical demand, compared to the other domains, is obviously always very low, because the simulation did not consider real instruments.

In general, according to the Sugarindra et al. score interpretation (Sugarindra, Suryoputro, and Permana 2017), NASA-TLX after each training part (with and without support) is on average high (high perceived workload for total scores of 50-79, on the 100-points scale)

(Table 5). Therefore, operators perceive the training as a high demanding activity, regardless of the virtual support.

5.3 Computed Stress and Mental Workload

Results about stress and mental workload computed through the proposed algorithm allows studying the effect of the virtual aid on the operators' physiological response, taking into account also performance and self-assessment.

Figure 12 shows the trends of stress and MWL from the 1st to the 2nd part for the first and second session (relying or not on the virtual aids). For both training parts, stress and mental workload increase when virtual support is provided. The increment in MWL is slightly higher in the 1st part, while the increase of stress is noteworthy only in the 2nd part. These results are not consistent with the ones of the self-assessment (red circles in Figure 12).

From the self-assessments, in the 1st part the perceived stress, on average, slightly increases with the use of VR support, while it minimally decreases with VR support in the 2nd part. Concerning the mental workload, from the NASA-TLX, when VR SOPs are provided, a decrement of the mental demand is found in the 1st part, while a small increment is observed in the 2nd part. Instead, computing the stress and the MWL with the proposed algorithm, when virtual support is provided, a remarkable increment of stress can be noticed in the 2nd part, and a small increase in MWL in the 1st part. These results could be explained referring to the intrinsic difficulty of the tasks. Indeed, the tasks of the 1st part are easier than the ones in the 2nd part. In the 2nd part, a greater number of small and similar components (e.g., screws, nuts), displaced in a wider action area, is present rather than in the 1st part. Thus, providing VR support when tasks are simple (1st part) could lead to an increment in the MWL: the trainee is asked to perform the tasks and read the virtual SOPs that may introduce an additional mental effort where not required (the operator could be able to correctly perform the tasks' sequence without external help). Similarly, supplying VR support with multiple information (e.g., SOPs, part numbers, colour-based suggestions) when tasks are difficult to perform or to remember (2nd part) helps the operator in the right execution (see performance results), but it may also increase the stress. Indeed, commonly recognized stressors include complications, interruptions, increased workload, which may all be introduced by information overload. Thus, the VR content may result in an overloading and ineffective support. For this reason, it is extremely important analysing the impact of VR training not only on the performance but also on the stress and cognitive conditions of the trainee. In this case, the virtual content should be re-design trying to avoid the increment of stress for the operator. Indeed, even if the perceived stress is not affected by the provided virtual content, the proposed methodology, combining the physiological response with the performance and subjective assessment, allows identifying potential stressful and overloading conditions, due to the VR content itself.

Even if subjective measures present low application costs and lack of interference with on-going tasks, they present some limitations, mostly connected to the difficulty in quantifying the perceived mental effort and level of stress. For this reason, although the high levels of inter- and intra-individual variability of biometric indices, current findings

suggest that physiological parameters are the most sensitive means for detecting variations in stress and cognitive load levels during training activities. However, the combination of the three methods (also including the performance) in a unique algorithm allows for a more precise and objective assessment.

Conclusions and Perspectives

The importance of training in the modern industry is widely highlighted in the literature. Modern factories should support novel operators in learning new skills and abilities, and digital solutions, like VR and AR, seem to be a good tool to introduce new technologies to the factory floor. In this context, work training processes need to be reshaped and new approaches are needed in order to support operators' continuous development of skills. The great increase in the use of advanced technologies in the manufacturing context arises the necessity to switch from a traditional training approach, based on paper and video instructions, to a more engaging and effective model.

VR applications have the potential to re-invent the entire training process; for this reason, VR training is becoming more and more adopted in the educational field. On one side, it provides a more immersive and safer environment but, on the other side, it could distract or overload the user, making the tasks more difficult, if it is not properly designed. As a matter of fact, VR training applications must be carefully designed considering human capabilities, limitations, perceptions, and the users' cognitive responses. Indeed, the use of VR devices such as HMDs or glasses, can require an additional mental and physical burden, demanding different skills and experience.

In this context, the adoption of a human-centered approach is compulsory for the creation of successful training paths. This paper proposed a structured protocol for the assessment of performance and cognitive conditions of operators during training with VR applications. The proposed approach relies on several wearable devices that record the user physiological signals in order to estimate the user's cognitive state through the use of an algorithm able to assess mental workload and stress and detect potentially dangerous situations. Experimental results allow a further optimization of the VR training app from a human-centric perspective, permitting a consequent improvement in the operators' knowledge and skills.

It is expected that VR training shortens the learning period thanks to a lower number of errors and a less demanding procedure. The obtained results highlighted these aspects in terms of performance: operators' errors decreases drastically when they are supported by additional information on the tasks procedure. However, in the meantime, the proposed algorithm showed a stress increment when providing virtual support, for complex tasks, with too much information to manage. For this reason, only by safeguarding the cognitive ergonomics of the operator, VR training would be effective and would have a positive effect on the productivity of the company, saving money and resources.

The main limitation of this study refers to the lack of comparison with other kinds of training in terms of effectiveness and cognitive impact on operators. Indeed, a comparison with traditional learning methods should be done to distinguish between the cognitive states (MWL and S) due the training tasks, the ones due to the developed VR application (in terms of content, layout, easiness of use, etc.), and the ones due to the use

of the new technology itself (i.e., HMD and virtual environment). Specifically, the comparison with reality could highlight several aspects such as the difference in the level of cognitive load (e.g., it may be higher because Human-Computer Interfaces are difficult to master, or it might be lower because the task itself is made simpler, compared to reality). Another limitation refers to the reduced sample of users and their lack of knowledge about the real assembly process.

Future works will consist in the assessment of the developed VR training application with a larger sample of users, involving both novel and expert operators from the plants. A comparative statistical randomized study will give the opportunity also to compare virtual training procedures and traditional training practices, performed through the use of documents and videos, analysing the differences between the standard training and, hopefully, the benefits of VR technology, in terms of performance, comfort, and cognitive wellness. A specific user experience and usability survey will be administered to analyse the operators' acceptance and subjective opinions about the use of VR and HMDs for training, and consequently to improve the developed application in a user-centred perspective. Thus, next steps will also include the optimization of the VR application in terms of usability and learning content. Indeed, after this analysis, it will be possible to re-design the VR training application, considering the requested mental demand, and thus avoiding stress, mental overload, and improving the overall performance.

