



## OPEN ACCESS

## EDITED BY

Maria Chiara Leva,  
Technological University Dublin, Ireland

## REVIEWED BY

Shijing Liu,  
Philips, United States  
Aleksandar Trifunović,  
University of Belgrade, Serbia

## \*CORRESPONDENCE

Peter Thorvald,  
✉ [peter.thorvald@his.se](mailto:peter.thorvald@his.se)

RECEIVED 04 April 2025

ACCEPTED 26 May 2025

PUBLISHED 05 June 2025

## CITATION

Fogelberg E, Cao H and Thorvald P (2025)  
Cognitive ergonomics: Triangulation of  
physiological, subjective, and performance-  
based mental workload assessments.  
*Front. Ind. Eng.* 3:1605975.  
doi: 10.3389/fieng.2025.1605975

## COPYRIGHT

© 2025 Fogelberg, Cao and Thorvald. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in  
other forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Cognitive ergonomics: Triangulation of physiological, subjective, and performance-based mental workload assessments

Emmie Fogelberg<sup>1</sup>, Huizhong Cao<sup>2</sup> and Peter Thorvald<sup>1\*</sup>

<sup>1</sup>Institute of Engineering, University of Skövde, Skövde, Sweden, <sup>2</sup>Industrial and Materials Science,  
Chalmers University of Technology, Gothenburg, Sweden

**Introduction:** As the manufacturing assembly industry advances, increased customizations and product variety results in operators' executing more cognitively complex tasks. To bridge these cognitive challenges, the assessment of operators' health and performance in relation to their tasks has become an increasingly important topic in the field of cognitive ergonomics.

**Methods:** This paper examines operators' mental workload through an integrated approach by implementing measures covering different mental workload signals: *physiological*, *performance-based*, and *subjective*, while assembling a 3D-printed drone. In this study, four validated mental workload instruments were used and their correlation levels were evaluated: *error rate*, *completion time*, the *Rating Scale Mental Effort* (RSME), and *Heart Rate Variability* (HRV).

**Results:** The results indicate that three out of four mental workload measures significantly correlate and can effectively be used to support the assessment of mental workload. More specifically, error rate, completion time, and RSME.

**Discussion:** Since current literature has stressed the importance of developing a multidimensional mental workload assessment framework, this paper contributes with new findings applicable to the manufacturing assembly industry.

## KEYWORDS

mental workload assessments, cognitive ergonomics, HRV, assembly, industry 5.0

## 1 Introduction

It is no coincidence that both the public and scientific community have become increasingly aware of the extensive impact that mental health has on our work performance. Today, this is considered one of the most challenging predicaments worldwide (Daniel, 2019). National agencies such as the Swedish Work Environment Authority (Arbetsmiljöverket, 2023) and the Swedish Social Insurance Agency (Försäkringskassan, 2023) report that work-related stress causes more than 700 deaths annually, with projections indicating continued escalation. To prevent and bridge this issue, research related to cognitive ergonomics seems to be on the rise and is currently applied to many different types of professions (Zoaktafi et al., 2020).

The manufacturing assembly industry, in particular, faces intensified cognitive demands due to greater product variety and customization. Ultimately, requiring

operators to execute increasingly complex tasks, and thus, increasing their *mental workload* (Bläsing and Bornewasser, 2020; Digiesi et al., 2020). These demands impose cognitive workloads that affect not only operator performance but also wellbeing (Galy, 2018; Muñoz-de-Escalona et al., 2024). At the same time, the adoption of advanced technologies such as automation, AI, and robotics contributes to additional cognitive strain as operators must adapt to hybrid human-technology systems (Longo et al., 2022; Tao et al., 2019).

These developments align with the evolving paradigm of Industry 5.0, where the emphasis shifts from technology-centered automation to more human-centered, resilient, and sustainable manufacturing ecosystems (Antonaci et al., 2024; Trstenjak et al., 2023). Cognitive ergonomics plays a pivotal role in this transition by making sure that operators' cognitive capacities are supported rather than burned out (Baldassarre et al., 2024; Longo et al., 2023). This requires moving beyond traditional workload assessments toward methods that capture real-time, multidimensional, and context-sensitive indicators of mental workload (Jafari et al., 2023; Chiacchio et al., 2023). In particular, triangulating physiological, subjective, and performance-based measures has been identified as a key research priority to better understand and support human operators in Industry 5.0 settings (Fan et al., 2020; Lagomarsino et al., 2022).

Mental workload is the balance between a task's complexity and the subject's cognitive capacity to cope with it. Excessive levels of mental workload arise from performing tasks that are too cognitively demanding for our current resources to handle (Galy et al., 2012). This causes subjects to make more errors, experience delayed information processing, and exhibit slower response time (DiDomenico and Nussbaum, 2011; Kantowitz, 1987; Longo et al., 2022), ultimately, affecting work performance negatively (Lindblom and Thorvald, 2014). Mental *underload* is instead caused by performing monotonous tasks that do not require any (or the absolute minimum) cognitive attention or resources (Basahel et al., 2010).

Despite the recognized need to elevate current assessment of mental workload, many studies continue to rely on isolated workload metrics, limiting their ecological validity and applicability in complex industrial environments (Longo et al., 2022; Van Acker et al., 2018). This paper addresses this gap by examining the triangulation of mental workload using error rate, completion time, RSME, and HRV during an assembly task. By evaluating these measures collectively, this study contributes to the development of practical, multidimensional workload assessment frameworks tailored for human-centered Industry 5.0 manufacturing environments.

The study's twofold objectives are: (i) to investigate the correlation between different mental workload indicators from multiple classes (physiological, performance-based, subjective), and (ii) to explore how these indicators can inform ergonomic interventions aimed at enhancing operator wellbeing and system resilience. Through this, we respond to calls in recent literature emphasizing the integration of cognitive ergonomics into Industry 5.0 strategies (Baldassarre et al., 2024; Longo et al., 2023).

## 2 Mental workload

Mental workload can be understood and interpreted from a number of perspectives. This is primarily because the phenomenon

lacks a standardized definition (Arana-De las Casas et al., 2023; Hertzum and Holmegaard, 2013; Miller, 2001). This paper views mental workload as “the amount of cognitive resources spent on tasks which essentially is determined by competence and demand” (Fogelberg et al., 2025, p. 2). Thus, mental workload is not to be viewed in binary terms; as a simple dichotomy of ‘on’ or ‘off’. Rather, it is more accurately represented as a spectrum, where both excessive and insufficient levels of mental workload negatively affect performance (Rubio et al., 2004; Ryu and Myung, 2005; Young and Stanton, 2002).

Acute stress and mental workload cause activity in the same networks of the brain (Hermans et al., 2014; Van Oort et al., 2017), and the brain responds to stress by inhibiting cognitive functions such as working memory, learning, and concentration (Schoofs et al., 2008); all essential functions for performing most work-related tasks. Although the two are often used indistinctively, it is necessary to separate them (Hidalgo-Muñoz et al., 2018) as acute stress and mental workload can occur independently of one another. Meaning, high mental workload does not automatically elicit acute stress and *vice versa*. Partially because they are thought to be evoked from different mechanisms (emotion vs. cognitive effort) (Parent et al., 2019).

Excessive levels of mental workload have been shown to alter physiological processes, including hormonal regulation. Similar to stress, high mental workload triggers elevated secretion of the hormone cortisol (Nomura et al., 2009; Zeier et al., 1996). In addition, Kusnanto et al. (2020) found that blood sugar increases in correlation with mental workload. Other associated symptoms are poorer body postures (Adams and Nino, 2024), fatigue, drowsiness, (Borghini et al., 2014), and decreased HRV (Yannakakis et al., 2016).

### 2.1 Definition

Adding to the complexity of the phenomenon, a recent review by Longo et al. (2022) identified 68 unique mental workload definitions. As a result, researchers typically focus on and prioritize different aspects of the concept they deem most pivotal. Some underline the importance of individual characteristics (Rasmussen, 1979) like motivation (Fairclough et al., 2019) and competence (Teoh Yi Zhe and Keikhosrokiani, 2021), while others focus on neurological functions (Mandrick et al., 2016), cognitive costs (Andre, 2001), and time (Cain, 2007; Longo et al., 2022).

The multidimensional nature of mental workload presumably complicates the search for a single definition; as more than two decades have passed without reaching a consensus (Jeffri and Rambli, 2021). One of the earliest mental workload depictions is thought to originate from Bornemann (1942) in which optimization of human-machine systems first is introduced (Radüntz, 2017). Modern perspectives phrase it as “... the relationship between demands placed upon individuals and their capacity to cope with it” (Alsuraykh et al., 2019, p. 372), or, “the total cognitive work needed to accomplish a specific task in a finite time period” (Longo et al., 2022, p. 8). No matter the viewpoint researchers adopt, it is crucial that they communicate their interpretation of the phenomenon given the lack of a standardized definition. Not only does this facilitate better comparison of findings but promote clarity and understanding to scholars.

Another terminology that often is used interchangeably with mental workload is *cognitive load*. Some researchers argue they are essentially the same, albeit more prevalent and derived from

different fields (Longo and Orrú, 2022). Cognitive load is predominately featured in educational psychology (Sweller et al., 1998) whereas mental workload has its roots in ergonomics, human factors (MacDonald, 2003), psychology (Hancock et al., 2021), and aviation (Hart, 2006).

