Cognitive workload assessment during VR forklift training

Saman Jamshid Nezhad Zahabi, Md Shafiqul Islam, Sunwook Kim, Nathan Lau, Maury A. Nussbaum, Sol Lim

Department of Industrial and Systems Engineering, Virginia Tech, VA, USA

Corresponding Author: Sol Lim, Ph.D.
Email: sol@vt.edu
Address:
549 Whittemore Hall
1185 Perry Street
Blacksburg, Virginia 24061
P: (540) 231-0083

Abstract: Virtual Reality (VR)-based training offers a safe and engaging environment for training forklift operators. Given the complexity of forklift operation, monitoring the cognitive workload of novice operators in these virtual settings is essential for optimizing the training process. This study investigated cognitive workload variation during a VR-based training for forklift operators due to varying levels of task difficulty and repeated training. Twenty novice participants completed two sessions in a VR simulator with each session including three forklift driving lessons at three difficulty levels. Perceived workload (NASA-TLX) and normalized encephalographic (EEG) activity were employed to assess cognitive workload. Five of the six NASA-TLX subscales and EEG activity in three distinct frequency bands (theta, alpha and beta) all significantly increased with increasing task difficulty. However, we did not observe significant changes in cognitive workload as measured by EEG in the second training session, highlighting a potential limitation in using EEG to track workload variations across days. Perceived workload and EEG measures showed moderate, positive correlations. Our results highlight the potential of EEG for real-time monitoring of workload during VR-based forklift training, particularly in differentiating tasks of varying difficulty. While more research is needed to confirm measurement consistency across sessions, this capability could facilitate worker monitoring to deliver timely alerts or assistance when workload levels exceed optimal thresholds.

Keywords: *Forklift operation, Virtual reality, Task difficulty, Electroencephalogram (EEG), Workload assessment*

1. Introduction

Forklift trucks are widely used material handling equipment in warehouses and other industrial settings. Operating a forklift often requires performing cognitively demanding tasks, such as pallet loading/unloading and precision driving in narrow aisles. These activities necessitate focused attention and skilled maneuvering (Ulutas & Ozkan, 2019), and the risk of accidents depends substantially on an operator's work performance (Cohen & Jensen, 1984; Al-Shaebi et al., 2017). Therefore, effective forklift operation is vital to ensure workplace safety and efficiency.

Cognitive workload plays a crucial role in forklift operation, since elevated workload levels can compromise situational awareness and work-related injuries (Choi et al., 2020). Cognitive workload reflects the balance between task demands and an operator’s mental resources (Wickens, 2002). Both cognitive overload and underload can lead to reduced performance and increased error rates (Young et al., 2015; Biondi et al., 2021; Yurko et al., 2010). Monitoring cognitive workload during operation offers practical benefits, such as enhancing operator safety by enabling timely interventions to prevent performance declines (Hoepf et al., 2016). However, to date no studies systematically investigated or compared the utility of different methods for measuring cognitive workload among forklift operators.

Directly measuring cognitive workload during real-world forklift operation poses several challenges, including high resource requirements, safety risks associated with exposing novice operators to hazardous scenarios, and limited experimental control over workload levels. Virtual Reality (VR) training platforms can be a promising alternative, by simulating the cognitive demands of real-world tasks in a safe and controlled environment. Prior studies have demonstrated that fully immersive VR platforms effectively simulate real-world cognitive workload, offering a deeper understanding of operator cognitive processes (Luque et al., 2024; Lackey et al., 2016). VR forklift training creates an immersive and interactive environment for novice operators (Islam et al., 2024; Abbas et al., 2023; Yuen et al., 2010), supporting simulation of hazardous scenarios without exposing trainees to real-world dangers (Neira-Tovar et al., 2022). Consequently, assessing forklift operators' cognitive workload in a controlled VR environment presents a unique opportunity to evaluate the efficacy of different cognitive workload measurement methods, without exposing operators and trainees to extreme workload conditions that could pose safety risks in real-world scenarios.

Selecting an appropriate method for measuring cognitive workload is crucial, as different approaches vary in reliability, applicability, and potential impact on task performance (Kramer, 1991; Miller, 2001). Among various methods or measures to assess cognitive workload, the NASA-Task Load Index (NASA-TLX) is a widely used, multidimensional tool to quantify perceived workload (Hart & Staveland, 1988). Despite its simplicity and ease of administration, NASA-TLX results can be influenced by individual biases and relatively high between-rater variability (Hart & Wickens, 1990). Administering the NASA-TLX can also disrupt work or VR immersive experience, because operators or trainees must momentarily shift their focus to complete the ratings of perceived workload. Physiological measures, such as electroencephalograms (EEG), offer an objective alternative for assessing cognitive workload without interfering with task performance (Hart & Wickens, 1990). EEG provides high temporal resolution, enabling near real-time monitoring of brain activity during task performance. It has been used widely to evaluate cognitive workload in activities such as driving (Di Flumeri et al., 2018), flight simulation (Kakkos et al., 2019), and VR-based tasks (Dey et al., 2019).

Nonetheless, EEG signals are susceptible to undesired noise, such as from poor device attachment or physiological artifacts (Jiang et al., 2019).

Combining subjective (e.g., NASA-TLX) and objective (e.g., EEG) measures is often recommended for a more comprehensive understanding of cognitive workload (Miller, 2001). Subjective methods offer a practical and relatively simple way to assess the participant's self-reported perceived workload (Cao et al., 2009), while objective methods provide a quantitative, real-time measure of cognitive workload during task engagement (Berka et al., 2007). Some recent studies have employed the combination of NASA-TLX and EEG measures to assess cognitive workload in a VR environment and have shown mixed relationships between these measures. For example, Mondellini et al. (2023) used NASA-TLX and EEG to evaluate mental workload during an n-back test in a VR setting. While they observed an overall increase in NASA-TLX scores across task difficulty levels, no significant changes were observed in EEG measurements, as well as no significant correlation with NASA-TLX. In a VR-based puzzle task, Li & Kim (2021) found greater NASA-TLX scores and less relative power in the beta band (12-30 Hz) at the frontal and parietal regions for a more difficult task. One study found a positive correlation between NASA-TLX and EEG (beta band power of the right temporal lobe) in a VR flight simulation (Ji et al., 2023).

Given the limited empirical evidence on the efficacy of NASA-TLX and EEG activity as measures of cognitive workload in VR-based training for forklift driver training, we had two objectives in our study. Our first objective was to investigate how different task difficulty levels impact changes in cognitive workload for novice operators during VR forklift training and to assess how this workload evolves with repeated training sessions. Second, we investigated the relationship between NASA-TLX and EEG activity to evaluate their respective utility and determine whether one method might serve as a practical substitute for the other.

2. Methods

2.1 Participants

Data analyzed here were obtained as part of a study reported elsewhere, in which a more detailed description of participant characteristics, experimental tasks, and procedures are available (Islam et al., 2024). In brief, 20 novice participants from the local university and community completed the study [14 males and 6 females; mean age (SD): 22.7 (4.3) years for males and 26.5 (5.8) years for females]. All participants were university students with a valid motor vehicle driver's license. All except two reported having no prior experience with forklift operation, with those two having less than one month of experience. Also, all participants reported having normal vision or corrected-to-normal vision (20/20). This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board (IRB #22-341) at Virginia Tech. Informed consent was obtained from each participant prior to participation.

2.2 Task Design

A high-fidelity VR-based order picker forklift simulator was used (The Raymond Corporation, USA). The simulator was designed to closely replicate real-world operations, featuring a VR headset (HTC Corporation, Taiwan) and a real cockpit with physical buttons and a steering wheel identical to those of an actual order picker forklift for controlling and maneuvering the vehicle (Figure 1). The VR environment featured scenarios that mirrored the tasks and conditions forklift drivers encounter during real-world operations, providing an immersive and realistic training experience.

Figure 1: VR training environment showing a participant wearing an EEG headset (a), and the participant's view within the VR setting (b).

Three VR training sessions with varying levels of task difficulty (categorized as low, medium and high difficulty level) were chosen to assess cognitive workload. The specific requirements for completing each of these training sessions are outlined below.

- Low (L): Drive along a specific path composed of multiple 45-90° turns (Figure 2a).
- Medium (M): Drive to a specific location, pick up a pallet using the forks, and then drive in reverse along a designated path to place the pallet at a predetermined location (Figure 2b).
- High (H): Drive to a specific aisle, raise the forks along with the operator's standing platform off the ground, to pick up the boxes placed on the racks. Five boxes, located in different aisles, were collected by the participants, by physically reaching out and "touching" the boxes with their hands in the simulated environment. (Figure 2c).