Acknowledgments

The authors wish to acknowledge CNH Industrial for the precious collaboration.

Funding Details

This work was supported by the project "FAR 2021 Attrezzature Misura 2" of the University of Modena and Reggio Emilia, Principal Investigator Prof. Margherita Peruzzini.

Disclosure statement

The authors report that there are no competing interests to declare.

References

- Abich, J., Parker, J., Murphy, J. S., & Eudy, M. (2021). A review of the evidence for training effectiveness with virtual reality technology. *Virtual Reality*, 25(4), 919-933.
- Abidi, Mustufa Haider, Abdulrahman Al-Ahmari, Ali Ahmad, Wadea Ameen, and Hisham Alkhalefah. 2019. "Assessment of Virtual Reality-Based Manufacturing Assembly Training System." *The International Journal of Advanced Manufacturing Technology* 105 (9). Springer: 3743-59.
- Atici-Ulusu, H., Ikiz, Y. D., Taskapilioglu, O., & Gunduz, T. (2021). Effects of augmented reality glasses on the cognitive load of assembly operators in the

- automotive industry. *International Journal of Computer Integrated Manufacturing*, 34(5), 487-499.
- Borsci, Simone, Glyn Lawson, and Simon Broome. 2015. "Empirical Evidence, Evaluation Criteria and Challenges for the Effectiveness of Virtual and Mixed Reality Tools for Training Operators of Car Service Maintenance." *Computers in Industry* 67. Elsevier: 17–26.
- Boucsein, Wolfram. 2012. *Electrodermal Activity*. Springer Science & Business Media.
- Brunzini, Agnese, Alessandra Papetti, Eleonora Brandoni Serrani, Martina Scafà, and Michele Germani. 2019. "How to Improve Medical Simulation Training: A New Methodology Based on Ergonomic Evaluation." In *International Conference on Applied Human Factors and Ergonomics*, 145–55. Springer.
- Brunzini, Agnese, Margherita Peruzzini, Fabio Grandi, Riccardo Karim Khamaisi, and Marcello Pellicciari. 2021. "A Preliminary Experimental Study on the Workers' Workload Assessment to Design Industrial Products and Processes." *Applied Sciences* 11 (24). MDPI: 12066.
- Cain, Brad. 2007. "A Review of the Mental Workload Literature." Defence research and development Toronto (Canada).
- Castaldo, Rossana, Paolo Melillo, Umberto Bracale, M Caserta, Maria Triassi, and Leandro Pecchia. 2015. "Acute Mental Stress Assessment via Short Term HRV Analysis in Healthy Adults: A Systematic Review with Meta-Analysis." *Biomedical Signal Processing and Control* 18. Elsevier: 370–77.
- Etzi, Roberta, Siyuan Huang, Giulia Wally Scurati, Shilei Lyu, Francesco Ferrise, Alberto Gallace, Andrea Gaggioli, Alice Chirico, Marina Carulli, and Monica Bordegoni. 2019. "Using Virtual Reality to Test Human-Robot Interaction during a Collaborative Task." In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 59179:V001T02A080. American Society of Mechanical Engineers.
- Gavish, Nirit, Teresa Gutiérrez, Sabine Webel, Jorge Rodríguez, Matteo Peveri, Uli Bockholt, and Franco Tecchia. 2015. "Evaluating Virtual Reality and Augmented Reality Training for Industrial Maintenance and Assembly Tasks." *Interactive Learning Environments* 23 (6). Taylor & Francis: 778–98.
- Gorecky, D., Khamis, M., & Mura, K. 2017. "Introduction and establishment of virtual training in the factory of the future." *International Journal of Computer Integrated Manufacturing*, 30(1), 182-190.
- Grandi, F., Peruzzini, M., Cavallaro, S., Prati, E., and Pellicciari, M.. 2022. "Creation of a UX Index to Design Human Tasks and Workstations." *International Journal of Computer Integrated Manufacturing* 35 (1). Taylor & Francis: 4–20.
- Grandi, Fabio, Luca Zanni, Margherita Peruzzini, Marcello Pellicciari, and Claudia Elisabetta Campanella. 2020. "A Transdisciplinary Digital Approach for Tractor's Human-Centred Design." *International Journal of Computer Integrated Manufacturing* 33 (4). Taylor & Francis: 377–97.
- Green, Thomas R G, and J-M Hoc. 1991. "What Is Cognitive Ergonomics?" *Le Travail Humain*. JSTOR, 291–304.
- Gualtieri, L., Rauch, E., & Vidoni, R. (2022). Human-robot activity allocation algorithm for the redesign of manual assembly systems into human-robot

- collaborative assembly. *International Journal of Computer Integrated Manufacturing*, 1-26.
- Hart, Sandra G., and Lowell E. Staveland. 1988. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." *Advances in Psychology* 52 (C): 139–83. doi:10.1016/S0166-4115(08)62386-9.
- Hoedt, Steven, Arno Claeys, Hendrik van Landeghem, and Johannes Cottyn. 2017. "The Evaluation of an Elementary Virtual Training System for Manual Assembly." *International Journal of Production Research* 55 (24). Taylor & Francis: 7496–7508.
- John, B., Raiturkar, P., Banerjee, A., & Jain, E. (2018, November). An evaluation of pupillary light response models for 2D screens and VR HMDs. *In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (pp. 1-11).
- John, B. (2019, June). Pupil diameter as a measure of emotion and sickness in VR. *In Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (pp. 1-3).
- Johnson, Addie, and Robert Proctor. 2013. *Neuroergonomics: A Cognitive Neuroscience Approach to Human Factors and Ergonomics*. Springer.
- Kadir, B. A., & Broberg, O. (2020). Human well-being and system performance in the transition to industry 4.0. *International Journal of Industrial Ergonomics*, 76, 102936.
- Karwowski, Waldemar. 2006. *International Encyclopedia of Ergonomics and Human Factors*, -3 Volume Set. Crc Press.
- Khalid, Shah, Sehat Ullah, Numan Ali, Aftab Alam, Nasir Rasheed, Muhammad Fayaz, and Masood Ahmad. 2021. "The Effect of Combined Aids on Users Performance in Collaborative Virtual Environments." *Multimedia Tools and Applications* 80 (6). Springer: 9371–91.
- Khamaisi, Riccardo Karim, Agnese Brunzini, Fabio Grandi, Margherita Peruzzini, Marcello Pellicciari. 2022. "UX assessment strategy to identify potential stressful conditions for workers." *Robotics and Computer-Integrated Manufacturing*, 78 102403. Elsevier.
- Lazarus, Richard S, and Susan Folkman. 1984. *Stress, Appraisal, and Coping*. Springer publishing company.
- Ledger, Hayley. 2013. "The Effect Cognitive Load Has on Eye Blinking." University of Plymouth.
- Leone, Alessandro, Gabriele Rescio, Pietro Siciliano, Alessandra Papetti, Agnese Brunzini, and Michele Germani. 2020. "Multi Sensors Platform for Stress Monitoring of Workers in Smart Manufacturing Context." *In 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 1–5. IEEE.
- Lesage, F-X, S Berjot, and F Deschamps. 2012. "Clinical Stress Assessment Using a Visual Analogue Scale." *Occupational Medicine* 62 (8). Oxford University Press UK: 600–605.
- Luque-Casado, Antonio, José C Perales, David Cárdenas, and Daniel Sanabria. 2016. "Heart Rate Variability and Cognitive Processing: The Autonomic Response to Task Demands." *Biological Psychology* 113. Elsevier: 83–90.