## 2.2 Assessment

Mental workload assessments are typically divided into three main classes, performance-based, physiological, and subjective ratings of mental workload (Young et al., 2015). Performance-based techniques quantify how successful subjects are when performing tasks (Butmee et al., 2019). Physiological instruments capture biomarkers of mental workload through nervous, auto, and hormonal regulation (Lean and Shan, 2012). Lastly, self-reports aim to collect subjects' perceived level of mental workload. Most research relies on a single mental workload instrument which ultimately fails to cover the multidimensional nature of the phenomenon (Longo et al., 2022; Tsang and Velazquez, 1996; Van Acker et al., 2018). Over the last 30 years, it has been stressed that employment of all mental workload classes simultaneously is essential (Di Stasi et al., 2009) as instruments fail to register pointers of mental workload beyond their reach (Cegarra and Chevalier, 2008). Assessment (from different classes) should be viewed as complementary elements (Jafari et al., 2020); allowing research to understand mental workload from a greater cognitive context. A more integrated approach, including assessment from all mental workload classes, has been proposed to resolve these challenges (Fan et al., 2020; Lagomarsino et al., 2022; Van Acker et al., 2021).

In combination with the absence of a standardized definition (Kramer, 2020), the pool of methods used to assess mental workload has increased. Not only makes this more difficult for studies to replicate the results, but the findings are often non-uniform due to methodological heterogeneity.

Multiple software and framework solutions have been developed to assess mental workload in a multifaceted way. For example, the Tholos created by Cegarra and Chevalier (2008), Cognitive Load Assessment in Manufacturing (CLAM) (Thorvald et al., 2017; 2019), and the Improved Performance Research Integration Tool (IMPRINT) (Rusnock and Borghetti, 2018). There are nonetheless some challenges posed with the already existing approaches, mainly in terms of intrusiveness, field of work, and mental workload sensitivity. Some of the methods employed in the suggested solutions are not appropriate in naturalistic settings (e.g., eye-tracking and the dual-task paradigm), causing difficulties with applicability. In other cases, triangulation of mental workload metrics has not been sufficiently reached as it does not include physiological parameters. As a final point, some of the tools incorporated are highly complex and are not suitable for the settings of an assembly industry and those who will manage the intricate systems.

## 2.3 Basis for assessment

Method triangulation faces many difficulties. Above all, research must identify an empirically valid triangulation. As there is a wide array of mental workload assessments, research must discover which instruments correlate and complement one another. Not only that,

but determine what instrument (from each of the three classes) best represents and captures the physiological, subjective, and performance-based mental workload signals; suited to the targeted environment. Given the large pool of instruments, this is no straightforward task.

After thoroughly scouring the field of mental workload assessments, we identified four methods that were incorporated into our study. These instruments were suitable to the targeted environment, held empirical weight, and did not require overly complicated technology and software.

### 2.3.1 Performance-based: Error rate and time

When researchers assess mental workload through performance-based methods, two approaches are available: primary and secondary tasks (Laviola et al., 2022). Measures of primary tasks collect relevant information concerned with the original task. Examples of some of the most common ways recordings of subjects' completion time, error rate, and response time (DiDomenico and Nussbaum, 2011; Longo et al., 2022). Research has shown that primary task measures are insensitive. They can only detect mental workload effectively when it is considerably high (Argyle et al., 2021; Bommer and Fendley, 2018). On the contrary, secondary task measures introduce a complementary task (more known as the dual-task paradigm). Researchers evaluate subjects' ability to switch attention and cognitive resources between different tasks. If subjects exhaust all of their resources on one task, they tend to perform poorer on the second task (Laviola et al., 2022). Although this method has demonstrated high mental workload sensitivity (Lagomarsino et al., 2022), it is hard enough to find secondary tasks that are suitable to controlled experiments and not the least to real-world settings.

### 2.3.2 Physiological instrument: HRV

Some of the most commonly reported physiological instruments to assess mental workload are cardiovascular, EEG (Electroencephalography), and eye movement (Tao et al., 2019). Notably, some physiological methods are more complex to execute due to their advanced technology, intricate analysis, and expensive equipment. These are some of the key benefits of HRV which likely is why so many implement it (Fista et al., 2019). The method allows for collecting data in real time (Hoover et al., 2012) and can quite effortlessly be applied in naturalistic settings (Rajcani et al., 2016). The advancements of wearables have made it increasingly easy to measure HRV in many different environments (Ishaque et al., 2021).

There are a couple of different ways to measure HRV, for example, electrocardiogram (ECG) and photoplethysmography (PPG). PPG essentially works by detecting changes in blood volume variation (Fortino and Giampà, 2010). In comparison to HR which is determined by beats per minute (BPM), HRV "is the temporal variation between sequences of consecutive heartbeats" (Karim et al., 2011, p.71).

### 2.3.3 Subjective rating scale: RSME

Among the most frequently implemented mental workload rating scales are the Cooper-Harper Rating Scale (Cooper and Harper, 1969), the NASA-TLX (NASA Task Load Index) (Hart and Staveland, 1988), and the Subjective Workload Assessment

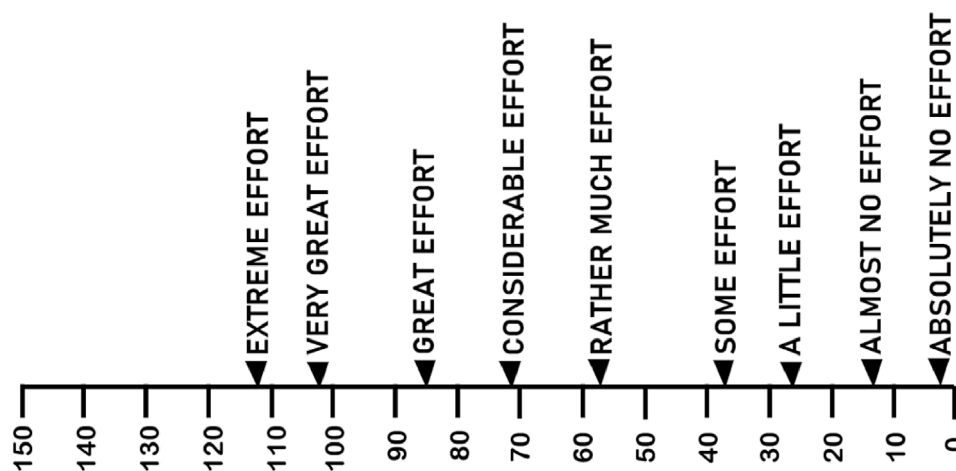


FIGURE 1  
The rating scale mental effort (RSME) by Zijlstra (1993).

Technique (SWAT) (Reid and Nygren, 1988). All listed scales take measurably more time to complete in comparison to the RSME. The RSME (illustrated in Figure 1) developed by Zijlstra (1993) is noticeably short which permits the questionnaire to interfere with the task as little as possible. Although the NASA-TLX likewise is relatively short, it has been shown to correlate with the RSME on multiple occasions (Ghanbary Sartang et al., 2016; Hoonakker et al., 2011; Rubio et al., 2004). It takes subjects around 1 minute to complete the RSME, whereas the NASA-TLX takes five to 10 minutes. More so, the RSME has shown good mental workload sensitivity (Herlambang et al., 2021; Longo and Orrú, 2022).

### 3 Methodology

As stated prior, the current study had two main purposes i) to examine VR, video, and robot as instructional methods, and ii) to correlate the different mental workload instruments HRV, time, error, and RSME. Based on the ii) purpose, one main hypothesis has been formulated and is tried accordingly.

**Hypothesis 1:** There is a significant correlation between the mental workload measurements error rate, completion time, the RSME, and HRV. As one measure changes, the others will reflect corresponding changes, though not necessarily linear, whilst performing an assembly task.

#### 3.1 Procedure

After the subjects had agreed and filled out the informed ethical consent, their baseline HRV levels were collected. Subjects were then randomly assigned to either one of the instructional method conditions (video, VR, and robot), creating a between-subject design. To assess performance, the experiment was divided into two central phases: *learning* and *operational* as can be observed in Figure 2.

During the learning phase, participants were asked to assemble a 3D-printed model of a drone while the researchers concurrently emphasized the importance of trying to remember the steps and the placement of all parts, see Figure 3. While assembling, subjects wore an ear clip sensor which collected their HRV. The subjects were allowed to spend as much time as they wanted in the learning phase and once the subjects stated they were finished, the RSME was filled out. Subjects rate their perceived level of mental workload on a scale from 0 to 150 (as can be seen in Figure 1). Immediately after, in the operational phase, participants were again asked to build a 3D-printed drone but without any instructions. Throughout the assembly, subjects' HRV was again recorded. Upon completion, subjects instantly filled out the RSME while the researchers composed the completion time and counted the number of errors.

#### 3.2 Conditions

In the video condition, subjects looked at a looped video throughout their drone assembly as shown in Figure 4. The video was separated into smaller segments showing a few seconds for one step in the assembly. Subjects had the ability to go back and forward between video fragments if they wanted to look more carefully.

In the VR condition, the entire learning phase took place in a virtual setup of the drone assembly, very similar to a real-world environment. Subjects were guided through 'ghost actions' where the VR instructed subjects how to assemble the drone by first showing what part goes where and then inviting the subject to follow. Subjects could not disassemble the drone once they were finished. They could nonetheless lift the drone and look more carefully at where all parts were located.

In the final condition, the conversational robot gave subjects instructions through automated AI, as shown in Figure 4. All of the drone parts that the conversational robot referred to, were listed and labeled on a screen in front of the subjects. The subjects could ask the conversational robot to specify and alter its instructions to a certain degree if they did not understand what was expected of them.