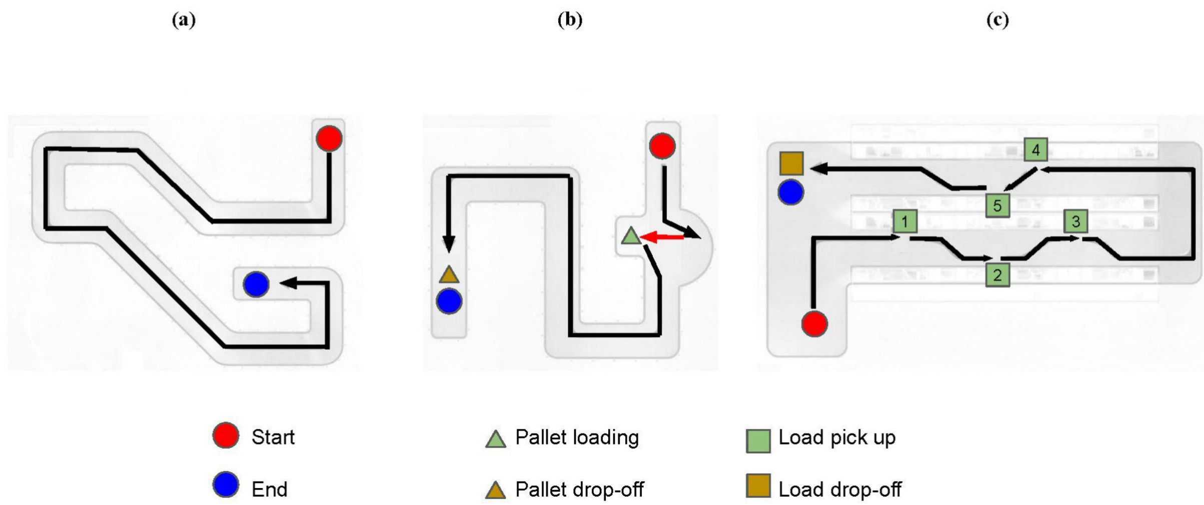

Figure 2: Driving lesson maps in the low (a), medium (b), and high (c) task difficulty levels. Black arrows indicate forward driving and the red arrow indicates reverse driving.

The task difficulty levels were determined based on our pilot study and the sequence of training techniques designed in the VR simulator. In the low task difficulty (i.e., lesson), participants learned basic truck control and navigation along a path. In the medium task difficulty, participants practiced both forward and reverse driving and learned how to operate the forks. In the high task difficulty, participants used both forward and reverse driving to position the forklift optimally for picking up loads from the shelves.

2.3 Experimental procedures

Participants completed two sessions with the same protocol, with the second session completed within five days of the first. In each session, participants performed three driving lessons in a fixed order, from low to higher difficulty levels (Figure 2). While randomizing the lesson order could have been more effective at controlling for confounding effects (e.g., to reduce potential learning or fatigue effects), such experimental design would have compromised ecological validity. In the real world, trainees naturally progress from basic to advanced driving skills. For this reason, the fixed order was used.

During each study session, participants began with an initial familiarization period to learn forklift operations and controls within the VR environment (Figure 1b). The initial familiarization period lasted until participants either indicated they felt ready to proceed to the main training tasks or reached a predetermined time limit of 10 minutes. Experimenters supervised the initial familiarization period, providing verbal instructions and feedback to ensure participants effectively learned how to operate the vehicle. Subsequently, participants donned an EEG headset and the VR headset over it. (Figure 1a). Specifically, we used a six channel dry-electrode headset (Arc device, Mindrove, Hungary; Figure 1a) to collect continuous EEG data at 500 Hz. Following recommendations by the International 10-20 system (Klem et al., 1999), the six main electrodes of the device were positioned to cover six locations over the frontal lobe (F5, F3, F1, F2, F4, and F6). Reference and bias electrodes were placed and secured behind the right and left ears, respectively. The experimenter confirmed visually that the EEG and VR headsets did not physically interfere with each other. Before starting the driving lessons, we collected open eyes resting (baseline) EEG measures while participants sat in a comfortable chair, wearing both EEG and VR headsets in a manner that would reduce the noise caused by the muscle and eye movement. During this resting EEG collection, participants were asked to stay in a relaxed and awake state and focus on a fixed point in the center of the VR screen for two minutes, minimizing body movements and eye blinking to the extent possible.

After the baseline measurement, participants underwent driving lessons during which we recorded EEG data. Participants completed one trial at each task difficulty level (i.e., low, medium, and high). No time restrictions were imposed, and participants were allowed to complete the tasks at their own pace. Trials were terminated if participants failed to complete the same lesson three times consecutively, either by crashing the truck into the racks or damaging the pallet. The mean (SD) completion times for the tasks were 1.2 (0.2) minutes for low difficulty, 2.0 (0.6) minutes for medium difficulty, and 4.0 (1.2) minutes for high difficulty.

Upon completing each lesson, participants were asked to provide ratings on the six NASA-TLX subscales (Hart & Staveland, 1988): mental workload, physical demand, temporal demand, performance, effort, and frustration. We excluded the pairwise comparison step to focus on the effect of workload on each NASA-TLX subscale, as examining subscale ratings rather than a single overall workload score has demonstrated effectiveness in identifying the sources of workload (Hart, 2006). Resting periods were provided after participants completed each NASA-TLX questionnaire and before beginning the next task. To minimize boredom and motion sickness, participants were encouraged to rest for as long as needed and resume the experiment only when they felt ready to continue.

2.4 EEG Data Processing

Raw EEG data from each training lesson were imported to the EEGLAB package (v2022.0, Delorme & Makeig, 2004) in MATLAB (R2022b, The MathWorks Inc, US) and visually inspected to identify potential artifacts. Noisy channels were removed and interpolated using the neighborhood channels. EEG data were downsampled from 500 to 250 Hz and band-pass filtered between 1 and 30 Hz, and an automated algorithm embedded in EEGLAB (Clean-raw data plugin -v2) was used to detect and remove any potential sources of noise and artifacts.

Independent Component Analysis was then applied to the multichannel EEG data to identify artifact components, and such components were removed from the filtered data (Yang et al., 2019). Subsequently, the power spectral density of each channel was calculated using a 1-second Hamming window with no overlap. Mean power for three different frequency bands—theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz)—were calculated separately for all six channels for each trial. The mean power in each frequency band across all the six channels, from each trial, was divided by the respective mean power at baseline (open eyes) for each participant, yielding *normalized band power* (Keskin et al., 2019).

2.5 Statistical analysis

Among the 20 participants, EEG data from one female participant were not recorded due to technical issues. EEG data from one male participant in session 1 were excluded due to noise caused by poor electrode attachment to the scalp (which was only identified after data collection). Three additional participants were unable to finish either the M or H task difficulty levels because they failed the trials after three consecutive attempts. Incomplete trial data were excluded from statistical analysis. NASA-TLX and EEG data were available from a total of 107 trials, including 50 trials from session 1 (18, 16, and 16 trials, respectively, for the L, M, and H difficulty levels) and 57 trials from session 2 (19 trials each for the L, M, and H difficulty levels). Separate repeated-measures analyses of variance (ANOVAs) were performed on each dependent variable to assess the effects of *Task difficulty* (L, M, and H) and *Session* (Session 1 *vs.* 2), including biological *Sex* as a blocking effect. The nine dependent variables were the six NASA-TLX subscales and normalized theta, alpha, and beta power. Significant effects were examined further using Tukey's HSD post-hoc pairwise comparison. Additionally, Pearson correlation analysis was performed between the six NASA-TLX subscales and the EEG band powers, to quantify the linear associations between subjective and objective measures of cognitive workload. All statistical analyses were performed in JMP Pro *v16.0.0* (SAS, NC, USA), using the restricted maximum likelihood (REML) method. Parametric model assumptions were confirmed, and statistical significance was determined when $p < .05$. Effect sizes are reported using partial eta squared (η^2_p).

3. Results

Descriptive statistics of the NASA-TLX and EEG (normalized theta, alpha, and beta power) is provided in Table 1 with respect to *Task difficulty* and *Session*. More detailed results are provided below.

Table 1: Mean (SD) of NASA-TLX ratings and EEG activity (normalized theta, alpha, and beta power) stratified by *Task difficulty* level (L, M, and H) and *Session*.