- Malik, A. A., Masood, T., & Bilberg, A. 2020. "Virtual reality in manufacturing: immersive and collaborative artificial-reality in design of human-robot workspace." *International Journal of Computer Integrated Manufacturing*, 33(1), 22-37.
- Mattsson, Sandra, Åsa Fast-Berglund, Dan Li, and Peter Thorvald. 2020. "Forming a Cognitive Automation Strategy for Operator 4.0 in Complex Assembly." *Computers & Industrial Engineering* 139. Elsevier: 105360.
- Naismith, Laura M, and Rodrigo B Cavalcanti. 2015. "Validity of Cognitive Load Measures in Simulation-Based Training: A Systematic Review." *Academic Medicine* 90 (11). LWW: S24–35.
- Otto, Michael, Eva Lampen, Philipp Agethen, Mareike Langohr, Gabriel Zachmann, and Enrico Rukzio. 2019. "A Virtual Reality Assembly Assessment Benchmark for Measuring VR Performance & Limitations." *Procedia CIRP* 81: 785–90. doi:10.1016/j.procir.2019.03.195.
- Ottogalli, K., Rosquete, D., Rojo, J., Amundarain, A., María Rodríguez, J., & Borro, D. 2021. "Virtual reality simulation of human-robot coexistence for an aircraft final assembly line: process evaluation and ergonomics assessment." *International Journal of Computer Integrated Manufacturing*, 34(9), 975-995.
- Papetti, Alessandra, Marianna Ciccarelli, Agnese Brunzini, and Michele Germani. 2021. "Design of Ergonomic Manufacturing Equipment by a Human-Centered Methodology." *International Journal on Interactive Design and Manufacturing (IJIDeM)* 15 (1). Springer: 107–11.
- Parasuraman, Raja, Thomas B Sheridan, and Christopher D Wickens. 2008. "Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs." *Journal of Cognitive Engineering and Decision Making* 2 (2). Sage Publications Sage CA: Los Angeles, CA: 140–60.
- Peruzzini, M., F. Grandi, S. Cavallaro, and M. Pellicciari. 2020. "Using Virtual Manufacturing to Design Human-Centric Factories: An Industrial Case." *International Journal of Advanced Manufacturing Technology*. doi:10.1007/s00170-020-06229-2.
- Peruzzini, Margherita, Fabio Grandi, and Marcello Pellicciari. 2020. "Exploring the Potential of Operator 4.0 Interface and Monitoring." *Computers and Industrial Engineering* 139 (December 2018). Elsevier: 105600. doi:10.1016/j.cie.2018.12.047.
- Pheasant, Stephan. 1986. "Bodyspace: Anthropometry." CUMINCAD.
- Romero, David, Johan Stahre, Thorsten Wuest, Ovidiu Noran, Peter Bernus, Åsa Fast-Berglund, and Dominic Gorecky. 2016. "Towards an Operator 4.0 Typology: A Human-Centric Perspective on the Fourth Industrial Revolution Technologies." *CIE 2016: 46th International Conferences on Computers and Industrial Engineering*, no. October: 0–11.
- Roscoe, Alan H. 1992. "Assessing Pilot Workload. Why Measure Heart Rate, HRV and Respiration?" *Biological Psychology* 34 (2–3). Elsevier: 259–87.

- Rossi, Marta, Alessandra Papetti, Michele Germani, and Marco Marconi. 2020. "An Augmented Reality System for Operator Training in the Footwear Sector." *Comput. Des. Appl* 18: 692–703.
- Rubio, E. M., Sanz, A., & Sebastián, M. A. (2005). Virtual reality applications for the next-generation manufacturing. *International Journal of Computer Integrated Manufacturing*, 18(7), 601-609.
- Sandroni, Claudio, Peter Fenici, Fabio Cavallaro, Maria Grazia Bocci, Andrea Scapigliati, and Massimo Antonelli. 2005. "Haemodynamic Effects of Mental Stress during Cardiac Arrest Simulation Testing on Advanced Life Support Courses." *Resuscitation* 66 (1). Elsevier: 39–44.
- Sugarindra, M., M. R. Suryoputro, and A. I. Permana. 2017. "Mental Workload Measurement in Operator Control Room Using NASA-TLX." *IOP Conference Series: Materials Science and Engineering* 277 (1). doi:10.1088/1757-899X/277/1/012022.
- Vosniakos, G-C, J Deville, and E Matsas. 2017. "On Immersive Virtual Environments for Assessing Human-Driven Assembly of Large Mechanical Parts." *Procedia Manufacturing* 11. Elsevier: 1263–70.
- Waard, Dick de, and K A Brookhuis. 1996. "The Measurement of Drivers' Mental Workload." Groningen University, Traffic Research Center Netherlands.
- Zolotová, Iveta, Peter Papcun, Erik Kajáti, Martin Miškuf, and Jozef Mocnej. 2020. "Smart and Cognitive Solutions for Operator 4.0: Laboratory H-CPPS Case Studies." *Computers & Industrial Engineering* 139. Elsevier: 105471.