## EXPERIMENT PROCEDURE

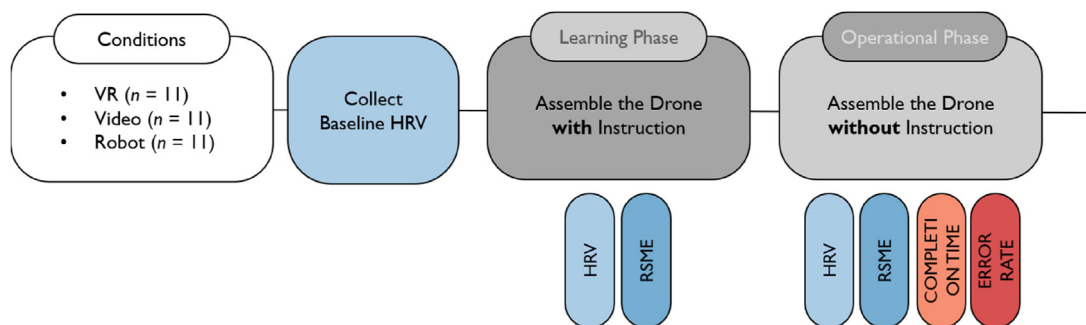


FIGURE 2  
Outline of the experiment procedure for the 4 × 3 factorial between-subject design.

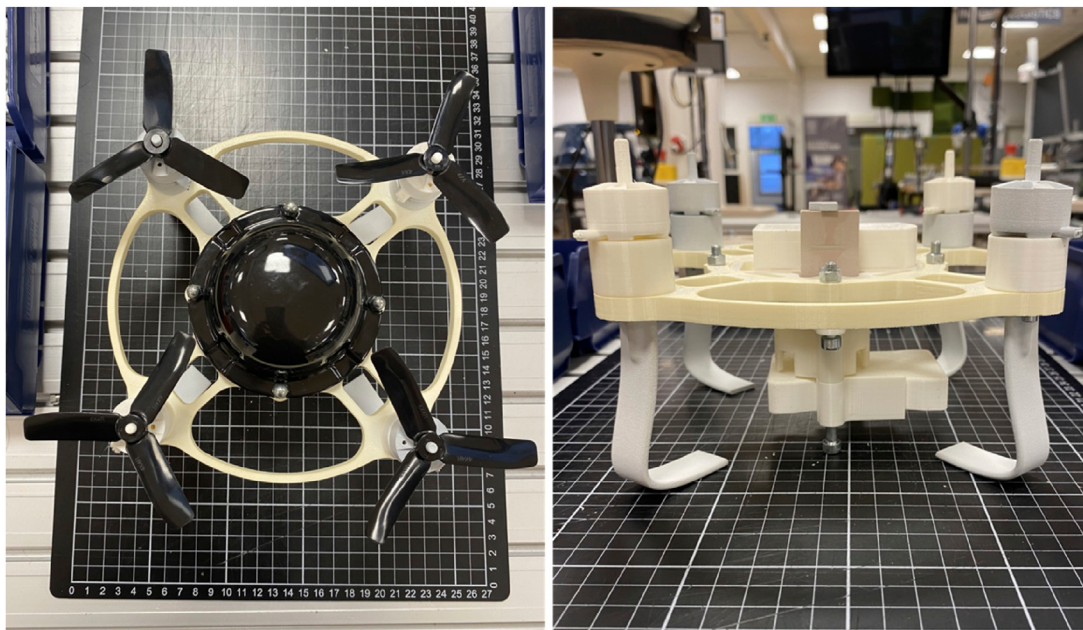


FIGURE 3  
The 3D-printed drone subjects assembled in the experiment. Here shown with the cover on, and the parts underneath the cover.

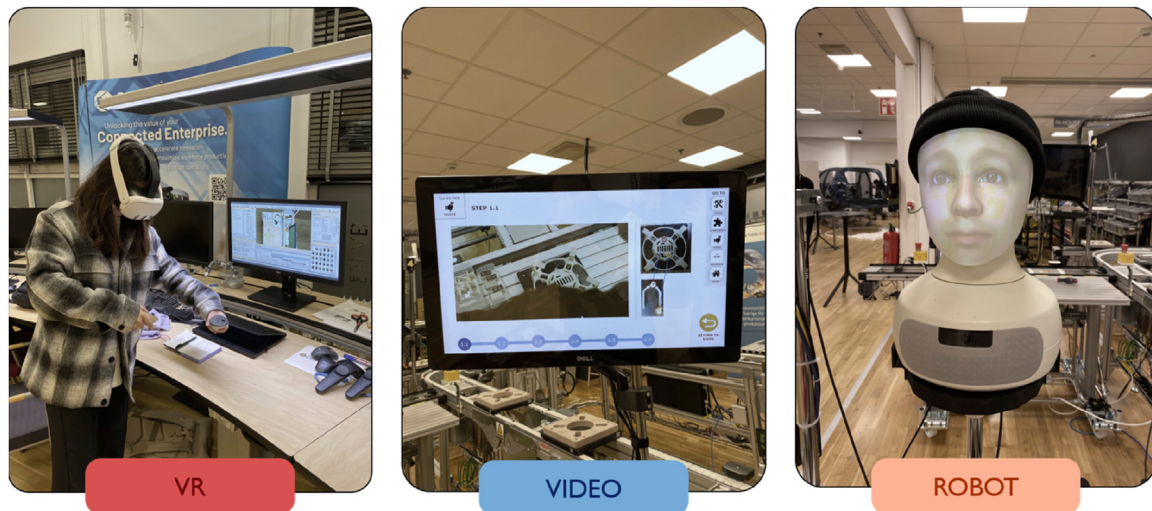
### 3.3 Equipment

The equipment used to collect subjects' HRV featured a PPG sensor clipped to their ears. The sensor was powered by an 'Arduino 33 BLE' microcontroller and connected to an external battery source (e.g., a computer or power bank) as shown in Figure 5. The Arduino transmits raw data in real-time to a computer via Bluetooth, consisting of digitized heart beat signals in time domain. This is used to calculate peak-to-peak interval and HRV directly. While acknowledging ongoing debates concerning its equivalence to ECG-derived HRV (Georgiou et al., 2018; Ishaque et al., 2021), PPG was preferred for many reasons; non-invasive design, subject comfort, portability, affordability, and suitability for rapid data acquisition in operational phases under

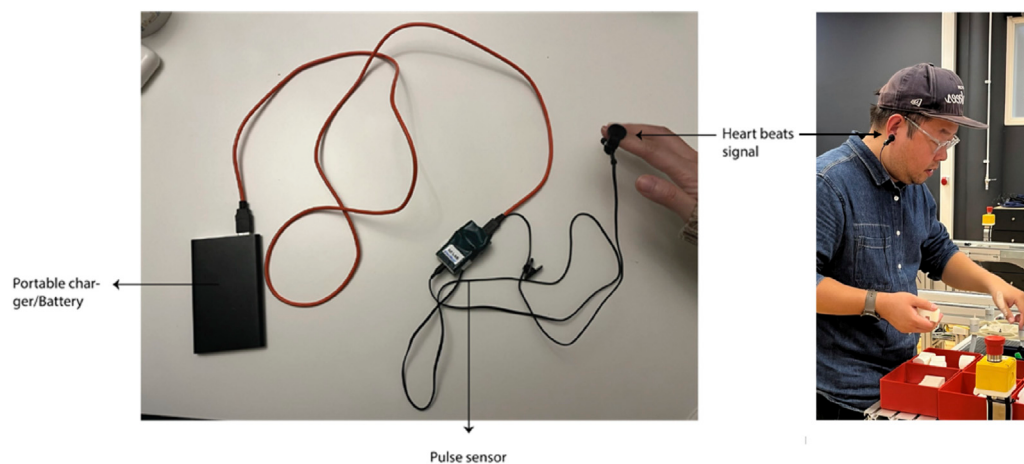
3 minutes. Traditional ECG placement (electrodes attached directly on the ribcage), was determined impractical considering the study's targeted population and settings. In real-world industrial environments, it is essential that the HRV instrument is easy to use, both for the operator and the researcher. More so, that the hands are kept free from sensors during assembly since this can disturb the HRV signal.

### 3.4 Signal acquisition and preprocessing

Following recommendations by Tarvainen et al. (2013), three critical pre-processing steps were applied to ensure robust HRV estimation.



**FIGURE 4**  
The three types of instructional methods used in the current experiment, VR, video, and robot.



**FIGURE 5**  
HRV sensor employed in the experiment.

#### i. Noise detection

Signal segments contaminated by motion artefacts (e.g., abrupt amplitude deviations  $>20\%$ , sudden spikes, or signal drops) were systematically flagged and excluded from analysis. This step minimized the influence of physiologically implausible data caused by sensor displacement or participant movement.

#### ii. Detrending considerations

Detrending process to remove slow, non-physiological variations in heart rate, was not applied in this study. Given to the short recording windows (under 5 min), long-term trends in

heart rate were negligible, aligning with [Tarvainen et al. \(2013\)](#), who suggest that detrending is less critical for ultra-short-term HRV analysis.

#### iii. Threshold-based beat correction

- Local baseline: A median-filtered moving window (duration: 10 s) generated a local average of RR intervals.
- Outlier removal: Individual RR intervals deviating from this baseline by  $> 400$  m (predefined threshold) were classified as outliers and excluded.
- Adaptive adjustment: The 400 m threshold was dynamically refined for each subject based on their mean heart rate, following best practices outlined in Kubios HRV's pre-processing guidelines ([Kubios, 2024](#)).

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N-1} (RR_{n+1} - RR_n)^2}$$

FIGURE 6  
Equation of the root mean square of successive RR interval differences (Malik et al., 1996).