	Session 1			Session 2		
Variable	**L**	**M**	**H**	**L**	**M**	**H**
NASA-TLX						
Mental workload (0-100)	16.5 (14.8)	19.4 (12.0)	35.5 (26.6)	8.2 (8.6)	12.2 (8.6)	19.1 (15.7)
Physical demand (0-100)	11.3 (12.0)	18.3 (21.6)	19.2 (15.9)	8.3 (8.0)	12.2 (12.3)	14.4 (11.2)
Temporal demand (0-100)	10.5 (12.9)	11.9 (11.1)	17.8 (19.3)	5.5 (6.6)	8.1 (9.6)	11.1 (10.8)
Performance (0-100)	2.6 (5.6)	13.4 (14.6)	13.4 (14.2)	2.1 (3.6)	7.6 (7.5)	8.2 (11.6)
Effort (0-100)	20.2 (25.6)	32.5 (27.5)	41.7 (31.5)	7.0 (6.8)	19.4 (25.8)	16.9 (15.4)
Frustration (0-100)	9.7 (13.7)	25.3 (26.0)	23.3 (22.0)	4.6 (6.03)	5.0 (6.8)	12.8 (16.6)
EEG						
Theta (4-8 Hz)	2.4 (1.3)	4.4 (2.1)	7.1 (3.6)	2.1 (0.9)	4.0 (2.1)	5.0 (1.7)
Alpha (8-12 Hz)	1.9 (1.0)	3.4 (1.4)	4.2 (1.6)	1.7 (0.9)	4.2 (3.5)	4.2 (2.8)
Beta (12-30 Hz)	1.6 (0.7)	2.6 (1.5)	6.4 (6.3)	1.3 (0.8)	2.3 (1.7)	2.4 (1.9)

3.1 NASA-TLX

ANOVA results for NASA-TLX are shown in Table 2, and significant pairwise comparisons are reported in Appendix Table A.1. Main effects of *Task difficulty* were significant for all NASA-TLX subscales except for the effort subscale (which approached significance, with p =

0.051). Overall, each of the subscale ratings increased in the more cognitively difficult tasks (Figure 3). Additionally, *Session* had significant main effects on three NASA-TLX subscales: temporal demand, effort, and frustration ratings were all significantly lower in Session 2 compared to Session 1. *Sex* x *Session* interaction effects were significant on mental demand ratings, where only male participants showed a significant difference between sessions. Results for the effort subscale were more complex, in that there was a significant three-way interaction effect of *Task Difficulty* × *Session* × *Sex*. Females showed no significant differences in their effort subscale ratings between different *Task difficulty* or *Session* conditions. However, males had significant differences in these ratings between different *Task Difficulty* levels within Session 1, and between sessions 1 and 2 across similar or different *Task Difficulty* levels (refer to Table A.1 for more details).

Table 2: ANOVA results [F value (p value, η^2_p)] for the main and interaction effects of *Task difficulty*, *Session*, and *Sex* on the six NASA-TLX subscales. Note that significant effects are highlighted using bold font.

NASA-TLX ratings	**Effect**						
	Task difficulty (T)	Session (S)	Sex	T × S	S × Sex	T × Sex	T × S × Sex
Mental demand	**16.48 (.001, 0.131)**	3.76 (.056, 0.002)	2.36 (.143, 0.015)	2.60 (.081, 0.005)	**4.69 (.033, 0.003)**	1.744 (.182, 0.002)	0.61 (.544, 0)
Physical demand	**7.22 (.001, 0.028)**	1.55 (.217, 0)	2.12 (.164, 0.013)	0.60 (.553, 0)	1.35 (.249, 0)	0.23 (.794, 0)	0.48 (.623, 0)
Temporal demand	**6.43 (.003, 0.023)**	**8.91 (.004, 0.010)**	3.27 (.090, 0.029)	1.32 (.272, 0.001)	0.02 (.890, 0)	0.22 (.798, 0)	1.45 (.240, 0.001)
Performance	**3.93 (.024, 0.011)**	0.70 (.404, 0)	0.64 (.447, 0.002)	0.24 (.790, 0)	1.72 (.194, 0.001)	0.91 (.409, 0.001)	0.42 (.656, 0)
Effort	3.11 (.051, 0.007)	**7.83 (.007, 0.009)**	2.05 (.171, 0.012)	**4.88 (.010, 0.016)**	2.53 (.116, 0.001)	2.53 (.087, 0.005)	**4.44 (.015, 0.013)**
Frustration	**6.43 (.003, 0.023)**	**6.72 (.011, 0.006)**	1.43 (.249, 0.007)	0.63 (.538, 0)	1.77 (.187, 0)	0.98 (.379, 0.001)	0.36 (.700, 0)

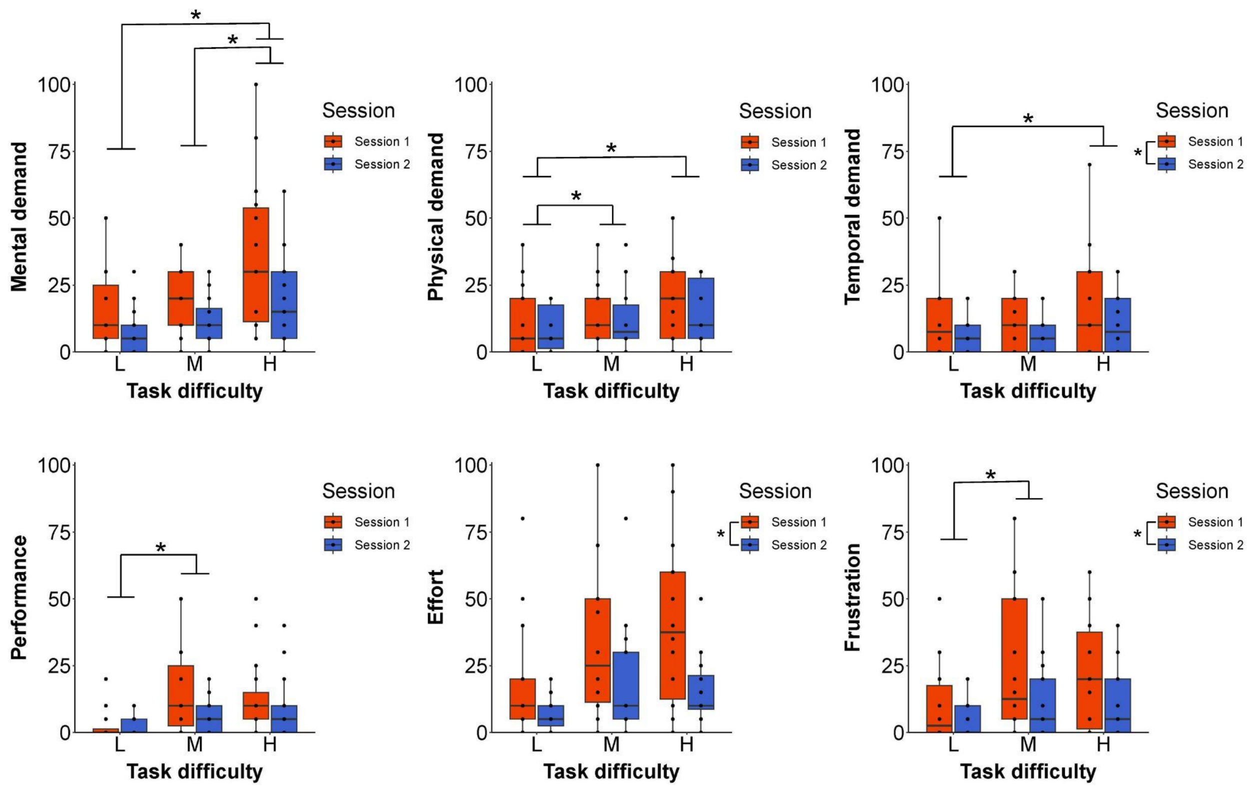

Figure 3: Effects of *Task difficulty* and *Session* on NASA-TLX mental workload rating (L: Low, M: Medium, H: High). "*" indicates a significant paired difference between *Task difficulty* and *Session* levels ($p < .05$). Whiskers are 1.5 times the interquartile range (IQR) from the first and third quartiles.

3.2 EEG

The main effect of *Task difficulty* was significant for normalized power in all three EEG bands (Table 3, Figure 4). Post hoc analysis revealed a significant increase of power between all task difficulty levels for the theta band. Normalized theta power significantly increased by 88.2% from L to M, 165.6% from L to H, and 41.1% from M to H task difficulty levels. Significant differences in normalized alpha power were observed between L and M and between L and H task difficulty levels. Specifically, normalized alpha power increased by 113% from L to M, and by 129% from

L to H. Lastly, normalized beta activity significantly increased by 128.2% from L to H task difficulty level.

Table 3: ANOVA results [F value (p value, η^2_p)] regarding the main and interaction effects of *Task difficulty*, *Session*, and *Sex* on EEG activity (normalized power in theta, alpha and beta band). Note that significant effects are highlighted using bold font.

