Proceedings of the Human Factors and Ergonomics Society Annual Meeting
2025, Vol. 69(1) 1435–1440

DOI: 10.1177/10711813251363210
journals.sagepub.com/home/pro
Sage

Investigating Cognitive Engagement from Training to Application Under Varied Workload Manipulations in Virtual Reality

Cara A. Spencer[1], Jalynn B. Nicoly[1], Lucas Plabst[2], Rebecca L. Pharmer[2], Christopher D. Wickens[2], Francisco Ortega[2], Benjamin A. Clegg[3], Joanna E. Lewis[4], and Leanne Hirshfield[1]

Abstract

This study examined how neural engagement differs between training and application under varied cognitive workload in a virtual-reality assembly task. Twenty-three participants completed four within-subject conditions (low vs. high intrinsic load and low vs. high extraneous load). We recorded three electroencephalography-derived engagement indicators: Cognitive Effort (parietal α/frontal θ), Sustained Attention ($\beta/(\alpha + \theta)$), and Integration & Execution (γ power). Changes in these indicators during learning were expected to reveal changes in engagement linked to unique cognitive resources recruited by workloads and phases of learning tasks. Analyses revealed a robust task phase effect: application activation exceeded training for Sustained Attention and Integration & Execution. Significant task phase workload condition interactions emerged for these indicators as well. Findings indicated that unguided application amplifies working-memory and integrative engagement and that this amplification depends on workload type. Our results inform the design of adaptive training systems monitoring phase-dependent neural engagement to optimize learning transfer.

Keywords

training, virtual reality, EEG, workload, engagement

Introduction

Imagine a novice technician completing a virtual reality (VR) training module to assemble a spacecraft panel. During training, they're guided step-by-step with visual cues and immediate feedback. Later, in a real mission context, they must reconstruct that same process from memory. Performance depends not just on task repetition, but on whether the training phase meaningfully engaged the cognitive resources needed for transfer to application. Did the new technician simply follow directions, or were they mentally invested enough to integrate and retain what they learned?

Effective training demands that learners exert cognitive effort, but also that they can retain and apply what they've learned when support is removed. Cognitive engagement has been identified as an orchestrator of that effort (i.e., cumulative cognitive resources) involved in learning and performance on the task learned (Dehais et al., 2020). We characterize engagement as the degree to which learners are actively focused, involved, and motivated during a learning task (Newman et al., 1992). Defining how and when learners engage remains a challenge in designing training materials, especially for domains like aerospace, healthcare, and manufacturing, where skills must transfer reliably from structured practice to unsupported application. Capturing engagement could, however, allow training instructors to improve instructional design and build adaptive learning environments (Dehais et al., 2020; Wickens et al., 2012).

Several aspects of engagement's role in learning remain unclear. First, it is unknown if commonly used indicators of engagement reflect the same cognitive processes during different task phases such as training, when guidance and feedback are present, versus transfer, retention, or application,

[1]University of Colorado, Boulder, USA
[2]Colorado State University, Fort Collins, USA
[3]Montana State University, Bozeman, USA
[4]University of Northern Colorado, Greeley, USA

Corresponding Author:
Cara A. Spencer, University of Colorado, 1777 Exposition Drive, Boulder 80309-0401, USA.
Email: cara.spencer@colorado.edu

when learners must perform the task by relying on memory and self-direction. This is further complicated when the nature of training imposes different types of cognitive demands on the learner. Due to poor design, resource restrictions or other reasons, real learning environments can vary in intrinsic load (i.e., how complex the task to be learned is inherently) and extraneous load (i.e., non-task relevant aspects to process) (Sweller, 1998). Research suggests that extraneous load, which stems from the presentation or format of instructional materials, tends to tax posterior brain areas associated with visual and spatial processing (Beauchemin et al., 2024; Makransky et al., 2019). In contrast, intrinsic load, arising from the inherent complexity of the material itself, engages frontal regions responsible for executive functions and working memory (Dehais et al., 2020; McDonnell et al., 2023).Both types of workloads have been found to interact with engagement but its relationship to task phases (i.e., training versus application) is not well understood (Bueno-Vesga et al., 2021; Dehais et al., 2020). If there is a relationship between the two, this could have implications for training design and adaptive systems.