### 3.5 HRV calculation

The pre-processed data was used to compute the root mean square of successive differences (RMSSD), a time-domain metric commonly employed for short-term PRV assessment (Bourdillon et al., 2022). All data and plots were automatically saved in a timestamped Python Jupyter notebook within a local “results” folder to ensure reproducibility. The root mean square of successive RR interval differences, or RMSSD, is given in Figure 6.

### 3.6 Participants

All participants were invited to take part in the experiment either through course participation at Chalmers or by being approached while on-site at campus on Chalmers Lindholmen in Sweden.

**Socioeconomic variables.** The experiment included a total sample size of  $N = 37$  participants. Two of them were part of the pilot study which resulted in minor changes, and therefore, they were not added to the final sample. Additionally, we were unable to collect complete and valid HRV recordings for five subjects. In the end, the experiment consisted of  $N = 33$  participants of which 11 were female and 22 males. Subjects' ages ranged from 18 to 42, with a mean age of 26 years. Although most reported Swedish as their native language (66.67%), others listed were Malayalam, Chinese, Portuguese, Telugu, Sinhalese, German, Arabic, Tamil, and Urdu. The majority of subjects stated high school as the highest level of education (36.67%), followed by a master's degree (33.33%), bachelor's (26.67%), and PhD (3.33%).

All subjects were required to read and fill out an informed consent. The document contained information about the experiment's purpose and all the instructional methods that subjects could be randomly assigned to. It further specified that participation is entirely voluntary from start to finish. More so, subjects are anonymous throughout, and all information is handled to not be traceable. To ensure safety, three disclaimers are listed in the consent: no severe visual or hearing impairments, no medical conditions posing a health risk, and at least 18 years old. In the end, we collected subjects' socioeconomic information and their level of experience with virtual reality (VR) and conversational robots.

## 4 Analysis

Although the instructional methods employed in the experiment (video, conversational robot, and VR) will not be expressively

discussed and stressed in this analysis, separation of the conditions in some analyses is still necessary to successfully triangulate the different measurements of mental workload (error rate, time, RSME, and HRV). The present paper's analyses are only interested in the operational phase since error rate, completion time, and RSME were collected during this stage. HRV and RSME were exclusively collected in both the learning and operational phases to assess the instructional methods.

We performed both descriptive and inferential statistics to summarize and examine the given data. Some of the subjects' HRV data was not used in the analyses ( $n = 6$ ) as it showed obvious outliers due to the sensor not sitting properly, subjects touching the sensor, etc. However, the same subjects were free from interferences in the other mental workload assessments. Hence, the sample size varies across different analyses in both descriptive and inferential statistics.

### 4.1 Descriptive statistics

The illustrations in Figure 7 show how the VR condition ( $n = 11$ ) generated longer completion times ( $\bar{x} = 08:18$  min) than both the robot ( $n = 11$ ,  $\bar{x} = 03:53$  min) and the video condition ( $n = 11$ ,  $\bar{x} = 03:30$  min). The same trend is observed in the number of errors made when VR ( $\bar{x} = 6$ ) is compared to the robot ( $\bar{x} = 1$ ) and video condition ( $\bar{x} = 0$ ).

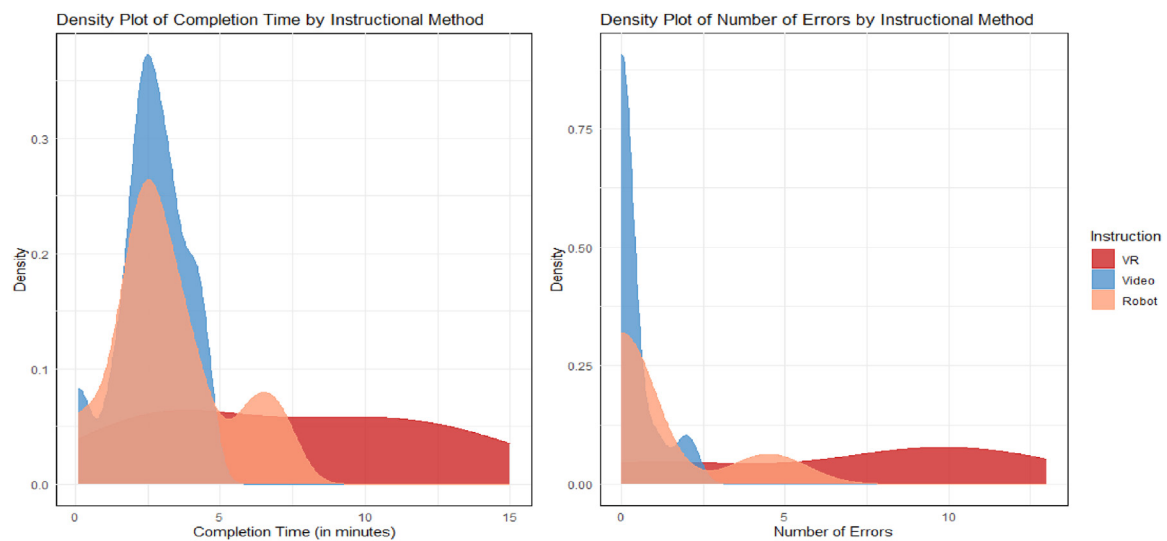
Figure 8 illustrates how the highest RSME ratings likewise came from the VR condition ( $\bar{x} = 71$ ), followed by the robot ( $\bar{x} = 46$ ) and video condition ( $\bar{x} = 25$ ). Purely based on the initial descriptive statistics it appears that performance-based and subjective measures of mental workload align during the operational phase. This, as the VR condition generated subjects the most mental workload in terms of both completion time, error rate, and RSME, and the same order is followed in the robot and video condition.

Lastly, a density plot of the subjects' average HRV during the operational phase suggests that the robot condition showed the highest average HRV followed by the video and VR conditions. It is important to note that a low HRV indicates a high mental workload. Therefore, subjects in the VR condition again appeared to experience the highest mental workload but the robot and video group switched places in comparison to the previous measures as shown in Figure 9.

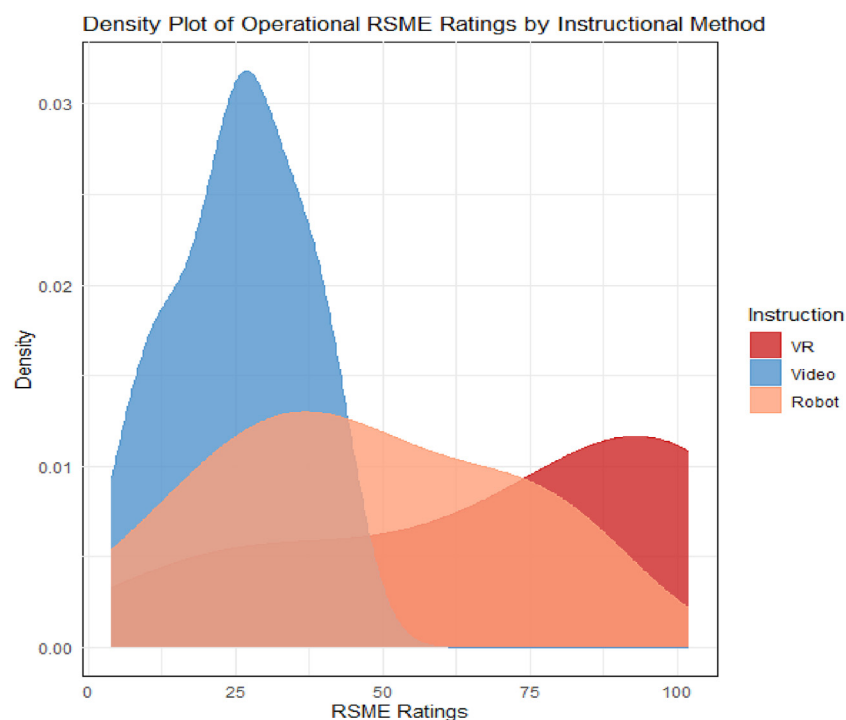
### 4.2 Inferential statistics

Multiple Pearson correlation analyses were conducted to examine the level of association between the different mental workload metrics. To illustrate the results, a correlation matrix was created containing all of the analyses made as can be observed in Figure 10.