TABLES

Table 1. Collected data

COLLECTED DATA	TOOLS	MONITORED PARAMETERS	USED IN THE ALGORITHM
PERFORMANCE	Reports per tasks unit	Number of attempts	-
		Number of errors	-
		Completion Time	UT
PHYSIOLOGICAL PARAMETERS	Chest band	HR, HRV	HA, HVA
	Wristband	EDA	EA
	Eye Tracker	PD	PA
SUBJECTIVE ASSESSMENT	NAS scale	Perceived stress	ω_4
	NASA-TLX questionnaire	Perceived workload:	-
		Mental Demand	ω_1
		Physical Demand	-
		Temporal Demand	-
		Performance	ω_5
		Effort	ω_3
Frustration	ω_2		

Table 2. Tasks sequence of the analyzed assembly procedure

FIRST PART	Pick the clamp hose exhaust and sub-assembly of DOC.
	Align the sub-assembly of DOC to the engine exhaust as marked and place over DOC mounting.
	Mount the DOC with the help of clamp.
	Pick the clamp hose exhaust and sub-assembly of DOC.
	Place the clamp assembly over DOC. By aligning the holes of DOC. Mounting. Insert the 4 flange bolts one by one manually up to 3 threads.
	Flange bolts with the help of gun-socket.
	Pick the shield DOC. and Insulation DOC.
	Align the DOC cover with the holes of DOC.
	Place the DOC cover over DOC.
	Pick the 4 flange bolts and 2 shield expansion tanks.
	Place the DOC cover over DOC. shield expansion tank over DOC. Cover and align the flange bolts.
	Torque up the bolts with help of gun.
	Pick the sensor temp 35 mm.
	Mount the sensor to the DOC.
	Connect the sensor pig tail to the electrical connector.
Pick the pipe vertical WA silencer with the help of tackle.	
Pick the exhaust hose clamp.	
SECOND PART	Insert the hose clamp over muffler DOC, do not tighten the clamp now.
	Align the silencer to the bkt. Exhaust system & fit the DOC pipe with silencer with the help of clamp. Now tighten the clamp now tighten the clamp.
	Pick 1 bolt and 2 nuts.

Insert the bolt to the bkt. Exhaust system and tighten manually up to 3 threads then with socket and gun.

Align both the nut one by one with the mounting studs.

Tighten the nuts manually up to 3 threads and then torque up with the help of socket – gun.

Pick the front hood sub-assembly with the help of front hood tackle.

Pick 2 NY lock nut M6 and 2 washers together.

Align the mounting studs of front hood with the holes of radiator sell and insert the studs to the radiator sell.

Pick each washer with each nut.

Insert the washer-nut over the mounting studs of front hood, tighten the nuts manually up to 3 threads and the with the help of socket-gun.

Lock the front hood with the help of 2 locking latch mounted on both side of front hood as shown. Be ensure the proper fitment of latch lock.

Pick the Centre panel sub-assembly by tackle as shown from the trolley carefully.

Place the centre panel over tractor move panel towards the indicating direction and pick 2 bolts, 2washer together by inserting washer over bolt.

Insert 2 bolts to the marked position with washer, tighten the bolts manually up to 3 threads then by socket-gun.

Pick 1 bolt, 1 washer, 1 spacer and 1 nut. Hold the spacer with nut to the inside of panel with spanner and insert the bolt with washer to the marked position and tighten the bolt by socket-gun.

Pick 1 bolt, 1 washer, 1 spacer and 1 nut. Hold the spacer with nut to the inside of panel with spanner and insert the bolt with washer to the marked position and tighten the bolt by socket-gun.

Table 3. Participants demographic information

CODE	GENDER	AGE
OP 1	M	32
OP 2	M	27
OP 3	M	29
OP 4	M	26
OP 5	M	30
OP 6	M	24
OP 7	M	24
OP 8	F	27
OP 9	F	23
OP 10	M	25

Table 4. NAS mean values and standard deviation

	MEAN	STANDARD DEVIATION
START - REST	3,30	1,95
1ST PART WITHOUT SUPPORT	4,60	1,58
1ST PART WITH SUPPORT	4,80	1,87
2ND PART WITHOUT SUPPORT	4,80	2,10
2ND PART WITH SUPPORT	4,70	1,64
END - REST	3,17	1,47

Table 5. NASA-TLX total score: mean values and standard deviation

	MEAN	STANDARD DEVIATION
1ST PART WITHOUT SUPPORT	52,27	12,45
1ST PART WITH SUPPORT	50,80	20,30
2ND PART WITHOUT SUPPORT	49,80	12,39
2ND PART WITH SUPPORT	51,45	20,82