Normalized Power	**Effect**						
	Task difficulty (T)	Session (S)	Sex	T × S	S × Sex	T × Sex	T × S × Sex
Theta (4-8 Hz)	**17.32 (.001, 0.177)**	0.59 (.445, 0)	3.13 (.094, 0.022)	0.21 (.814, 0)	2.62 (.110, 0.001)	0.05 (.949, 0)	1.46 (.241, 0.002)
Alpha (8-12 Hz)	**8.12 (.001, 0.061)**	0.08 (.772, 0)	0.01 (.986, 0)	0.17 (.842, 0)	0.14 (.712, 0)	0.14 (.873, 0)	0.12 (.884, 0)
Beta (12-30 Hz)	**5.07 (.009, 0.024)**	3.37 (.071, 0.002)	0.01 (.967, 0)	2.12 (.129, 0.005)	0.25 (.618, 0)	0.15 (.856, 0)	0.07 (.929, 0)

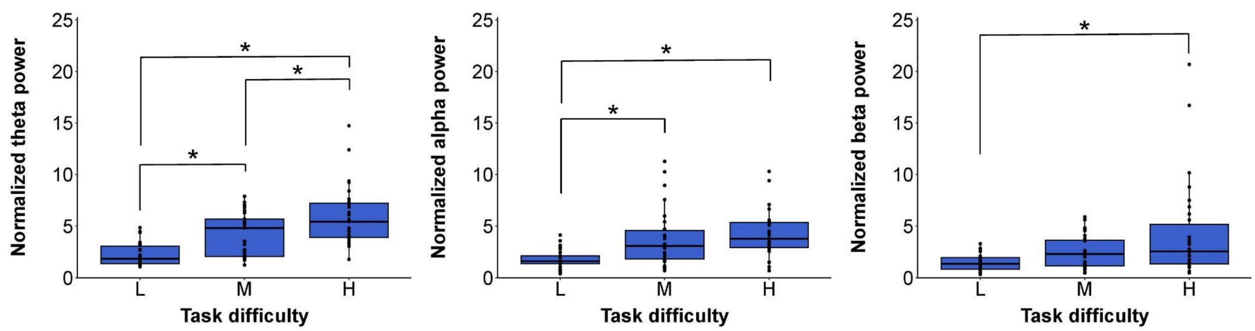

Figure 4: Normalized EEG power in three frequency bands with respect to *Task difficulty* (L: Low, M: Medium, H: High). "*" indicates a significant paired difference ($p < .05$). Whiskers are 1.5 times the interquartile range (IQR) from the first and third quartiles.

3.3 Correlations between NASA-TLX subscales and normalized EEG powers

Normalized power in the theta band was significantly and positively correlated with mental demand, performance, and frustration (Table 4). Normalized power in the alpha band showed a positive correlation only with physical demand. Notably, normalized power in the beta band had significant positive correlations with all NASA-TLX subscales except performance. Overall, the correlations between NASA-TLX subscales and normalized EEG powers appeared linear, though the strength of these correlations was at best modest and varied between subscales and frequency bands (Figures 5 and A.1, A.2).

Table 4: Correlation analysis results [R (p value)] between NASA-TLX subscales and normalized EEG power in three frequency bands.

Variables	NASA-TLX ratings					
	Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration
Normalized Power						
Theta	**0.362 (.001)**	0.138 (.194)	0.118 (.279)	**0.305 (.001)**	0.437 (.222)	**0.247 (.023)**
Alpha	0.210 (.069)	**0.255 (.022)**	0.173 (.129)	0.158 (.185)	0.055 (.647)	0.192 (.092)
Beta	**0.395 (.001)**	**0.292 (.008)**	**0.239 (.035)**	0.105 (.378)	**0.290 (.012)**	**0.385 (.001)**

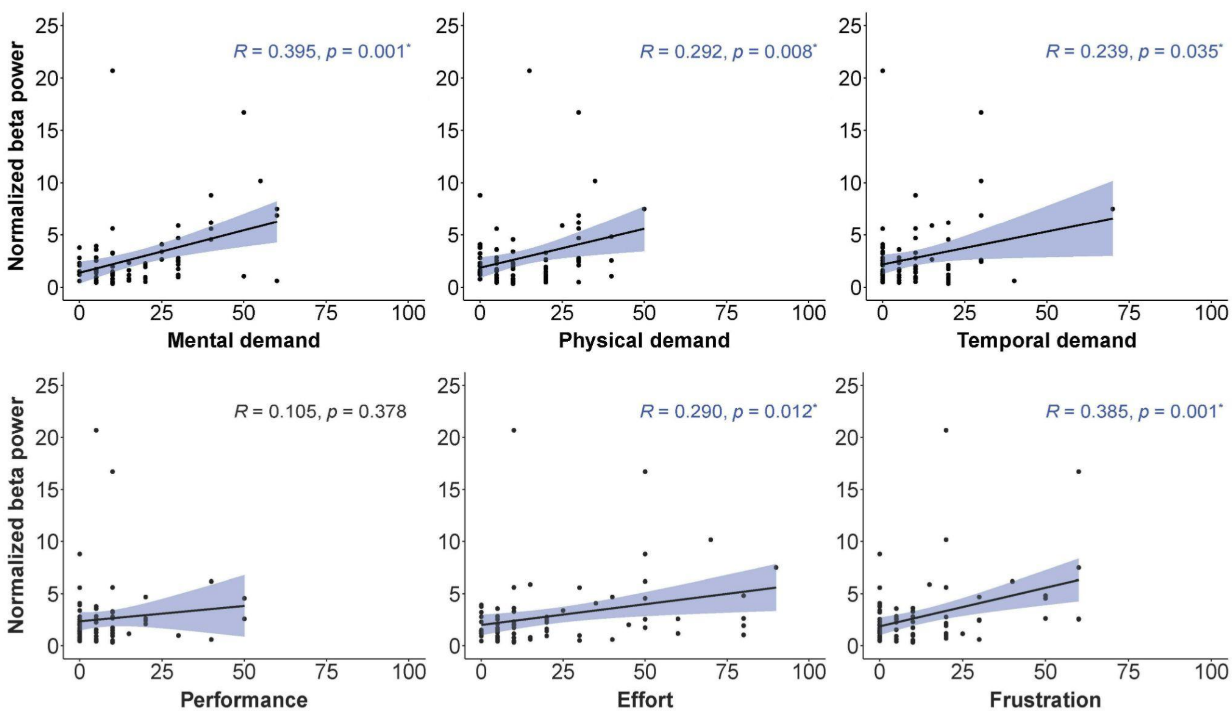

Figure 5: Pearson correlation with 95% confidence interval between NASA-TLX subscales and normalized EEG power in the beta band.

4. Discussion

We investigated the effect of task difficulty and repeated training on perceived workload and brain activity during VR-based forklift training. Both the NASA-TLX (except for the effort subscale) and EEG measures effectively differentiated cognitive workload between different levels of task difficulty during VR training. Among the measures, normalized EEG power in the theta band was the only one sufficiently sensitive in differentiating between all three levels of task difficulty (Figure 4). The larger effect size of normalized EEG power in the theta band versus other measures further suggests its greater sensitivity to variations in workload for the tasks considered in this study. An additional training session did not significantly change cognitive workload measured by the EEG activity, while ratings of three NASA-TLX subscales (temporal demand, effort, frustration) declined in the second training session. The following discussion addresses these results in more detail.

4.1 Task Difficulty

Significant main effect of task difficulty was observed for all NASA TLX subscales except for the effort, which approached significance with $p = 0.051$ (Table 2; Figure 3). The significant increase of NASA-TLX in the more difficult tasks implies that workload manipulations in this study effectively induced the perceived workload of participants across various task difficulty levels.

Normalized EEG power in all three frequency bands (theta, alpha, beta) also significantly increased with task difficulty (Table 3; Figure 4). The highest sensitivity was observed in the theta band, which showed significant differences in all pairwise comparisons among three task difficulty levels. This observation is consistent with previous findings that frontal theta activity increases when performing a highly demanding task (Jensen & Tesche, 2002; Depestele et al., 2023), and that theta band synchronization can serve as a reliable indicator of cognitive workload elevation (Borghini et al., 2014; Shou & Ding, 2013). Theta activity is closely linked to different cognitive processes, such as attention, decision making and encoding the new information into episodic memory (Klimesch, 1999; Ishii et al., 1999; Senftleben & Scherbaum, 2021). Compared to a 2D VR environment, the highly immersive nature of a 3D VR environment can trigger an increase in theta activity, driven by the need to encoding spatial information and search for spatial cues and landmarks (Slobounov et al., 2015; Kober et al., 2012). In our study, the increase in theta band activity with increasing task demands may be attributed to the processing of visuo-spatial information within the 3D VR environment, especially when task complexity increased (Bischof & Boulanger, 2003; Jaiswal et al., 2010). Another possible explanation is that during more challenging tasks participants had to increase their attention and arousal levels to meet the task demands, which may have influenced their brain activation (Diaz-Piedra et al., 2020).