A Pearson correlation analysis between RSME ratings and completion time during the operational phase ( $N = 33$ ) showed a strong positive correlation  $r(31) = 0.70$ ,  $p < 0.001$ , 95% CI [0.47, 0.84]. Indicating that higher RSME scores are associated with longer completion times. RSME and error rate also revealed a moderate positive correlation  $r(31) = 0.51$ ,  $p < 0.002$ , 95% CI [0.20, 0.72]. Higher RSME scores also appear to be associated with higher error



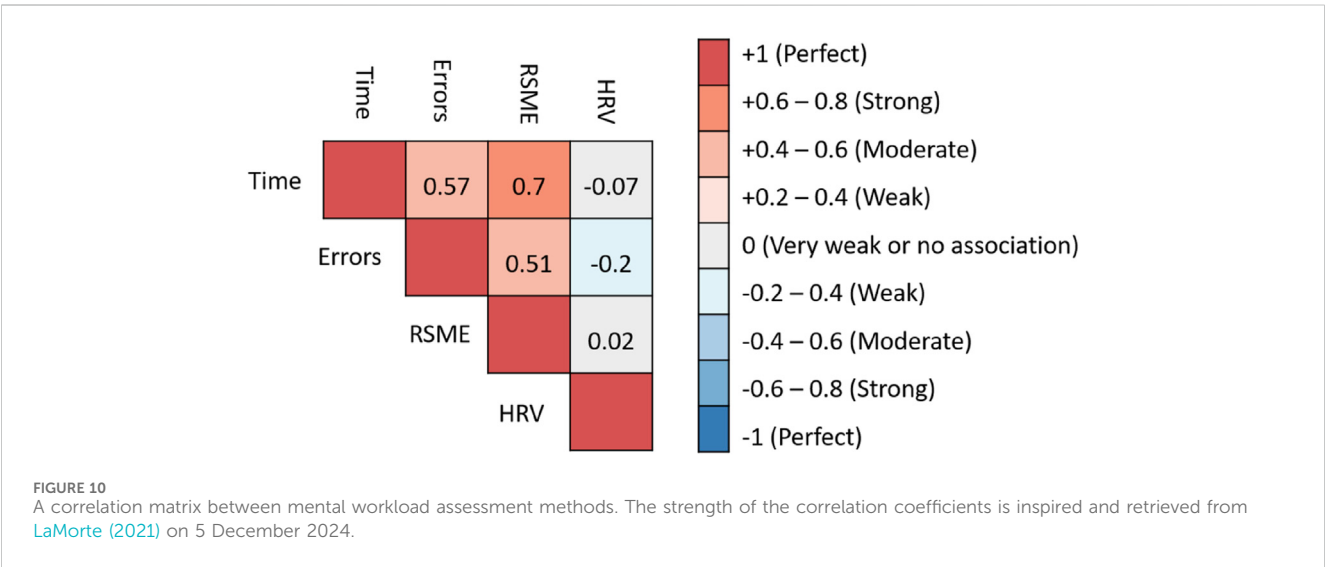
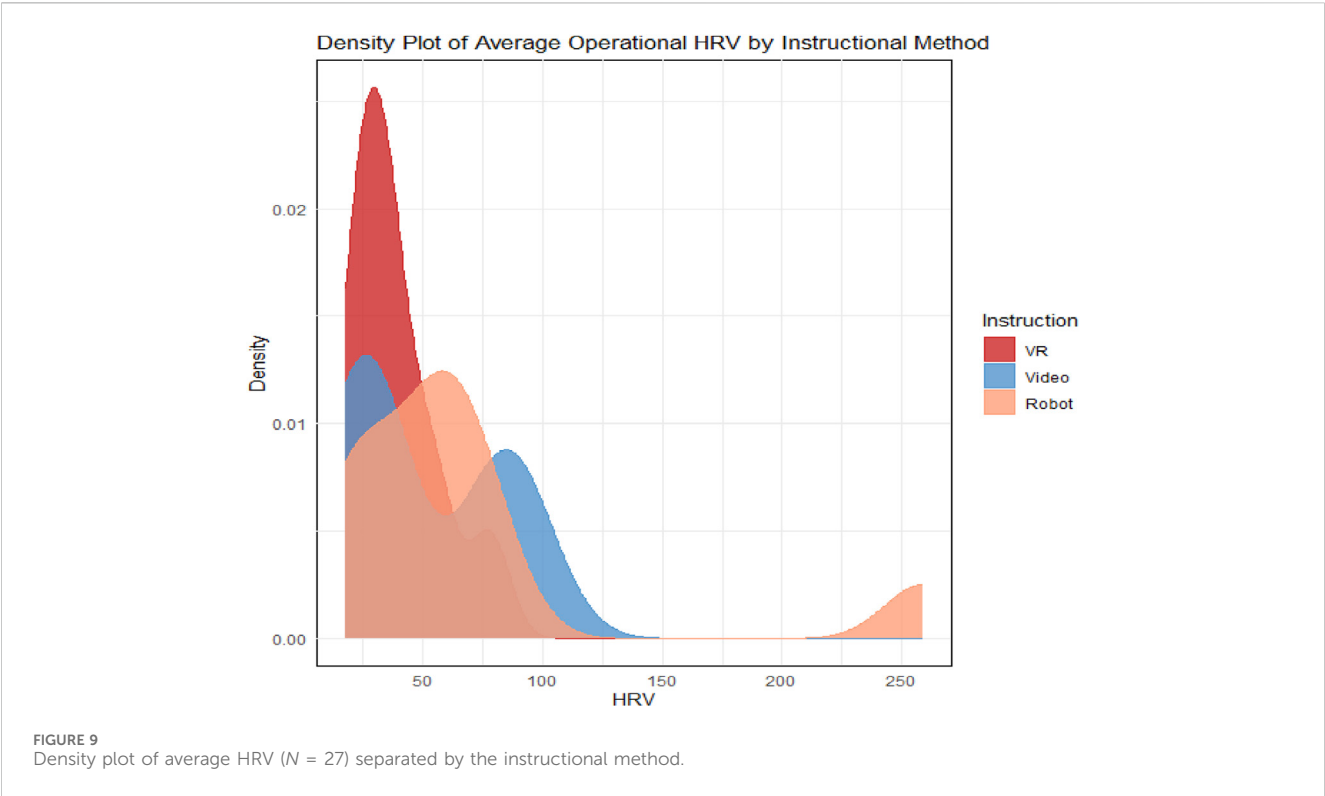
**FIGURE 7**  
To the left: density plot of completion times separated by instructional method ( $N = 33$ ). To the right: density plot of error rate separated by instructional method ( $N = 33$ ).



**FIGURE 8**  
Density plot of RSME ratings ( $N = 33$ ) separated by the instructional method.

rates. An additional Pearson correlation between error rate and completion time also demonstrated a moderate positive correlation  $r(31) = 0.57$ ,  $p < 0.001$ , 95% CI [0.28, 0.76]. This suggests that higher error rates are associated with longer completion times. These results align well with the illustrations presented in the descriptive statistics.

Pearson correlation analyses with a smaller sample size ( $N = 27$ ) were conducted to assess the relationship between average HRV and the other mental workload metrics. Between HRV and error rate, the correlation coefficient was found to be  $(r = -0.2)$ , indicating a weak relationship. The same trend was found for HRV and time ( $r = 0.07$ ), and RSME ( $r = 0.012$ ). Thus, demonstrating a very weak correlation



as illustrated in Figure 10. Based on the inferential statistics, *a priori* power analysis was conducted which revealed a power of 0.283 and a small effect size of  $d = 0.344$ . Normally,  $(1-\beta)$  of 0.80 statistical power and  $\alpha = 0.05$ , or greater, is generally the recommended power level (Rawson et al., 2008).

### 5 Discussion

This study has examined four mental workload assessment methods (HRV, error rate, completion time, and RSME) while

subjects assembled a 3D-printed drone with one out of the three instructional methods (VR, video, or robot). Two main objectives were explored, the first was to scrutinize the designated instructional methods, with the findings disclosed and discussed in the paper by Cao et al. (2024). The second, and principal focus of this paper, is to investigate the level of correlation between physiological (HRV), performance-based (error rate and completion time), and subjective (RSME) instruments as mental workload increases. We found evidence that some of the measurements significantly correlate. More specifically, completion time, error rate, and subjects' RSME ratings. The

results indicate that these mental workload instruments can be successfully employed simultaneously while picking up different signals of elevated mental workload levels.

The results make evident that a high mental workload both led to longer completion times and more errors made. This finding is supported by prior research, suggesting that an elevated mental workload causes subjects to perform poorer because their cognitive abilities (such as focus) become more strained (Bläsing and Bornewasser, 2020; DiDomenico and Nussbaum, 2011; Longo et al., 2022). The correlation between error rate and completion time (in many cases) seems to exhibit a speed-accuracy trade-off (Standage et al., 2014) that could not be confirmed here. This trade-off means that there is a conscious or unconscious compromise between being fast or accurate. Shorter completion times usually result in more errors made, while longer completion times are the result of fewer errors.

In this study, elevated mental workload influenced both time and accuracy negatively. A possible reason for the absence of a speed-accuracy trade-off is that subjects were not deliberately instructed to prioritize one metric over the other. Rather, they were instructed from the beginning that subjects' performance would be based on time and accuracy. Nonetheless, the correlation between time and accuracy was only found moderate. This could be due to the great spread in both speed and accuracy among subjects, as presented in Figure 7. Despite being moderate, it indicates that the performance-based measures applied are sensitive to mental workload changes to some extent, and can empirically be used together.

The results further revealed that the subjective mental workload measure (RSME) significantly correlated with both of the performance-based instruments. As subjects reported higher RSME scores, they also exhibited longer completion times and made more errors. Stressing that high mental workload shows up in both subjective and performance-based data collections. This is also coherent with earlier research proposing that mental workload is a multidimensional phenomenon triggering different types of signals (Van Acker et al., 2018). This finding supports the concern that solely relying on a single mental workload instrument may provide the researcher with insufficient data on the phenomenon; which earlier has been raised by, for example, Jafari et al. (2020) who state that the complexity and multidimensionality of mental workload necessitates multiple assessment methods from different classes.

Interestingly, the only mental workload measure that showed no correlation with the other variables in the study was HRV. Previous studies have shown that HRV is sensitive to mental workload changes (Grandi et al., 2022; Matthews et al., 2015). However, the explicit correlation between HRV, completion time, error rate, and RSME has not been extensively studied. There are many potential explanations for this. Given that the sample size was relatively small ( $N = 27$ ), we performed a power analysis (0.283) which revealed that there is a relatively low probability for the current study to find a significant effect for HRV—if we assume a power of  $(1-\beta)$  of 0.80 and  $\alpha = 0.05$  (Rawson et al., 2008). Meaning that there is a noticeable risk for the study to conduct a Type II error. More so, the lowest HRV average in the operational phase was 18.18 (indicating the most mental workload) whereas the highest HRV average was documented to 258.54. The individual differences might

have been too large for the small sample, consequently diluting group-level correlations.