FIGURES

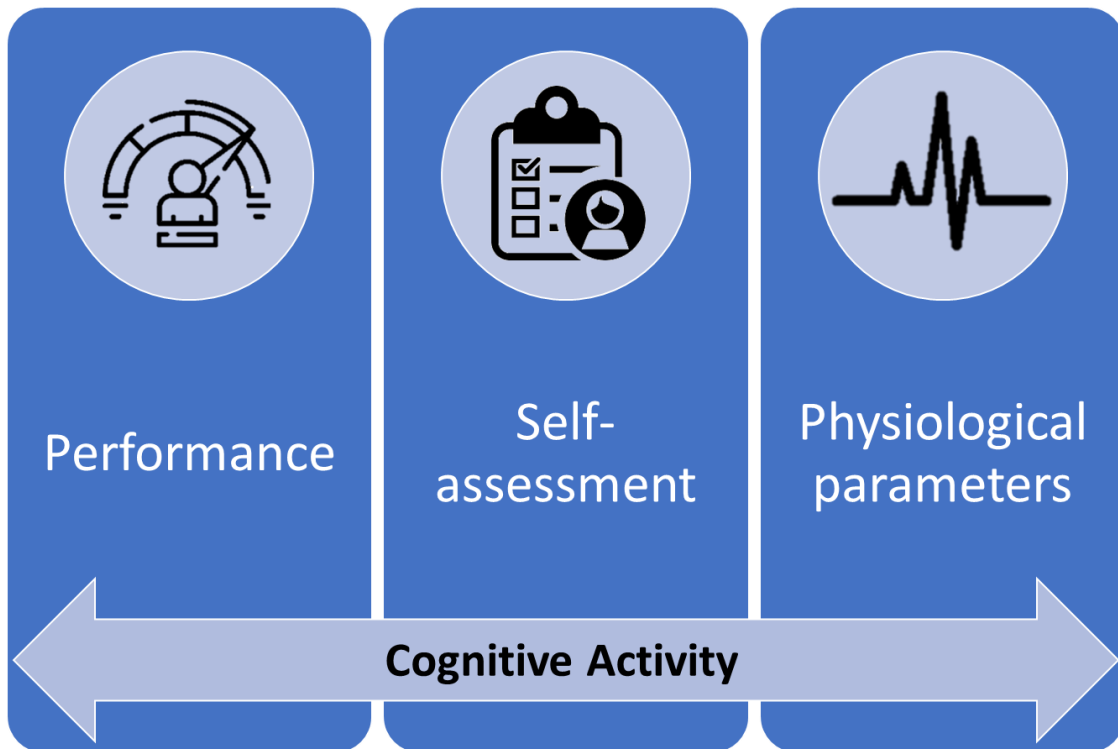


Figure 1. Overall methodology for the assessment of mental workload and stress

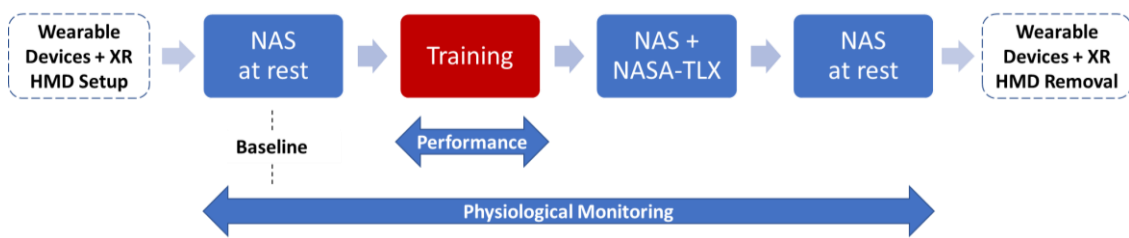


Figure 2. Protocol for the assessment of cognitive conditions (MWL and stress)

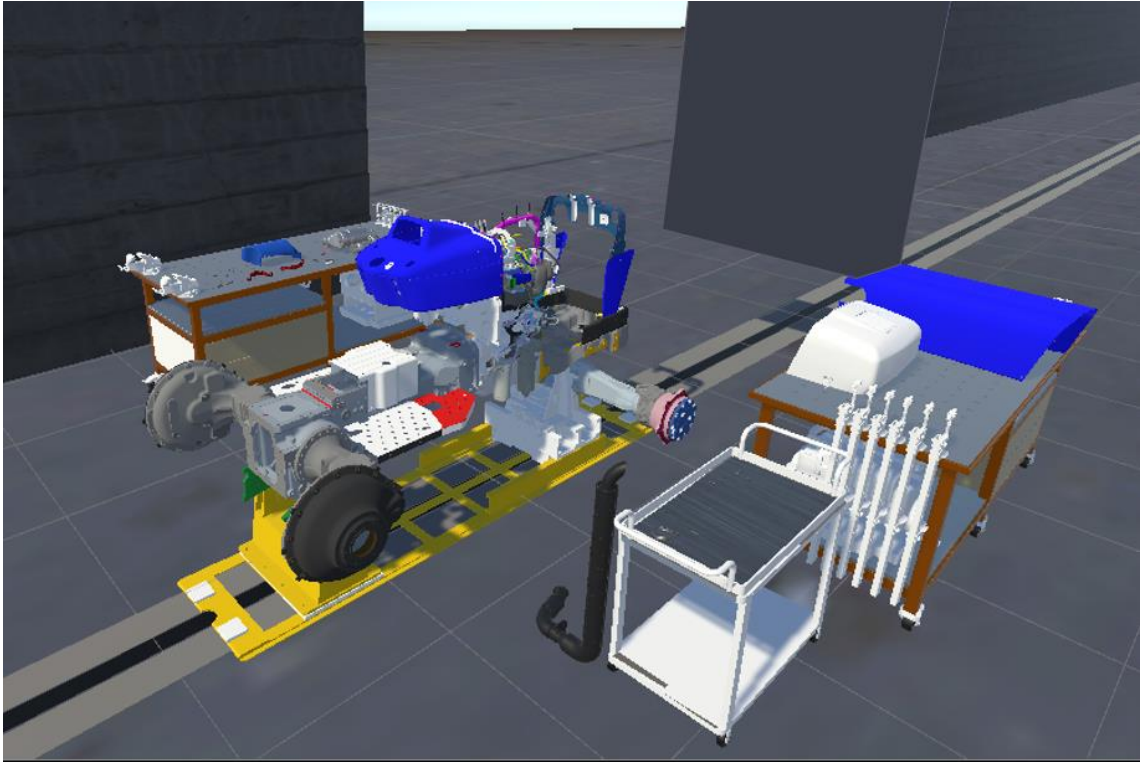


Figure 3. Assembly workstation in VR environment

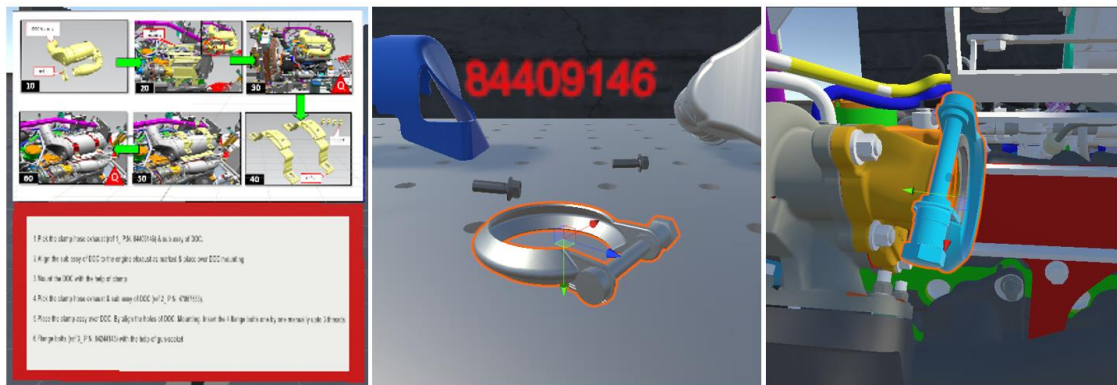


Figure 4. Assembly supports tools in VR environment (SOP panel, part number and final position)