While not as sensitive as theta, normalized EEG power in the alpha band also showed a significant increase with increasing task difficulty (from L to M and L to H). Most prior research findings showed decreased power in the EEG alpha band with increasing cognitive workload (Lei & Roetting, 2011; Morton et al., 2022), or alpha desynchronization (Klimesch et al., 1997) . However, some researchers have reported the opposite observation, similar to ours, which is known as alpha synchronization (Pope et al., 1995, Michels et al., 2008). These mixed findings can be explained by two contradicting pathways, in which alpha band *desynchronization* is related to memory maintenance while *synchronization* is associated with suppressing irrelevant stimuli (Puma et al., 2018). We postulate that our result might indicate participant effort in suppressing distractors (Rihs et al., 2007) in the VR simulation, to maintain focus or attention on the driving task. The lack of a significant correlation between alpha activity and NASA-TLX subscales (except for physical demand) further supports this interpretation, suggesting that changes in alpha activity might not be primarily linked to mental workload but could instead be driven by underlying factors, such as efforts to avoid distractions. Furthermore, since the decrease in alpha band activity associated with heightened cognitive workload is commonly reported in the parietal regions of the brain (Borghini et al., 2014), making direct comparisons between our findings and others may be challenging. In addition, variation in the alpha frequency band is dependent on some individual differences such as age, sex (Chiang et al., 2011), and memory performance (Klimesch et al., 1993).

The EEG beta band reflects attentional state and active information processing (Güntekin et al., 2013; Ray & Cole, 1985); thus, increased power in the beta band is often used as an indication of increased workload (Örün & Akbulut, 2019) or cognitive overload (Matthews et al., 2017). Here, we found a significant increase in the normalized EEG power in the beta band from low to high task difficulty level (Table 3, Figure 4), which may reflect a heightened cognitive workload and attention to meet the demands of a task (Gong et al., 2019).

4.2 Effects of repeated training

The training session treatment did not significantly impact workload for three NASA-TLX subscales (mental demand, physical demand and performance). However, participants reported a significant decrease in NASA-TLX ratings during the second training session for the effort, frustration, and temporal demand subscales. Effort is influenced by the interplay between alertness and task demands (Galy et al., 2018). In the second session, the reduced effort could suggest that participants' increased alertness allowed them to easily manage the necessary resources, possibly due to task familiarization. Frustration can be related to the participant's level of tension (Galy et al., 2018). The substantial decrease in frustration could therefore be seen as a reduction in tension, likely due to becoming more familiar with the task during the second session, which may also indicate learning, adaptation to the VR environment, or a diminished novelty effect. Repeated training did not affect EEG activity in any of the three frequency bands measured. In another analysis of this study (Islam et al., 2024) reported a decrease in the range of motion and movement variability of participants in the second session, which we interpreted as a result of an increase in their forklift operation skills. However, we did not observe such a decrease in cognitive workload measured by EEG signals.

The absence of significant main effects of training sessions on EEG activity and three NASA-TLX subscales (mental demand, physical demand, performance) suggest that only two training sessions might not be sufficient to substantially reduce the cognitive effort needed to perform the tasks. Borghini et al. (2013) demonstrated that EEG theta activity (as an indicator of cognitive workload) initially increased for novice participants during a VR flight task training. However, with further training sessions, participants enhanced their skills to accomplish the tasks using less cognitive resources, and this was reflected in reduced theta activity in the fifth session of training. We speculate that two training sessions may not be sufficient for participants to develop their own strategies, achieve full proficiency in the task, and reduce the cognitive resources required for its completion. Future studies with more repeated sessions are suggested. Extended training will be useful to assess the sensitivity of EEG activity measures and to determine the optimal training sessions required for substantial reduction in cognitive workload due to learning the task among novice forklift drivers undergoing VR training.

The inability of EEG to detect changes in cognitive workload across different observation days in our study also raises concerns about the test-retest reliability of EEG measurements over multiple sessions. Factors such as variations in electrode placement, differences in the contact between the electrodes and the scalp and the participants' hair condition may contribute to inconsistencies in EEG data collection (Gamano et al., 2018; Alarcón-Segovia et al., 2024; Scrivener & Reader, 2022). Although we maintained consistent EEG procedures, using the device on the second day likely introduced some variations in electrode positioning and scalp contact, affecting signal quality. Future studies are needed to better understand and improve the between-session reliability of EEG measures.

4.3 Differences related to biological sex

While Sex had no main effect on any NASA-TLX subscales or EEG measures, we found significant interactive effects between task difficulty, session and sex for two NASA-TLX subscales (effort and mental demand). For the males, mental demand and effort subscales showed significant changes affected by the session, task difficulty, or their interaction (Table A.1), indicating they were more sensitive to these factors or adapted differently under varying conditions. For the females, the mental demand and effort subscales remained relatively unchanged, suggesting that their perception of mental workload and effort was less influenced across different task difficulties or sessions. This difference in perceived workload ratings between male and female participants, highlights potential sex differences in how novice male and female forklift drivers experience workload in their forklift operations and the need for optimizing training strategies to reduce workload disparities between genders.

4.4 Correlations between NASA-TLX ratings and normalized EEG powers

There was a moderate positive correlation between some NASA-TLX subscales and EEG powers (Table 4; Figure 5), highlighting the potential utility of EEG in measuring workload changes in VR-based forklift driving. EEG power in the beta band showed significant correlations with all NASA-TLX subscales, except for performance. Similar to our findings, Ji et al. (2023) observed a strong correlation between the NASA-TLX and beta activity for pilots during VR-based flight tasks. The correlation between NASA-TLX and EEG power in the beta band suggests the potential of the beta band in the continuous assessment of workload in VR training, in a manner distinct from that of NASA-TLX. Our findings point to potential for future research to create real-time cognitive workload monitoring systems using EEG headsets for forklift driving. These headsets could be worn by forklift operators for seamless integration and data collection to obtain indications of cognitive workload in (near) real-time. Incorporating real-time brain monitoring into forklift operations presents multiple opportunities to enhance training effectiveness, operator safety, and operational efficiency in both virtual and real-world contexts. In VR settings, monitoring the cognitive workload of novice drivers during training could support the design of adaptive, effective programs that align with the learner's available cognitive capacity (Rodenburg et al., 2018; Dey et al., 2019). In real-world scenarios, such monitoring systems could offer immediate feedback and alerts to operators when workload levels approach critical thresholds (Aricò et al., 2017), thereby improving safety and enhancing operational performance.

4.5 Limitations

Some limitations of the study should be acknowledged. First, EEG data were collected from only 6 electrodes at the frontal region of the brain, due to limitations in attaching an EEG device beneath the VR headset. Considering that diverse cognitive processes are associated with distinct brain regions, EEG data collected beyond the frontal region could provide further information for workload measurements. Second, participants were engaged in moderate movement (i.e., load pick-up from a shelf or turning their bodies to reverse driving and to operate the forklift forks) during certain segments of the difficult task conditions. Despite applying several signal processing steps to minimize noise resulting from body movements, we are unsure of the consistency in signal-to-noise ratio across all difficulty levels. In addition, we did not explore the effect of task completion time on perceived workload and EEG measures. While longer task duration could potentially increase perceived workload and EEG measures, the mean duration of lessons were relatively short for all task difficulties (L: 1.2 ± 0.2; M: 2.0 ± 0.6; H: 4.0 ± 1.2 minutes). As mental fatigue is expected to happen over an extended period during driving tasks (Tran et al., 2020), we consider the effect of different task durations between different task difficulty levels to not be of a level sufficient to have generated mental fatigue and drowsiness in our training. The sample size of this study was relatively small. However, our post-hoc power analysis for the effect of task difficulty on most of the variables showed sufficient statistical power (higher than 90%) to detect meaningful differences . Another limitation of this study was the small number of female participants. To better understand the impact of biological sex on the workload of novice forklift drivers in VR training, future research should include a larger sample of female participants. Finally, due to the exploratory nature of our study, we did not consider factors such as gaming experience or cognitive function as inclusion criteria for recruiting the participants or as control variables for the analysis. Future work should incorporate these factors into the study design to explore their potential influence on cognitive workload measurement across various VR-based training platforms.

5. Conclusions

In summary, we studied cognitive workload in a VR-based training environment using NASA-TLX and EEG activity. Both NASA-TLX and EEG power (in theta, alpha, and beta bands) were sensitive at discriminating different levels of cognitive workload between different forklift driving lessons with varying difficulties. Both types of measures also showed moderate positive correlations with each other. However, we could not find evidence that EEG activity can effectively detect changes in cognitive workload (measured by three NASA-TLX subscales, i.e., temporal demand, effort, and frustration) across repeated training sessions. Our results highlight the potential utility of EEG in monitoring cognitive workload induced by distinct levels of task difficulty in a VR training environment, though evidence supporting its utility across repeated sessions remains limited. EEG-based workload monitoring systems hold substantial practical potential for enhancing VR training by enabling adaptive learning tailored to operators' cognitive workload levels. Additionally, in real-time operations, such systems could deliver timely interventions to improve operator safety and enhance operational performance.