In the second paper of the study (Cao et al., 2024), a significant difference in HRV across instructional methods in the *learning phase* was found. Although the learning phase has not been deliberately discussed here, the findings could provide some potential explanations to why HRV showed no correlation with the other measures during the operational phase. As explained in the procedure, HRV and RSME were the only mental workload instruments that were collected during the learning phase. Arguably, it could be that subjects had enough time to regulate their physiological responses between phases which is why a more leveled HRV is observed in the operational phase.

Another potential explanation is the drone task itself. It is possible that the assembly task was too short, or not intensive enough, to allow subjects' bodies time to respond physiologically. Almost 60% of the sample ( $N = 33$ ) completed the task in under 4 minutes. Other research that has incorporated HRV as a mental workload measure has had subjects perform the task for a minimum of five to seven (Grandi et al., 2022), 20 (Alhaag et al., 2021), and even 30 min (Ma et al., 2024).

Other physiological metrics have earlier been found to statistically correlate with subjective and performance-based mental workload assessments in research. Different forms of eye tracking such as fixation duration (Bommer and Fendley, 2018) and pupil diameter (Grandi et al., 2022). However, these are not suitable to naturalistic settings, nor the assembly industry where light sources often are insufficient in terms of intensity and quantity. To record proper eye-tracking readings appropriate and controlled lighting is required (Jongerius et al., 2021). It is crucial that the mental workload instruments employed are suitable to the targeted environment and population. In our focus group and environment, instruments that are not feasible in real-world situations have not been prioritized.

## 6 Conclusion

The objective to triangulate all of the mental workload classes (physiological, subjective, performance-based) and thus, create a multifaceted framework targeted toward the assembly industry, takes one step forward by showing that as mental workload increases, RSME, error rate, and completion time statistically significantly correlate. Previous literature has stressed that a multi-dimensional tool to detect multiple mental workload signals is needed to cover the complexity of the phenomenon. Findings from the current study are beneficial for future research and implementation of different mental workload measures. The key findings of the study are the following.

- Higher RSME ratings are significantly correlated with increased error rates and longer completion times.
- Completion time and error rate demonstrate a moderate correlation, proposing that heightened mental workload has a negative impact on *both* speed and accuracy.
- While HRV has been demonstrated to be sensitive to mental workload changes in previous research, it did not statistically correlate with any of the other measures in the current study.

With respect to the small power value (0.283) and effect size ( $d = 0.344$ ), it also appears pivotal to consider the task duration when collecting physiological signals.

## 7 Limitations

The most evident limitation that this study holds is the relatively small sample size. Naturally, with more participants, the potential to find significant results increases as was depicted in the power analysis. Provided with a larger sample, it is feasible that a correlation between HRV and the other mental workload instruments could have been established. Especially considering since metrics like HRV are prone to exhibit substantial individual differences. Thus, the current study is potentially subject to conducting a Type II error as a result of the limited sample size. Another limitation pertaining to the HRV measures is the short task duration. It is probable that the subjects were too quick assembling the 3D drone, and hence, the task lacking enough time to provoke any physiological changes. Consequently, failing to triangulate HRV with any of the other mental workload instruments. Future studies ought to take note of this and consider extending the task. A further limitation is the general validity. Seeing that mental workload is a subjective experience, all types of physiological and performance-based measures introduce validity concerns. This, as the phenomenon is not directly observable, nor measurable for researchers. Lastly, some of the factors display relatively high variation which raises some concern in regards to the reliability of the instruments employed. This study opted for a PPG sensor owing to its non-intrusiveness and easy implementation to real-world environments, however, it was not without fault as six subjects' HRV data were incomplete. The sensor is sensitive to accidental contact, and although being instructed otherwise, some of the subjects unintentionally touched it which led to disturbances during the data collection. Future studies are recommended to investigate if other reliable HRV instruments, like ECG, can be successfully used on subjects without being intrusive (i.e., electrodes placed on the chest area) or disturb the task (i.e., electrodes placed fingers and hands).

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The requirement of ethical approval was waived by Chalmers Institutional Ethics Advisory Board (IEAB) for the studies involving

humans because Chalmers Institutional Ethics Advisory Board (IEAB). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

EF: Formal Analysis, Writing – original draft, Data curation, Writing – review and editing, Resources, Conceptualization, Visualization, Investigation, Methodology, Validation, Project administration. HC: Data curation, Writing – original draft, Software, Methodology, Resources, Project administration. PT: Supervision, Writing – review and editing, Funding acquisition, Validation.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was carried out in the DIGITALIS (DIGITAL work instructions for cognitive work) project, funded by Swedish innovation agency Vinnova through their strategic innovation program, Produktion 2030.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Adams, R., and Nino, V. (2024). Work-related psychosocial factors and their effects on mental workload perception and body postures. *Int. J. Environ. Res. public health* 21 (7), 876. doi:10.3390/ijerph21070876
- Alhaag, M. H., Ghaleb, A. M., Mansour, L., and Ramadan, M. Z. (2021). Investigating the immediate influence of moderate pedal exercises during an assembly work on performance and workload in healthy men. *Healthcare* 9 (12), 1644. doi:10.3390/healthcare9121644
- Alsurraykh, N. H., Wilson, M. L., Tennent, P., and Sharples, S. (2019). "How stress and mental workload are connected. *Proc. 13th EAI Int. Conf. Pervasive Comput. Technol. Healthc.* 371–376. doi:10.1145/3329189.3329235