Several indicators of engagement based on electroencephalography (EEG) have been developed that offer real-time neural markers useful for addressing these gaps (Apicella et al., 2022; Beauchemin et al., 2024; Halderman et al., 2021; McDonnell et al., 2023; Pope et al., 1995). EEG is a noninvasive sensor placed on the scalp to record electrical brain activity. Human factors research has utilized EEG for decades due to its temporal precision, cost effectiveness, and sensitivity to rapid changes in brain activity (Rahman et al., 2019). Further, EEG has been shown to be quickly informative and minimally obtrusive when used with other technology like VR headsets (Makransky et al., 2019). The minimal spatial specificity in the local field potential (i.e., 3–4 cm) also provides some flexibility on where exactly electrodes are placed to capture regional signals, improving practicality for real world applications.

EEG indicators traditionally capture engagement through distinct spectral dynamics reflecting different cognitive subprocesses. Specifically, EEG engagement indicators in prior literature include: **(1) Cognitive Effort.** EEG features include the ratio of (parietal α/frontal θ), as used during driving in different difficulty conditions (McDonnell et al., 2023). Their findings suggest that the Cognitive Effort indicator decreases with higher engagement and increases with higher workload due to the inverse relationship between the alpha and theta band. **(2) Sustained Attention.** EEG features include the average $(\beta)/(\alpha + \theta)$ ratio across channels, as used while learning to use an interface (Apicella et al., 2022) and flying with an adaptive autopilot (Pope et al., 1995), and **(3) Information Integration and Execution.** This EEG feature is the z-transformed log power of γ-band along three midline electrodes, as used while taking a standardized test (Halderman et al., 2021). Importantly, these engagement indicators have roots in different cognitive resources, and we

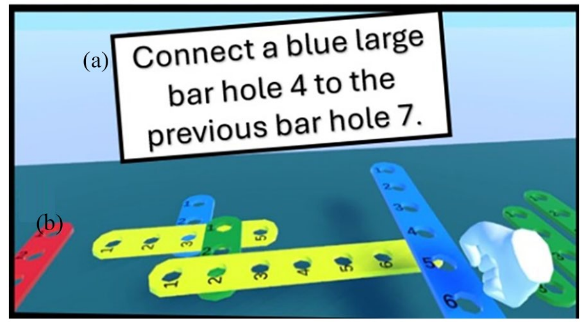

Figure 1. VR environment. Manipulations include (a) extraneous load (depicted here is the low load condition) and (b) intrinsic load (depicted here is high intrinsic load).

expect that their collective changes during learning will illuminate possible changes in engagement with more sensitivity than just using one indicator.

The objectives of the current study are thus to identify the impact of task phase on engagement, and to investigate if task phase interacts with different types of cognitive workload, and if EEG indicators of engagement are sensitive to these interactions. Participants underwent training on how to assemble a complex shape in an adaptive virtual reality (VR) environment while levels of **extraneous and intrinsic workload** were manipulated. As shown in Figure 1, participants read instructions and place bars on a workbench in VR, with successive steps taken to build out a given shape. Extraneous load manipulations included adding superfluous text into the instructions (Figure 1a Intrinsic load manipulations included the number of relevant bar characteristics (Figure 1b, as color, size, and hole number).

Following their training, they were instructed to assemble the object from memory. Thus, there was a **training phase** and, following a short delay, an **application phase** to the study. Meanwhile, EEG was recorded to track changes in engagement, and EEG features were extracted to investigate the Cognitive Effort, Sustained Attention, and Integration and Execution indicators described previously. In our current study, we seek to investigate how the EEG engagement indicators (Cognitive Effort, Sustained Attention, and Information and Execution) are differentially sensitive to the load manipulations in our study (intrinsic vs. extraneous load), as well which task phase (training vs. application phase) participants are in.

We expected the training process to prolong the use of attentional resources while an individual follows along with external guidance. Furthermore, the generative, integrative, and working-memory based operations during application are predicted to increase the cognitive effort and integration and execution indicators higher than during training. We thus hypothesize that (H1) the training phase will invoke an amplified Sustained Attention indicator response compared to application phase, and (H2) the other indicators of engagement are expected to amplify during application more than training. Prior literature has indicated a relationship between workload and engagement (Dehais et al., 2020) motivating hypothesis (H3), that changes in engagement will vary depending on the interaction between task phase *and* type of workload.