6. Declaration of Competing Interest

The authors declare that they have no competing financial interests that could have influenced the work reported in this paper.

7. Acknowledgement:

This study was supported by The Raymond Corporation (a Toyota Industries Company) University Research Program. Any opinions expressed in this paper are those of the authors and do not necessarily reflect the views of The Raymond Corporation. The authors are grateful to Hsiang-Wen Hsing for providing technical support with the study.

Appendix

Table A.1: Results from post-hoc pairwise comparisons (Tukey HSD test) for interaction and main effects of *Task difficulty*, *Session* and *Sex* on nine dependent variables including six NASA-TLX subscales and EEG activity in the theta, alpha and beta bands.

Variable	Main/Interaction Effect	Pair 1	Pair 2	95% CI	*p-value*
Mental demand	*Task difficulty (T)*	L	H	[-23.90, -9.23]	<.001
		M	H	[-21.88, -6.14]	<.001
	Sex × S	Session 1, Male	Session 2, Male	[5.36, 17.17]	<.001
Physical demand	*Task difficulty (T)*	L	M	[-9.34, -.004]	.048
		L	H	[-11.90, -2.63]	.001
Temporal demand	*Task difficulty (T)*	L	H	[-11.79, -2.22]	.002
	Session (S)	Session 1	Session 2	[1.75, 8.74]	.004
Performance	*Task difficulty (T)*	L	M	[-14.12, -0.88]	.023
Effort	*Task difficulty (T)*	Session 1	Session 2	[2.39, 14.20]	.007
	T × S	Session 1, L	Session 1, H	[-31.58, -4.62]	.003
		Session 1, H	Session 2, L	[4.65, 37.46]	.005
		Session 1, H	Session 2, M	[0.60, 27.71]	.036
		Session 1, H	Session 2, H	[6.21, 35.40]	.001
	T × Sex × S	Session 1, L, Male	Session 1, M, Male	[-33.07, -2.95]	.007
		Session 1, L, Male	Session 1, H, Male	[-39.14, -9.02]	<.001
		Session 1, M, Male	Session 2 , L, Male	[7.18, 37.30]	<.001
		Session 1, M, Male	Session 2 , M, Male	[2.18, 32.30]	.012
		Session 1, H, Male	Session 2 , L, Male	[13.25, 43.37]	<.001
		Session 1, H, Male	Session 2 , M, Male	[8.25, 38.37]	<.001
		Session 1, H, Male	Session 2 , H, Male	[2.48, 32.60]	.010
Frustration	*Task difficulty (T)*	L	M	[-18.71, -3.59]	.002
	Session (S)	Session 1	Session 2	[1.59, 12.10]	.011
Theta activity	*Task difficulty (T)*	L	M	[-3.34, -0.58]	.003
		L	H	[-5.01, -2.08]	<.001
		M	H	[-3.08, -0.09]	<.035
Alpha activity	*Task difficulty (T)*	L	M	[-3.60, -0.55]	.005

		L	H	[-3.84, -0.72]	.002
Beta activity	*Task difficulty (T)*	L	H	[-4.97, -0.69]	.007

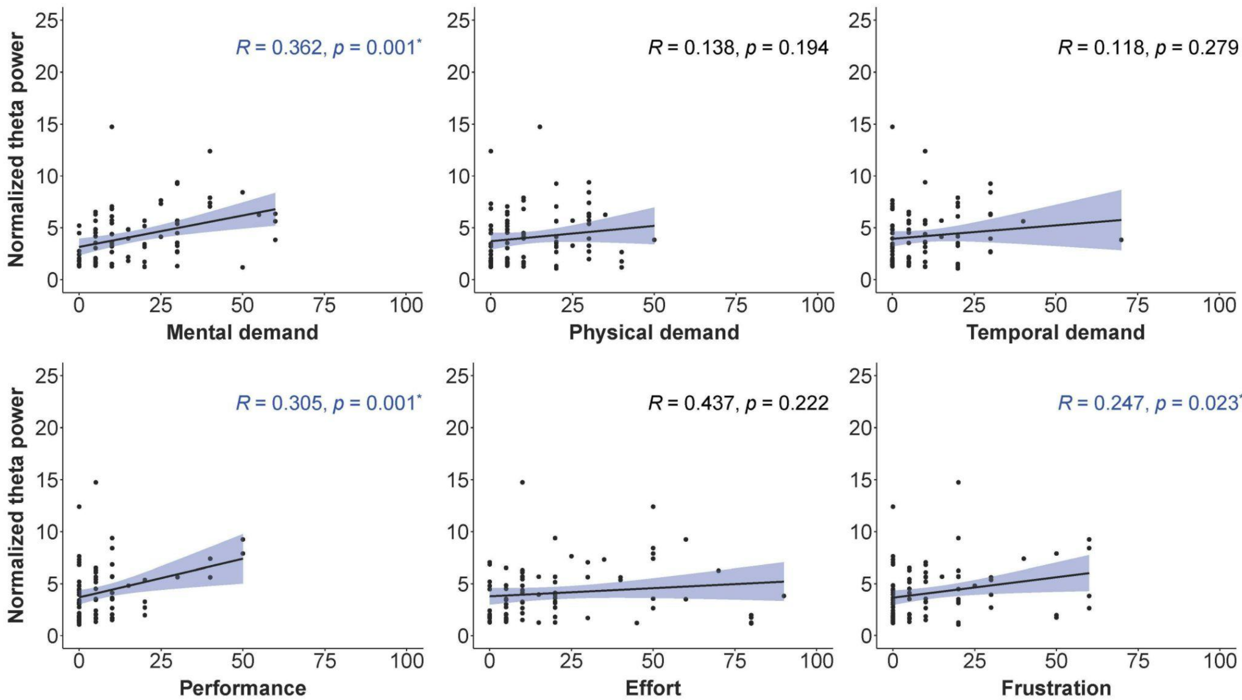

Figure A.1: Correlations between NASA-TLX subscales and normalized EEG power in the *theta* band.

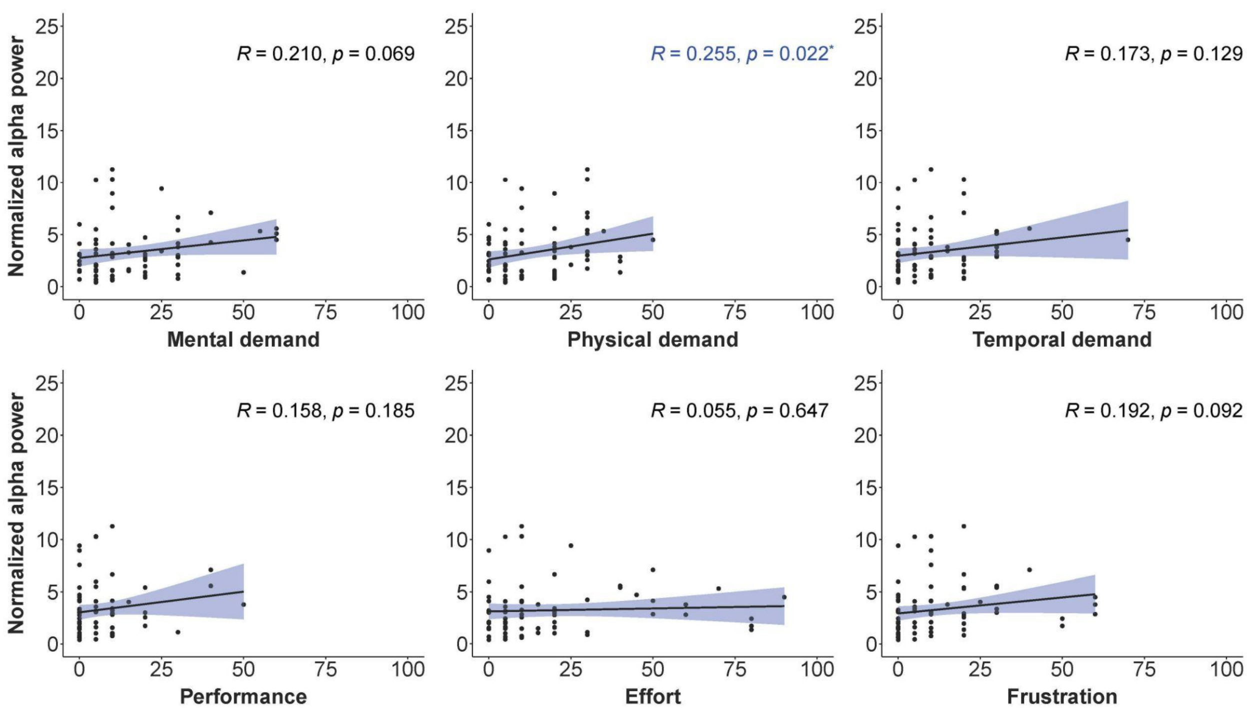

Figure A.2: Correlations between NASA-TLX subscales and normalized EEG power in the *alpha* band.