- Andre, A. D. (2001). "The value of workload in the design and analysis of consumer products," in *Stress, workload, and fatigue*. Editors P. A. Hancock, and P. A. Desmond (Mahwah, NJ: L. Erlbaum), 373–382.
- Antonaci, F. G., Olivetti, E. C., Marcolin, F., Castiblanco Jimenez, I. A., Eynard, B., Vezzetti, E., et al. (2024). Workplace well-being in industry 5.0: a worker-centered systematic review. *Sensors* 24 (17), 5473. doi:10.3390/s24175473
- Arana-De las Casas, N. I., De la Riva-Rodriguez, J., Maldonado-Macias, A. A., and Sáenz-Zamarrón, D. (2023). Cognitive analyses for interface design using dual N-back tasks for mental workload (MWL) evaluation. *Int. J. Environ. Res. Public Health* 20 (2), 1184. doi:10.3390/ijerph20021184
- Arbetsmiljöverket (2023). Ohälsosam arbetsbelastning vanligare orsak till dödlighet. *Arbetsmiljöverket*. Available online at: <https://www.av.se/press/ohalsosam-arbetsbelastning-vanligare-orsak-till-dodlighet/> [press release].
- Argyle, E. M., Marinescu, A., Wilson, M. L., Lawson, G., and Sharples, S. (2021). Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments. *Int. J. Human-Computer Stud.* 145, 102522. doi:10.1016/j.ijhcs.2020.102522
- Baldassarre, M. T., Chiacchio, F., and Lambiase, A. (2024). Trust, cognitive load, and adaptive operator support in Industry 5.0: a review of challenges and solutions. *Hum. Factors Ergon. Manuf. and Serv. Ind., early online*. doi:10.1002/hfm.21015
- Basahel, A. M., Young, M. S., and Ajovalasit, M. (2010). "Impacts of physical and mental workload interaction on human attentional resources performance," in Proceedings of the 28th Annual European Conference on Cognitive Ergonomics, 215–217.
- Blasing, D., and Bornewasser, M. (2020). "Influence of complexity and noise on mental workload during a manual assembly task," in Human Mental Workload: Models and Applications: 4th International Symposium, H-WORKLOAD 2020, Granada, Spain, December 3–5, 2020, Proceedings 4 (Springer International Publishing), 147–174.
- Bommer, S. C., and Fendley, M. (2018). A theoretical framework for evaluating mental workload resources in human systems design for manufacturing operations. *Int. J. Ind. Ergon.* 63, 7–17. doi:10.1016/j.ergon.2016.10.007
- Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. and Biobehav. Rev.* 44, 58–75. doi:10.1016/j.neubiorev.2012.10.003
- Bornemann, E. (1942). Untersuchungen über den Grad der geistigen Beanspruchung. 2. Teil: Praktische Ergebnisse. *Arbeitsphysiologie* 12, 173–191. doi:10.1007/BF02605156
- Bourdillon, N., Yazdani, S., Vesin, J., Schmitt, L., and Millet, G. P. (2022). RMSSD is more sensitive to artifacts than Frequency-Domain parameters: implication in athletes' monitoring. *J. Sports Sci. and Med.* 21, 260–266. doi:10.52082/jssm.2022.260
- Butmee, T., Lansdown, T. C., and Walker, G. H. (2019). "Mental workload and performance measurements in driving task: a review literature," in Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018) Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics, March 2022 (Springer International Publishing), 286–294.
- Cain, B. (2007). *A review of the mental workload literature*. Toronto, Canada: Defence Research and Development, 4–32.
- Cao, H., Fogelberg, E., Thorvald, P., Tengelin, E., Salunkhe, O., and Stahre, J. (2024). A study on assessing mental workload of the digitalization methods in learning and operation phase: VR training, social robot, and video.
- Cegarra, J., and Chevalier, A. (2008). The use of Tholos software for combining measures of mental workload: toward theoretical and methodological improvements. *Behav. Res. Methods* 40 (4), 988–1000. doi:10.3758/BRM.40.4.988
- Chiacchio, F., Rotunno, P., and Lambiase, A. (2023). Human-centric digital twins for cognitive support in smart manufacturing. *J. Ind. Inf. Integr.* n 39, 100456. doi:10.1016/j.jii.2023.100456
- Cooper, G. E., and Harper, R. P., Jr. (1969). The use of pilot rating in the evaluation of aircraft handling qualities (NASA Technical Note D-5153). *National Aeronautics and Space Administration, Ames Research Center*. Available online at: <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19690013177.pdf>
- Daniel, C. O. (2019). Effects of job stress on employee's performance. *Int. J. Bus. Manag. Soc. Res.* 6 (2), 375–382. doi:10.18801/ijbmsr.060219.40
- DiDomenico, A., and Nussbaum, M. A. (2011). Effects of different physical workload parameters on mental workload and performance. *Int. J. Ind. Ergon.* 41 (3), 255–260. doi:10.1016/j.ergon.2011.01.008
- Digiesi, S., Manghisi, V. M., Facchini, F., Klose, E. M., Foglia, M. M., and Mummolo, C. (2020). Heart rate variability based assessment of cognitive workload in smart operators. *Manag. Prod. Eng. Rev.* 11 (3), 56–64. doi:10.24425/MPER.2020.134932
- Di Stasi, L. L., Álvarez-Valbuena, V., Cañas, J. J., Maldonado, A., Catena, A., Antolí, A., et al. (2009). Risk behaviour and mental workload: multimodal assessment techniques applied to motorbike riding simulation. *Transp. Res. part F traffic Psychol. Behav.* 12 (5), 361–370. doi:10.1016/j.trf.2009.02.004
- Fairclough, S., Ewing, K., Burns, C., and Kreplin, U. (2019). "Neural efficiency and mental workload: locating the red line," in *Neuroergonomics* (Academic Press), 73–77. doi:10.1016/B978-0-12-811926-6.00012-9
- Fan, X., Zhao, C., Zhang, X., Luo, H., and Zhang, W. (2020). Assessment of mental workload based on multi-physiological signals. *Technol. health care Official J. Eur. Soc. Eng. Med.* 28 (S1), 67–80. doi:10.3233/THC209008
- Fista, B., Aziz, H. A., Aprilya, T., Saidatul, S., Sinaga, M. K., Pratama, J., et al. (2019). Review of cognitive ergonomic measurement tools. *IOP Conf. Ser. Mater. Sci. Eng.* 598 (1), 012131. doi:10.1088/1757-899X/598/1/012131
- Fogelberg, E., Thorvald, P., and Kolbeinsson, A. (2024). Mental workload assessments in the assembly industry and the way forward: a literature review. *Int. J. Hum. Factors Ergon.* 11, 412–438. doi:10.1504/ijhfe.2024.144213
- Försäkringskassan (2023). *Psykisk ohälsa i dagens arbetsliv: Lägesrapport 2024. Försäkringskassan. Försäkringskassans lägesrapport – sjuk av stress: kvinnor löper dubbel risk*.
- Fortino, G., and Giampà, V. (2010). "PPG-based methods for non invasive and continuous blood pressure measurement: an overview and development issues in body sensor networks," in 2010 IEEE International Workshop on Medical Measurements and Applications (IEEE), 10–13.
- Galy, E. (2018). Consideration of several mental workload categories: perspectives for elaboration of new ergonomic recommendations concerning shiftwork. *Theor. issues Ergon. Sci.* 19 (4), 483–497. doi:10.1080/1463922X.2017.1381777
- Galy, E., Cariou, M., and Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *Int. J. Psychophysiol.* 83 (3), 269–275. doi:10.1016/j.psychophys.2011.09.023
- Georgiou, K., Larentzakis, A. V., Khamis, N. N., Alsuhailani, G. I., Alaska, Y. A., and Giallafos, E. J. (2018). Can wearable devices accurately measure heart rate variability? A systematic review. *Folia Medica* 60 (1), 7–20. doi:10.2478/folmed-2018-0012
- Ghanbary Sartang, A., Ashnagar, M., Habibi, E., and Sadeghi, S. (2016). Evaluation of Rating Scale Mental Effort (RSME) effectiveness for mental workload assessment in nurses. *J. Occup. Health Epidemiol.* 5 (4), 211–217. doi:10.18869/acadpub.johe.5.4.211
- Grandi, F., Peruzzini, M., Cavallaro, S., Prati, E., and Pellicciari, M. (2022). Creation of a UX index to design human tasks and workstations. *Int. J. Comput. Integr. Manuf.* 35 (1), 4–20. doi:10.1080/0951192X.2021.1972470
- Hancock, G. M., Longo, L., Young, M. S., and Hancock, P. A. (2021). Mental workload. *Handb. Hum. factors Ergon.*, 203–226. doi:10.1002/9781119636113.ch7
- Hart, S. G. (2006). "NASA-task load index (NASA-TLX); 20 years later," in Proceedings of the human factors and ergonomics society annual meeting (Sage CA: Los Angeles, CA: Sage publications). doi:10.1177/154193120605000909
- Hart, S. G., and Staveland, L. E. (1988). "Development of NASA-TLX (task load index): results of empirical and theoretical research. *Adv. Psychol.* 52, 139–183. doi:10.1016/S0166-4115(08)62386-9
- Herlambang, M. B., Cnossen, F., and Taatgen, N. A. (2021). The effects of intrinsic motivation on mental fatigue. *PLoS One* 16 (1), e0243754. doi:10.1371/journal.pone.0243754
- Hermans, E. J., Henckens, M. J., Joëls, M., and Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends Neurosci.* 37 (6), 304–314. doi:10.1016/j.tins.2014.03.006
- Hertzum, M., and Holmegaard, K. D. (2013). Perceived time as a measure of mental workload: effects of time constraints and task success. *Int. J. Human-Computer Interact.* 29 (1), 26–39. doi:10.1080/10447318.2012.676538
- Hidalgo-Muñoz, A. R., Mouratille, D., Matton, N., Causse, M., Rouillard, Y., and El-Yagoubi, R. (2018). Cardiovascular correlates of emotional state, cognitive workload and time-on-task effect during a realistic flight simulation. *Int. J. Psychophysiol.* 128, 62–69. doi:10.1016/j.psychophys.2018.04.002
- Hoonakker, P., Carayon, P., Gurses, A. P., Brown, R., Khunlertkit, A., McGuire, K., et al. (2011). Measuring workload of ICU nurses with a questionnaire survey: the NASA Task Load Index (TLX). *IEE Trans. Healthc. Syst. Eng.* 1 (2), 131–143. doi:10.1080/19488300.2011.609524
- Hoover, A., Singh, A., Fishel-Brown, S., and Muth, E. (2012). Real-time detection of workload changes using heart rate variability. *Biomed. Signal Process. Control* 7 (4), 333–341. doi:10.1016/j.bspc.2011.07.004
- Ishaque, S., Khan, N., and Krishnan, S. (2021). Trends in heart-rate variability signal analysis. *Front. Digital Health* 3, 639444. doi:10.3389/fdgh.2021.639444
- Jafari, M. J., Hassanzadeh-Rangi, N., and Al-Qaisi, S. (2023). Toward real-time mental workload monitoring in hybrid human-AI work environments: a cognitive ergonomics perspective. *Appl. Ergon.* 113, 104028. doi:10.1016/j.apergo.2023.104028
- Jafari, M. J., Zaeri, F., Jafari, A. H., Payandeh Najafabadi, A. T., Al-Qaisi, S., and Hassanzadeh-Rangi, N. (2020). Assessment and monitoring of mental workload in subway train operations using physiological, subjective, and performance measures. *Hum. Factors Ergon. Manuf. and Serv. Ind.* 30 (3), 165–175. doi:10.1002/hfm.20831
- Jeffri, N. F. S., and Awang Rambli, D. R. (2021). A review of augmented reality systems and their effects on mental workload and task performance. *Heliyon* 7 (3), e06277. doi:10.1016/j.heliyon.2021.e06277
- Kantowitz, B. H. (1987). "3. Mental workload. *Adv. Psychol.*, 47, 81–121. doi:10.1016/S0166-4115(08)62307-9