Method

Twenty-three participants engaged in a VR-based shape assembly task (shown in Figure 1) with systematically manipulated levels of intrinsic and extraneous workload. The ages in the sample ranged from 18 to 49 (M=24 years, SD=9.95) and biological sex was split 47% females (N=17), 53% males (N=19), and 0% intersex. Compensation for participation was given to all participants. Research procedures were approved by the University's Institutional Review Boards and each participant was consented.

There were two task phases: **training** and **application**. In the training phase, participants were provided step-by-step instructions to build an object using a set of five differently sized and colored bars. Once they completed the assembly with all of the instructions, they repeated the assembly with the last step's instruction omitted requiring them to rely on memory to complete the last step. This instruction omission was incrementally increased over the course of six assemblies. Each subsequent assembly had one less instruction until the last assembly where no instructions were given. The program provided a green check mark upon successful step completion or a red "X" upon an error. In addition to the "X," the step's instruction was shown again, and the participant could reference it before trying the step again. Thus, the training increased difficulty (reliance on memory, rather than perception) in an adaptive fashion (Landsberg, 2012). Following training, participants were given a 90-s distractor task and 15-s rest period. In the following application phase, they were tested on their ability to build the shape without any instructions or feedback at any point. An assembly task was chosen because of its generalizability to a wide range of real-world occupations that involve visual, procedural learning.

Two types of cognitive workload were manipulated using elements of the learning environment (e.g., superfluous task instructions, details of bars required), as shown in Figure 1. All manipulations took place only during training due to their expected impact on learner performance and the desire for the application task phase to be uniform across all conditions.

EEG Apparatus and Processing

EEG data were collected using a BrainVision LiveAmp eight channel device connected to Lab Recorder software. We selected three EEG electrode locations at Fz, Cz, and Pz in the 10 to 20 system because these were three commonly used regions in our review of engagement studies. Reference and ground electrodes were placed at PO9 and PO10, respectively. We prioritized a smaller system and number of electrodes to better apply to real-world field training since unobtrusiveness is advantageous. The VR headset position under the eye prevented the use of electrooculography and thus, independent component analysis was used to identify and remove eye movement related artifacts (Mennes et al.,