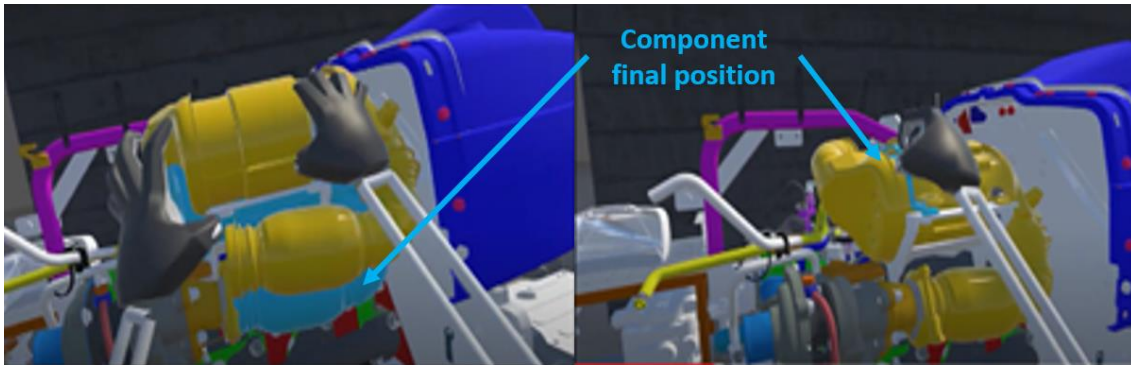


Figure 5. Components final position indications

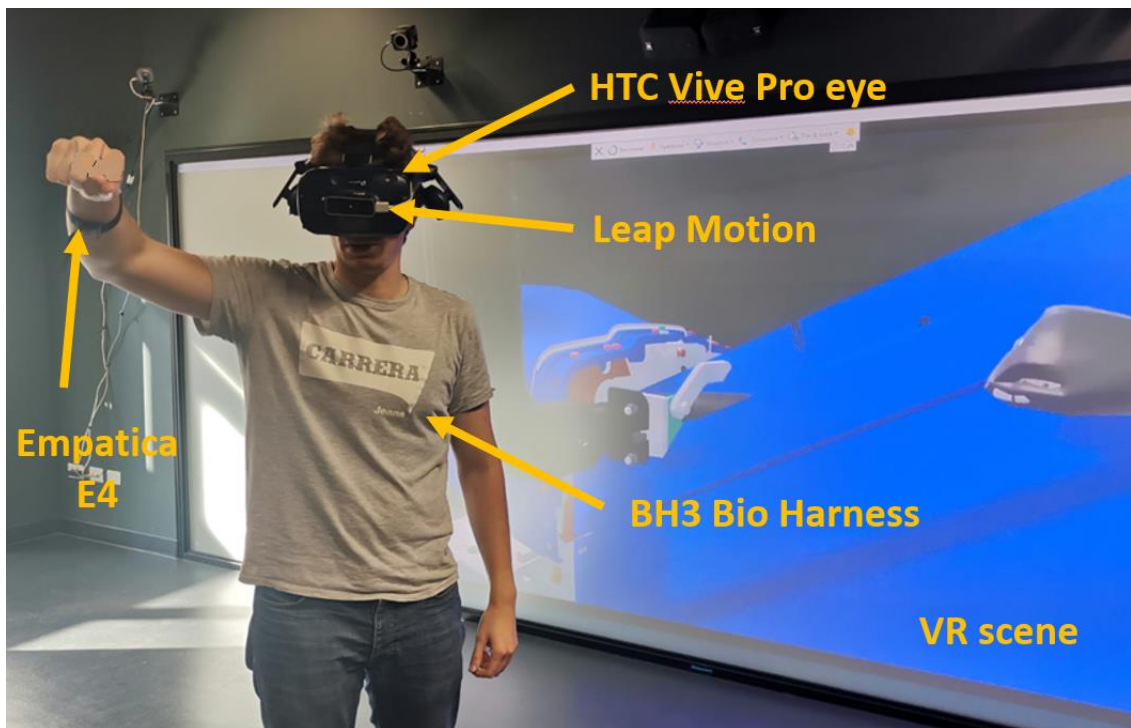


Figure 6. External view of the trial and the used technological setup

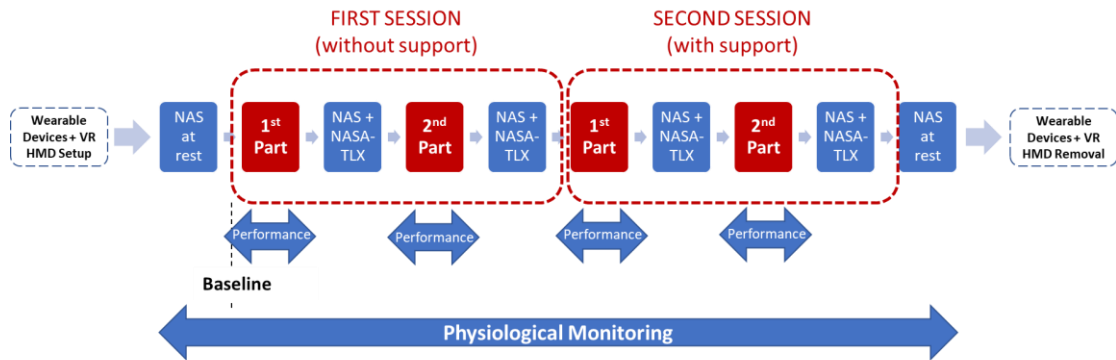


Figure 7. Experimental testing protocol

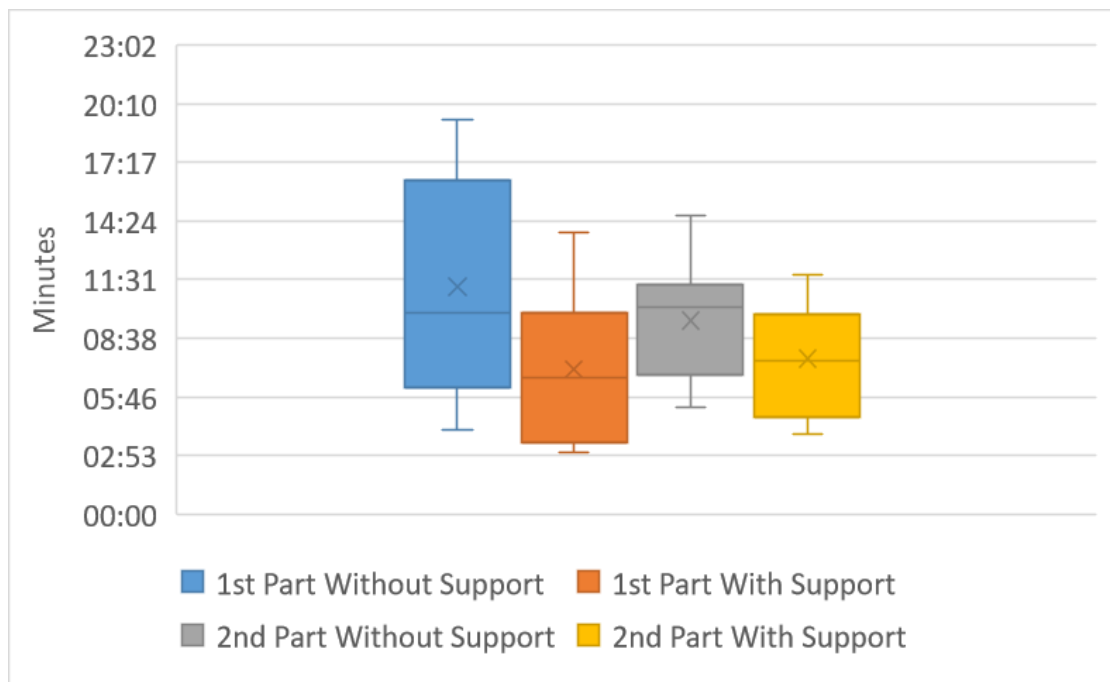


Figure 8. Boxplots of execution times (1st and 2nd parts, with and without VR support)