- Karim, N., Hasan, J. A., and Ali, S. S. (2011). Heart rate variability-a review. *J. Basic and Appl. Sci.* 7 (1).
- Kramer, A. F. (2020). Physiological metrics of mental workload: a review of recent progress. *Mult. Task Perform.*, 279–328. doi:10.1201/9781003069447-14
- Kubios (2024). *HRV preprocessing*. Available online at: <https://www.kubios.com/blog/preprocessing-of-hrv-data/> (Accessed March 16, 2025).
- Kusnanto, K., Rohmah, F. A., Andri, S. W., and Hidayat, A. (2020). Mental workload and stress with blood glucose level: a correlational study among lecturers who are structural officers at the university. *Syst. Rev. Pharm.* 11 (7), 253–257. doi:10.31838/srp.2020.7.40
- Lagomarsino, M., Lorenzini, M., De Momi, E., and Ajoudani, A. (2022). An online framework for cognitive load assessment in industrial tasks. *Rob. Comput.- Integr. Manuf.* 78, 102380. doi:10.1016/j.rcim.2022.102380
- LaMorte, W. W. (2021). *Correlation and regression*. Boston University School of Public Health. Available online at: <https://sphweb.bumc.bu.edu/otlt/MPH-Modules/PH717-QuantCore/PH717-Module9-Correlation-Regression/PH717-Module9-Correlation-Regression4.html>.
- Laviola, E., Gattullo, M., Manghisi, V. M., Fiorentino, M., and Uva, A. E. (2022). Minimal AR: visual asset optimization for the authoring of augmented reality work instructions in manufacturing. *Int. J. Adv. Manuf. Technol.* 119, 1769–1784. doi:10.1007/s00170-021-08449-6
- Lean, Y., and Shan, F. (2012). Brief review on physiological and biochemical evaluations of human mental workload. *Hum. Factors Ergon. Manuf. and Serv. Ind.* 22 (3), 177–187. doi:10.1002/hfm.2026
- Lindblom, J., and Thorvald, P. (2014). Towards a framework for reducing cognitive load in manufacturing personnel. *Adv. cognitive Eng. Neuroergonomics* 11, 233–244. doi:10.54941/ahfe100234
- Longo, L., Hancock, P. A., and Hancock, G. M. (2023). Cognitive ergonomics in Industry 5.0: challenges and emerging directions. *Front. Psychol.* 14, 1198763. doi:10.3389/fpsyg.2023.1198763
- Longo, L., and Orrù, G. (2022). Evaluating instructional designs with mental workload assessments in university classrooms. *Behav. and Inf. Technol.* 41 (6), 1199–1229. doi:10.1080/01044929X.2020.1864019
- Longo, L., Wickens, C. D., Hancock, G., and Hancock, P. A. (2022). Human mental workload: a survey and a novel inclusive definition. *Front. Psychol.* 13, 883321. doi:10.3389/fpsyg.2022.883321
- Ma, X., Monfared, R., Grant, R., and Goh, Y. M. (2024). Determining cognitive workload using physiological measurements: pupillometry and heart-rate variability. *Sensors* 24 (6), 2010. doi:10.3390/s24062010
- MacDonald, W. (2003). The impact of job demands and workload on stress and fatigue. *Aust. Psychol.* 38 (2), 102–117. doi:10.1080/00050060310001707107
- Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., et al. (1996). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Eur. heart J.* 17 (3), 354–381. doi:10.1093/oxfordjournals.eurheartj.a014868
- Mandrick, K., Chua, Z., Causse, M., Perrey, S., and Dehais, F. (2016). Why a comprehensive understanding of mental workload through the measurement of neurovascular coupling is a key issue for neuroergonomics? *Front. Hum. Neurosci.* 10, 250. doi:10.3389/fnhum.2016.00250
- Matthews, G., Reinerman-Jones, L. E., Barber, D. J., and Abich, J. IV (2015). The psychometrics of mental workload: multiple measures are sensitive but divergent. *Hum. Factors J. Hum. Factors Ergon. Soc.* 57 (1), 125–143. doi:10.1177/0018720814539505
- Miller, S. (2001). *Workload measures*. National advanced driving simulator. Iowa City, United States.
- Muñoz-de-Escalona, E., Leva, M. C., and Cañas, J. J. (2024). Mental workload as a predictor of ATCO's performance: lessons learnt from ATM task-related experiments. *Aerospace* 11 (8), 691. doi:10.6084/m9.figshare.17433440
- Nomura, S., Mizuno, T., Nozawa, A., Asano, H., and Ide, H. (2009). "Salivary cortisol as a new biomarker for a mild mental workload," (IEEE), 127–131. doi:10.1109/ICBAKE.2009.322009 international conference on biometrics and kansei engineering
- Parent, M., Peysakhovich, V., Mandrick, K., Tremblay, S., and Causse, M. (2019). The diagnosticity of psychophysiological signatures: can we disentangle mental workload from acute stress with ECG and fNIRS? *Int. J. Psychophysiol.* 146, 139–147. doi:10.1016/j.jpsycho.2019.09.005
- Radüntz, T. (2017). Dual frequency head maps: a new method for indexing mental workload continuously during execution of cognitive tasks. *Front. Physiology* 8, 1019. doi:10.3389/fphys.2017.01019
- Rajcani, J., Solarikova, P., Turonova, D., Brezina, I., and Rajcáni, J. (2016). Heart rate variability in psychosocial stress: comparison between laboratory and real-life setting. *Act. Nerv. Super. Rediviva* 58 (3), 77–82.
- Rasmussen, J. (1979). Reflections on the concept of operator workload. *Ment. workload theory Meas.*, 29–40. doi:10.1007/978-1-4757-0884-44
- Rawson, E. S., Lieberman, H. R., Walsh, T. M., Zuber, S. M., Harhart, J. M., and Matthews, T. C. (2008). Creatine supplementation does not improve cognitive function in young adults. *Physiol. and Behav.* 95 (1–2), 130–134. doi:10.1016/j.physbeh.2008.05.009
- Reid, G. B., and Nygren, T. E. (1988). The subjective workload assessment technique: a scaling procedure for measuring mental workload. *Adv. Psychol.* 52, 185–218. doi:10.1016/S0166-4115(08)62387-0
- Rubio, S., Díaz, E., Martín, J., and Puente, J. M. (2004). Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and workload profile methods. *Appl. Psychol.* 53 (1), 61–86. doi:10.1111/j.1464-0597.2004.00161
- Rusnock, C. F., and Borghetti, B. J. (2018). Workload profiles: a continuous measure of mental workload. *Int. J. Ind. Ergon.* 63, 49–64. doi:10.1016/j.ergon.2016.09.003
- Ryu, K., and Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int. J. Ind. Ergon.* 35 (11), 991–1009. doi:10.1016/j.ergon.2005.04.005
- Schoofs, D., Preuß, D., and Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology* 33 (5), 643–653. doi:10.1016/j.psyneuen.2008.02.004
- Standage, D., Blohm, G., and Dorris, M. C. (2014). On the neural implementation of the speed-accuracy trade-off. *Front. Neurosci.* 8, 236. doi:10.3389/fnins.2014.00236
- Sweller, J., Van Merriënboer, J. J., and Paas, F. G. (1998). Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10, 251–296. doi:10.1023/A:1022193728205
- Tao, D., Tan, H., Wang, H., Zhang, X., Qu, X., and Zhang, T. (2019). A systematic review of physiological measures of mental workload. *Int. J. Environ. Res. public health* 16 (15), 2716. doi:10.3390/ijerph16152716
- Tarvainen, M. P., Niskanen, J., Lipponen, J. A., Ranta-Aho, P. O., and Karjalainen, P. A. (2014). Kubios HRV – heart rate variability analysis software. *Comput. Methods Programs Biomed.* 113 (1), 210–220. doi:10.1016/j.cmpb.2013.07.024
- Teoh Yi Zhe, I., and Keikhosrokiani, P. (2021). Knowledge workers mental workload prediction using optimised ELANFIS. *Appl. Intell.* 51 (4), 2406–2430. doi:10.1007/s10489-020-01928-5
- Thorvald, P., Lindblom, J., and Andreasson, R. (2017). "CLAM-A method for cognitive load assessment in manufacturing," in *Advances in manufacturing technology XXXI* (IOS Press), 114–119. doi:10.3233/978-1-61499-792-4-114
- Thorvald, P., Lindblom, J., and Andreasson, R. (2019). On the development of a method for cognitive load assessment in manufacturing. *Rob. Computer-Integrated Manuf.* 59, 252–266. doi:10.1016/j.rcim.2019.04.012
- Trstenjak, M., Hegedić, M., Cajner, H., Opetuk, T., and Tošanović, N. (2023). "Cognitive ergonomics in industry 5.0," in *International Conference on Flexible Automation and Intelligent Manufacturing* (Cham: Springer Nature Switzerland), 763–770. doi:10.1007/978-3-031-38165-2\_88
- Tsang, P. S., and Velazquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics* 39 (3), 358–381. doi:10.1080/00140139608964470
- Van Acker, B. B., Parmentier, D. D., Conradie, P. D., Van Hove, S., Biondi, A., Bombeke, K., et al. (2021). Development and validation of a behavioural video coding scheme for detecting mental workload in manual assembly. *Ergonomics* 64 (1), 78–102. doi:10.1080/00140139.2020.1811400
- Van Acker, B. B., Parmentier, D. D., Vlerick, P., and Saldien, J. (2018). Understanding mental workload: from a clarifying concept analysis toward an implementable framework. *Cognition, Technol. and work* 20, 351–365. doi:10.1007/s10111-018-0481-3
- Van Oort, J., Tendolkar, I., Hermans, E. J., Mulders, P. C., Beckmann, C. F., Schene, A. H., et al. (2017). How the brain connects in response to acute stress: a review at the human brain systems level. *Neurosci. and Biobehav. Rev.* 83, 281–297. doi:10.1016/j.neubiorev.2017.10.015
- Yannakakis, G. N., Martinez, H. P., and Garbarino, M. (2016). "Psychophysiology in games," in *Emotion in games: theory and praxis* (Cham: Springer International Publishing), 119–137.
- Young, M. S., Brookhuis, K. A., Wickens, C. D., and Hancock, P. A. (2015). State of science: mental workload.
- Young, M. S., and Stanton, N. A. (2002). Attention and automation: new perspectives on mental underload and performance. *Theor. issues Ergon. Sci.* 3 (2), 178–194. doi:10.1080/14639220210123789
- Zeier, H., Brauchli, P., and Joller-Jemelka, H. I. (1996). Effects of work demands on immunoglobulin A and cortisol in air traffic controllers. *Biol. Psychol.* 42 (3), 413–423. doi:10.1016/0301-0511(95)05170-8
- Zijlstra, F. R. H. (1993). *Efficiency in work behaviour: a design approach for modern tools*. Delft: Delft University Press.
- Zoakafi, M., Kazemi, R., Choobineh, A., SaboorYaraghi, A., Nematollahi, S., and Zakerian, S. A. (2020). Relationship between mental workload and salivary cortisol levels: a field study. *Work* 67 (2), 381–386. doi:10.3233/WOR-203287