References

Abbas, A., Seo, J., Ahn, S., Luo, Y., Wyllie, M. J., Lee, G., & Billinghurst, M. (2023). How Immersive Virtual Reality Safety Training System Features Impact Learning Outcomes: An Experimental Study of Forklift Training. *Journal of Management in Engineering*, *39*(1), 04022068. https://doi.org/10.1061/(ASCE)ME.1943-5479.0001101

Alarcón-Segovia, L. C., Morel, R., Spies, R., & Rintoul, I. (2024). THE CHALLENGING SCALP–ELECTRODE INTERFACE AND THE EVOLUTION OF MATERIALS AND ELECTRODE INTEGRATED ICTs FOR ELECTROENCEPHALOGRAPHY. *Surface Review and Letters*, *31*(09), 2430009. https://doi.org/10.1142/S0218625X24300090

Al-Shaebi, A., Khader, N., Daoud, H., Weiss, J., & Yoon, S. W. (2017). The Effect of Forklift Driver Behavior on Energy Consumption and Productivity. *Procedia Manufacturing*, *11*, 778–786. https://doi.org/10.1016/j.promfg.2017.07.179

Aricò, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., Colosimo, A., & Babiloni, F. (2017). Passive BCI in Operational Environments: Insights, Recent Advances, and Future Trends. *IEEE Transactions on Biomedical Engineering*, *64*(7), 1431–1436. IEEE Transactions on Biomedical Engineering. https://doi.org/10.1109/TBME.2017.2694856

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*(5 Suppl), B231-244.

Biondi, F. N., Cacanindin, A., Douglas, C., & Cort, J. (2021). Overloaded and at Work: Investigating the Effect of Cognitive Workload on Assembly Task Performance. *Human Factors*, *63*(5), 813–820. https://doi.org/10.1177/0018720820929928

Bischof, W. F., & Boulanger, P. (2003). Spatial Navigation in Virtual Reality Environments: An EEG Analysis. *CyberPsychology & Behavior*, *6*(5), 487–495. https://doi.org/10.1089/109493103769710514

Borghini, G., Arico, P., Astolfi, L., Toppi, J., Cincotti, F., Mattia, D., Cherubino, P., Vecchiato, G., Maglione, A. G., Graziani, I., & Babiloni, F. (2013). Frontal EEG theta changes assess the training improvements of novices in flight simulation tasks. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, *2013*, 6619–6622. https://doi.org/10.1109/EMBC.2013.6611073

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, *44*, 58–75. https://doi.org/10.1016/j.neubiorev.2012.10.003

Cao, A., Chintamani, K. K., Pandya, A. K., & Ellis, R. D. (2009). NASA TLX: Software for assessing subjective mental workload. *Behavior Research Methods*, *41*(1), 113–117. https://doi.org/10.3758/BRM.41.1.113

Chiang, A. K. I., Rennie, C. J., Robinson, P. A., van Albada, S. J., & Kerr, C. C. (2011). Age trends and sex differences of alpha rhythms including split alpha peaks. *Clinical Neurophysiology*, *122*(8), 1505–1517. https://doi.org/10.1016/j.clinph.2011.01.040

Choi, M., Ahn, S., & Seo, J. (2020). VR-Based investigation of forklift operator situation awareness for preventing collision accidents. *Accident Analysis & Prevention*, *136*, 105404. https://doi.org/10.1016/j.aap.2019.105404

Cohen, H. H., & Jensen, R. C. (1984). Measuring the effectiveness of an industrial lift truck safety training program. *Journal of Safety Research*, *15*(3), 125–135. https://doi.org/10.1016/0022-4375(84)90023-9

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Depestele, S., van Dun, K., Verstraelen, S., Ross, V., Van Hoornweder, S., Brijs, K., Brijs, T., & Meesen, R. (2023). The effect of Dual-Task modality on midfrontal theta power and driving performance across the adult lifespan. *Transportation Research Part F: Traffic Psychology and Behaviour*, *95*, 188–201. https://doi.org/10.1016/j.trf.2023.04.007

Dey, A., Chatburn, A., & Billinghurst, M. (2019). Exploration of an EEG-Based Cognitively Adaptive Training System in Virtual Reality. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 220–226. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). https://doi.org/10.1109/VR.2019.8797840

Di Flumeri, G., Borghini, G., Aricò, P., Sciaraffa, N., Lanzi, P., Pozzi, S., Vignali, V., Lantieri, C., Bichicchi, A., Simone, A., & Babiloni, F. (2018). EEG-Based Mental Workload Neurometric to Evaluate the Impact of Different Traffic and Road Conditions in Real Driving Settings. *Frontiers in Human Neuroscience*, *12*, 509. https://doi.org/10.3389/fnhum.2018.00509

Diaz-Piedra, C., Sebastián, M. V., & Di Stasi, L. L. (2020). EEG Theta Power Activity Reflects Workload among Army Combat Drivers: An Experimental Study. *Brain Sciences*, *10*(4), 199. https://doi.org/10.3390/brainsci10040199

Galy, E., Paxion, J., & Berthelon, C. (2018). Measuring mental workload with the NASA-TLX needs to examine each dimension rather than relying on the global score: An example with driving. *Ergonomics*, *61*(4), 517–527. https://doi.org/10.1080/00140139.2017.1369583

Gamano, Y., Saito, Y., Takamori, K., & Morishige, K.-I. (2018). Effects of EEG Electrode Positional Deviations for Classification Accuracy on Different Days. *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2277–2282. https://doi.org/10.1109/SMC.2018.00391

Gong, D., Li, Y., Yan, Y., Yao, Y., Gao, Y., Liu, T., Ma, W., & Yao, D. (2019). The high-working load states induced by action real-time strategy gaming: An EEG power spectrum and network study. *Neuropsychologia*, *131*, 42–52. https://doi.org/10.1016/j.neuropsychologia.2019.05.002

Güntekin, B., Emek-Savaş, D. D., Kurt, P., Yener, G. G., & Başar, E. (2013). Beta oscillatory responses in healthy subjects and subjects with mild cognitive impairment. *NeuroImage : Clinical*, *3*, 39–46. https://doi.org/10.1016/j.nicl.2013.07.003

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908. https://doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hart, S. G., & Wickens, C. D. (1990). Workload Assessment and Prediction. In H. R. Booher (Ed.), *Manprint: An Approach to Systems Integration* (pp. 257–296). Springer Netherlands. https://doi.org/10.1007/978-94-009-0437-8_9

Hoepf, M., Middendorf, M., Mead, J. O., Gruenwald, C., Credlebaugh, C., Middendorf, P., & Galster, S. (2016, July 13). *Evaluation of Physiologically-Based Artificial Neural Network Models to Detect Operator Workload in Remotely Piloted Aircraft Operations*. https://www.semanticscholar.org/paper/Evaluation-of-Physiologically-Based-Artificial-to-Hoepf-Middendorf/015bfad4c98fa9ace20e98cd30a6674269b30ac9

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., Hirabuki, N., Asada, H., Kihara, T., Robinson, S. E., & Takeda, M. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport*, *10*(4), 675–679. https://doi.org/10.1097/00001756-199903170-00003

Islam, M. S., Zahabi, S. J. N., Kim, S., Lau, N., Nussbaum, M. A., & Lim, S. (2024). Changes in forklift driving performance and postures among novices resulting from training using a high-fidelity virtual reality simulator: An exploratory study. *International Journal of Industrial Ergonomics*, *104*, 103648. https://doi.org/10.1016/j.ergon.2024.103648

Jaiswal, N., Ray, W., & Slobounov, S. (2010). ENCODING OF VISUAL-SPATIAL INFORMATION IN WORKING MEMORY REQUIRES MORE CEREBRAL EFFORTS THAN RETRIEVAL: EVIDENCE FROM EEG AND VIRTUAL REALITY STUDY. *Brain Research*, *1347*, 80–89. https://doi.org/10.1016/j.brainres.2010.05.086

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *The European Journal of Neuroscience*, *15*(8), 1395–1399. https://doi.org/10.1046/j.1460-9568.2002.01975.x