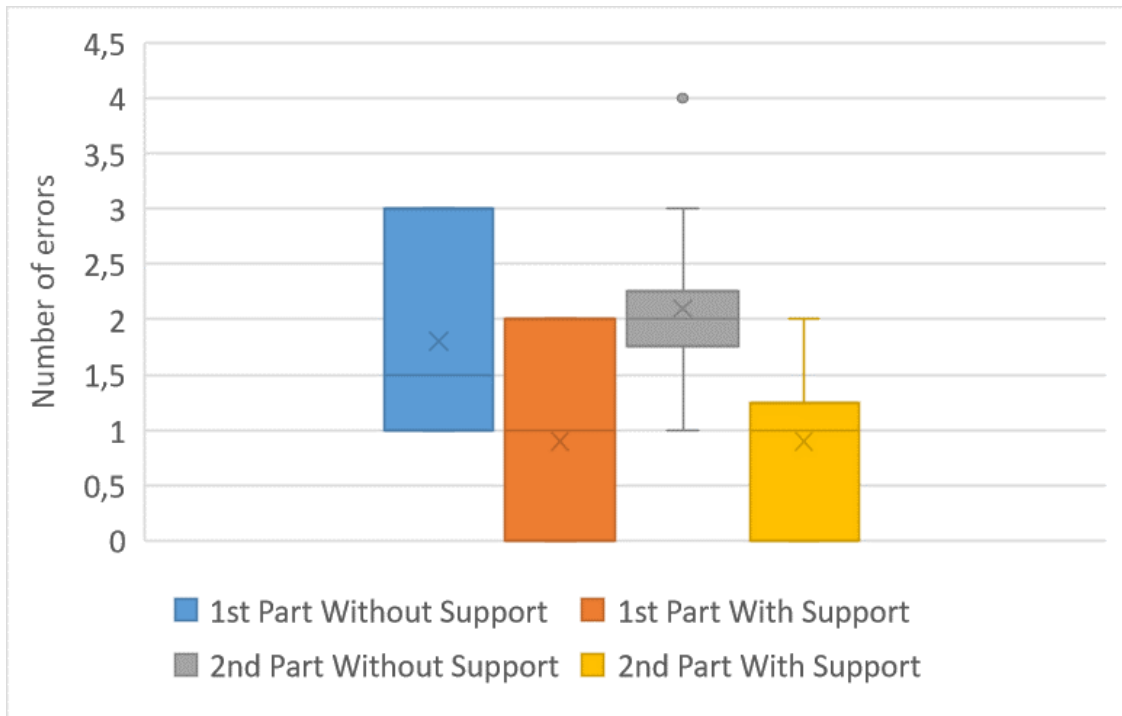


Figure 9. Boxplots of committed errors (1st and 2nd parts, with and without VR support)

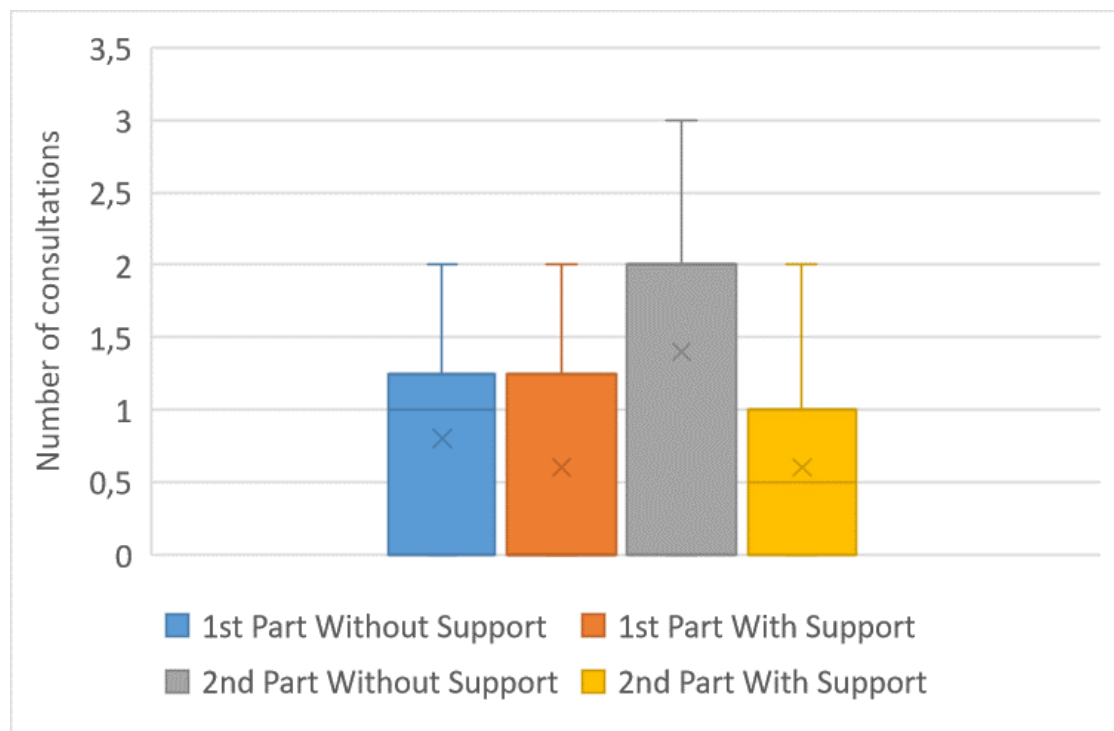


Figure 10. Boxplots of consultations (1st and 2nd parts, with and without VR support)