Ji, L., Zhang, C., Li, H., Zhang, N., Guo, C., Zhang, Y., & Tang, X. (2023). A comprehensive experimental framework based on analysis of the pilot's EEG and NASA-TLX questionnaire in a VR environment. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine*, *237*(7), 869–878. https://doi.org/10.1177/09544119231179842

Jiang, X., Bian, G.-B., & Tian, Z. (2019). Removal of Artifacts from EEG Signals: A Review. *Sensors (Basel, Switzerland)*, *19*(5), 987. https://doi.org/10.3390/s19050987

Kakkos, I., Dimitrakopoulos, G. N., Gao, L., Zhang, Y., Qi, P., Matsopoulos, G. K., Thakor, N., Bezerianos, A., & Sun, Y. (2019). Mental Workload Drives Different Reorganizations of Functional Cortical Connectivity Between 2D and 3D Simulated Flight Experiments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, *27*(9), 1704–1713. https://doi.org/10.1109/TNSRE.2019.2930082

Keskin, M., Ooms, K., Dogru, A. O., & De Maeyer, P. (2019). EEG & Eye Tracking User Experiments for Spatial Memory Task on Maps. *ISPRS International Journal of Geo-Information*, *8*(12), Article 12. https://doi.org/10.3390/ijgi8120546

Klem, G. H., Lüders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalography and Clinical Neurophysiology. Supplement*, *52*, 3–6.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research. Brain Research Reviews*, *29*(2–3), 169–195. https://doi.org/10.1016/s0165-0173(98)00056-3

Klimesch, W., Doppelmayr, M., Schimke, H., & Ripper, B. (1997). Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology*, *34*(2), 169–176. https://doi.org/10.1111/j.1469-8986.1997.tb02128.x

Klimesch, W., Schimke, H., & Pfurtscheller, G. (1993). Alpha frequency, cognitive load and memory performance. *Brain Topography*, *5*(3), 241–251. https://doi.org/10.1007/BF01128991

Kober, S. E., Kurzmann, J., & Neuper, C. (2012). Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: An EEG study. *International Journal of Psychophysiology*, *83*(3), 365–374. https://doi.org/10.1016/j.ijpsycho.2011.12.003

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In *Multiple Task Performance*. CRC Press.

Lackey, S. J., Salcedo, J. N., Szalma, J. L., & Hancock, P. A. (2016). The stress and workload of virtual reality training: The effects of presence, immersion and flow. *Ergonomics*, *59*(8), 1060–1072. https://doi.org/10.1080/00140139.2015.1122234

Lei, S., & Roetting, M. (2011). Influence of task combination on EEG spectrum modulation for driver workload estimation. *Human Factors*, *53*(2), 168–179. https://doi.org/10.1177/0018720811400601

Li, J., & Kim, J.-E. (2021). The Effect of Task Complexity on Time Estimation in the Virtual Reality Environment: An EEG Study. *Applied Sciences*, *11*(20), Article 20. https://doi.org/10.3390/app11209779

Luque, F., Armada, V., Piovano, L., Jurado-Barba, R., & Santamaría, A. (2024). Understanding Pedestrian Cognition Workload in Traffic Environments Using Virtual Reality and Electroencephalography. *Electronics*, *13*(8), 1453. https://doi.org/10.3390/electronics13081453

Matthews, G., Reinerman-Jones, L., Abich, J., & Kustubayeva, A. (2017). Metrics for individual differences in EEG response to cognitive workload: Optimizing performance prediction. *Personality and Individual Differences*, *118*, 22–28. https://doi.org/10.1016/j.paid.2017.03.002

Michels, L., Moazami-Goudarzi, M., Jeanmonod, D., & Sarnthein, J. (2008). EEG alpha distinguishes between cuneal and precuneal activation in working memory. *NeuroImage*, *40*(3), 1296–1310. https://doi.org/10.1016/j.neuroimage.2007.12.048

Miller, S. (2001). Workload measures. *National Advanced Driving Simulator. Iowa City, United States*.

Mondellini, M., Pirovano, I., Colombo, V., Arlati, S., Sacco, M., Rizzo, G., & Mastropietro, A. (2023). A Multimodal Approach Exploiting EEG to Investigate the Effects of VR Environment on Mental Workload. *International Journal of Human–Computer Interaction*, *0*(0), 1–13. https://doi.org/10.1080/10447318.2023.2258017

Morton, J., Zheleva, A., Van Acker, B. B., Durnez, W., Vanneste, P., Larmuseau, C., De Bruyne, J., Raes, A., Cornillie, F., Saldien, J., De Marez, L., & Bombeke, K. (2022). Danger, high voltage! Using EEG and EOG measurements for cognitive overload detection in a simulated industrial context. *Applied Ergonomics*, *102*, 103763. https://doi.org/10.1016/j.apergo.2022.103763

Neira-Tovar, L., Castañeda-Mancillas, E. E., Rossano, G., Cupich, J., & Cruz, H. (2022). Implementation of Virtual Reality for Training in Education and Industry: A Forklift Case. In J. Trojanowska, A. Kujawińska, J. Machado, & I. Pavlenko (Eds.), *Advances in Manufacturing III* (pp. 292–305). Springer International Publishing. https://doi.org/10.1007/978-3-030-99310-8_23

Örün, Ö., & Akbulut, Y. (2019). Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. *Computers in Human Behavior*, *92*, 216–229. https://doi.org/10.1016/j.chb.2018.11.027

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, *40*(1), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3

Puma, S., Matton, N., Paubel, P.-V., Raufaste, É., & El-Yagoubi, R. (2018). Using theta and alpha band power to assess cognitive workload in multitasking environments. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *123*, 111–120. https://doi.org/10.1016/j.ijpsycho.2017.10.004

Ray, W. J., & Cole, H. W. (1985). EEG activity during cognitive processing: Influence of attentional factors. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *3*(1), 43–48. https://doi.org/10.1016/0167-8760(85)90018-2

Rihs, T. A., Michel, C. M., & Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by alpha-band EEG synchronization. *The European Journal of Neuroscience*, *25*(2), 603–610. https://doi.org/10.1111/j.1460-9568.2007.05278.x

Rodenburg, D., Hungler, P., Etemad, S. A., Howes, D., Szulewski, A., & Mclellan, J. (2018). Dynamically adaptive simulation based on expertise and cognitive load. *2018 IEEE Games, Entertainment, Media Conference (GEM)*, 1–6. https://doi.org/10.1109/GEM.2018.8587618

Scrivener, C. L., & Reader, A. T. (2022). Variability of EEG electrode positions and their underlying brain regions: Visualizing gel artifacts from a simultaneous EEG-fMRI dataset. *Brain and Behavior*, *12*(2), e2476. https://doi.org/10.1002/brb3.2476

Senftleben, U., & Scherbaum, S. (2021). Mid-frontal Theta during Conflict in a Value-based Decision Task. *Journal of Cognitive Neuroscience*, *33*(10), 2109–2131. https://doi.org/10.1162/jocn_a_01741

Shou, G., & Ding, L. (2013). Frontal theta EEG dynamics in a real-world air traffic control task. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, *2013*, 5594–5597. https://doi.org/10.1109/EMBC.2013.6610818

Slobounov, S. M., Ray, W., Johnson, B., Slobounov, E., & Newell, K. M. (2015). Modulation of cortical activity in 2D versus 3D virtual reality environments: An EEG study. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *95*(3), 254–260. https://doi.org/10.1016/j.ijpsycho.2014.11.003

Tran, Y., Craig, A., Craig, R., Chai, R., & Nguyen, H. (2020). The influence of mental fatigue on brain activity: Evidence from a systematic review with meta-analyses. *Psychophysiology*, *57*(5), e13554. https://doi.org/10.1111/psyp.13554

Ulutas, B. H., & Ozkan, N. F. (2019). Assessing occupational risk factors for forklift drivers. *Le Travail Humain*, *82*(2), 129–149.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Yang, L., Guan, W., Ma, R., & Li, X. (2019). Comparison among driving state prediction models for car-following condition based on EEG and driving features. *Accident Analysis & Prevention*, *133*, 105296. https://doi.org/10.1016/j.aap.2019.105296

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics*, *58*(1), 1–17. https://doi.org/10.1080/00140139.2014.956151

Yuen, K. K., Choi, S. H., & Yang, X. B. (2010). A Full-immersive CAVE-based VR Simulation System of Forklift Truck

Operations for Safety Training. *Computer-Aided Design and Applications*, *7*(2), 235–245. https://doi.org/10.3722/cadaps.2010.235-245

Yurko, Y. Y., Scerbo, M. W., Prabhu, A. S., Acker, C. E., & Stefanidis, D. (2010). Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, *5*(5), 267–271. https://doi.org/10.1097/SIH.0b013e3181e3f329