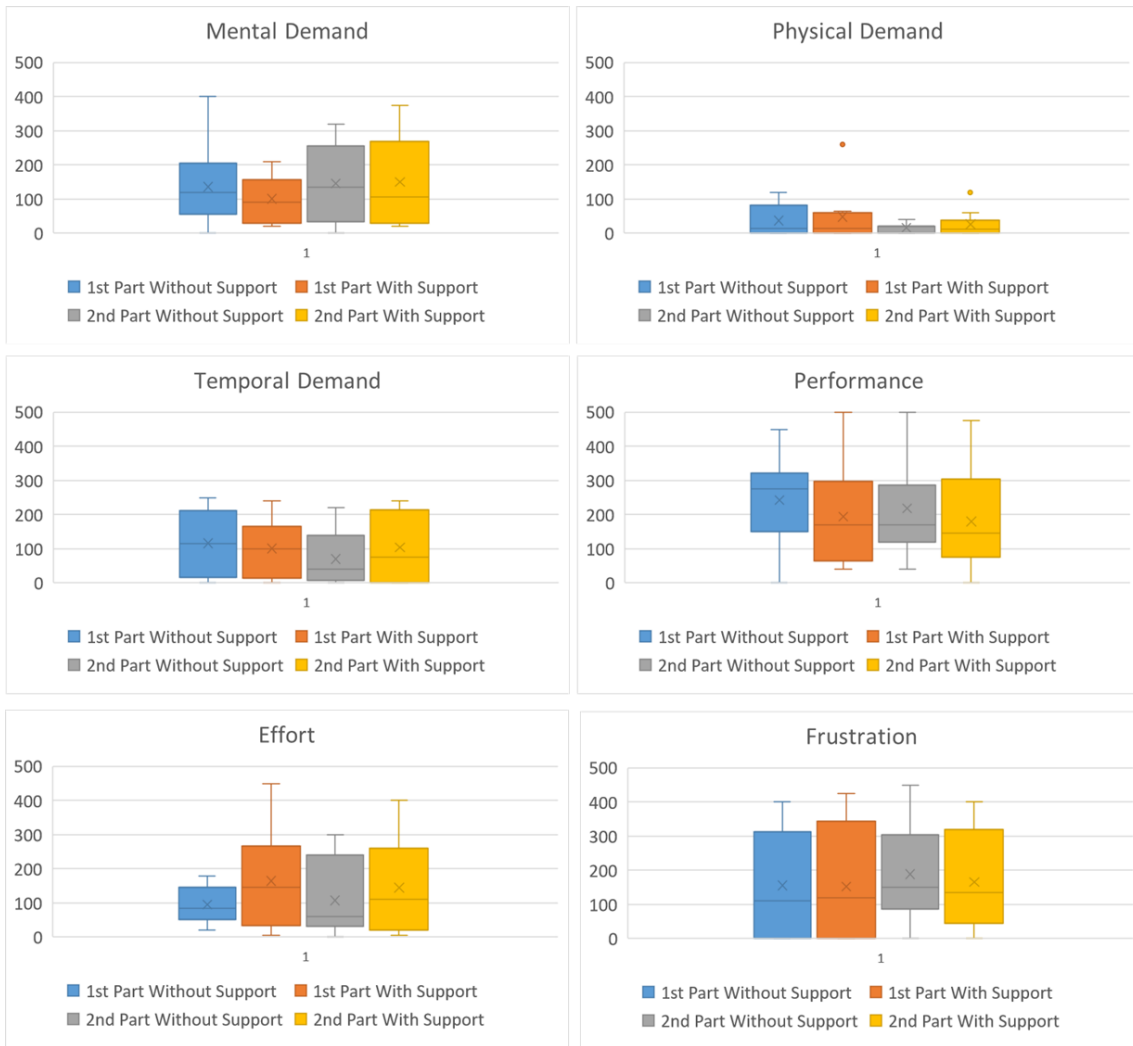


Figure 11. NASA-TLX boxplots for the six domains (500-points scale)

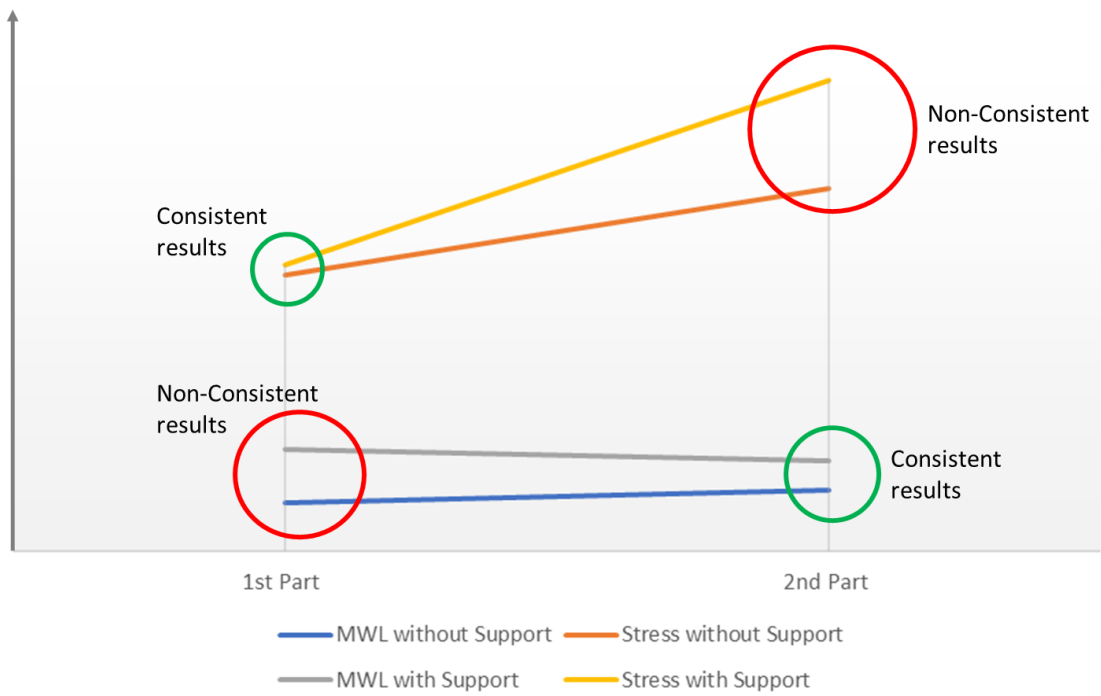


Figure 12. Trends of the computed stress and mental workload from the 1st to the 2nd part; comparison with self-assessment in the circles.