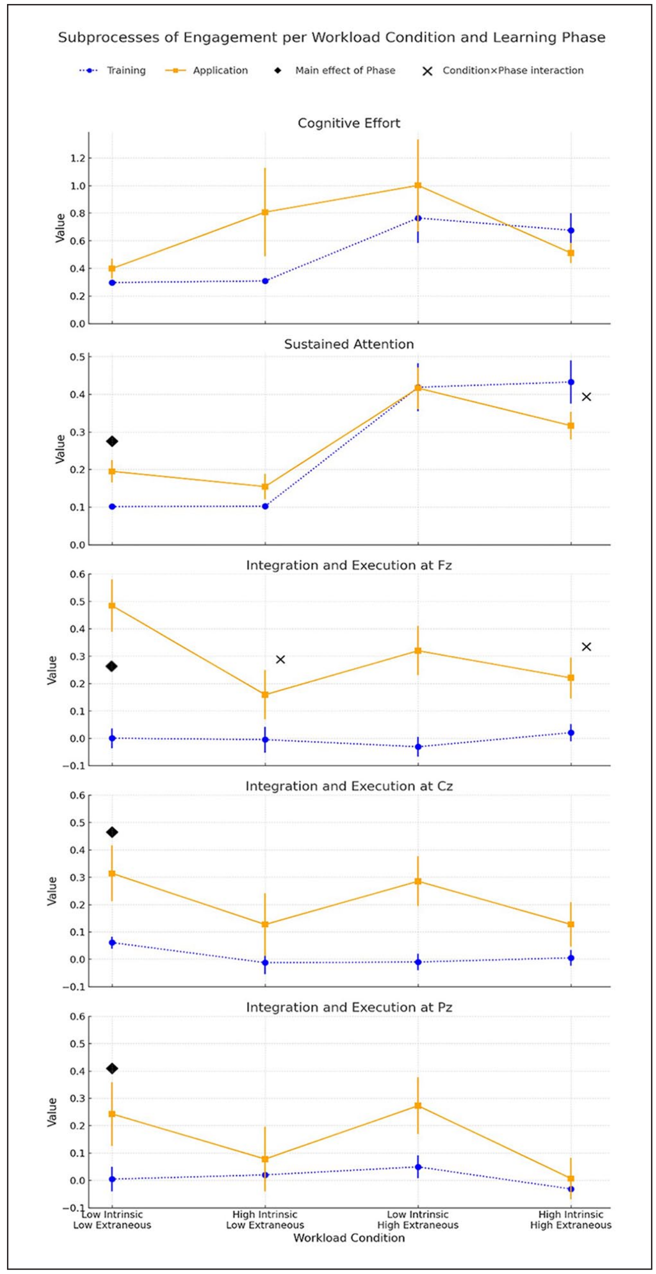

Figure 2. Depiction of main effects and interactions for each EEG indicator of engagement for each workload condition and task phase. Each plot shows the activation from each indicator's formulation.

2010; Sumathi et al., 2023). The EEG signal was processed and analyzed using Python MNE library (Larson et al., 2024). The signal was downsampled to 250 Hz, bandpass filtered from 0.1 Hz to 30 Hz, and epoched into 1 s intervals. Available rest intervals or a default 15 s baseline from the start of the EEG recording were used for baseline correction. Training and application phase data were separated for each condition.

In addition to the EEG measurements, participant performance (i.e., number of correctly assembled shapes and training time) was recorded throughout the task. Participants filled out a NASA-TLX workload self-report after each workload condition (Hart & Staveland, 1988). These data were used to confirm that the manipulations were indeed effective.

Results

Self-report survey (NASA-TLX) data and performance measures confirmed that manipulating intrinsic and extraneous load impacted participants as expected. Specifically, higher intrinsic load led to increased self-reported cognitive workload, longer training and applications times, and increased training error rates. In contrast, increasing extraneous load showed a limited effect of only increasing training time.

EEG Results

For the EEG hypothesis testing, we ran an analysis of variance (ANOVA) and fit a linear mixed-effects model (LME; lmerTest in R) for each engagement indicator: Cognitive Effort, Sustained Attention, and Integration and Execution at Fz, Cz, and Pz. Fixed effects included: **workload condition** (four levels: defined by low and high intrinsic and extraneous load); **task phase** (training vs. application), and their **interaction.** A random intercept for Subject was included. Task phase and workload condition were coded as treatment contrasts, and Satterthwaite's method was used to approximate denominator degrees of freedom in an ANOVA. Results for each engagement index are described below and are depicted in Figure 2. In line with the purpose of the study, only the main effects of task phase and interactions between task phase and each of the four workload conditions are reported. When there was a significant interaction, then a significant Condition main effect was also included (ANOVA results then the LME contrasts).

Cognitive Effort. The ANOVA revealed no significant main effect of Phase, $F(1, 161) = 2.07$, $p = .152$, $\eta^2 = .01$, nor a Condition × Phase interaction, $F(3, 161) = 1.39$, $p = .247$, $\eta^2 = .03$. The contrast estimates confirmed that Training versus Application did not significantly differ: $b = 0.102$, $SE = 0.234$, 95% CI [−0.357, 0.560], $t(161) = 0.43$, $p = .665$. No interaction contrasts reached significance (all $p > .23$).

Sustained Attention. The ANOVA did not reveal a task phase effect $F(1, 161) = 0.10$, $p = .755$, $\eta_p^2 < .001$. However, Condition exerted a large main effect, $F(3, 161) = 44.24$, $p < .001$, $\eta_p^2 = .45$ and there was a reliable Condition × Phase interaction, $F(3, 161) = 4.06$, $p = .008$, $\eta_p^2 = .07$ and the Training versus Application contrast showed a modest Phase effect, $b = 0.094$, $SE = 0.045$, 95% CI [0.006, 0.183], $t(161) = 2.08$, $p = .039$. The interaction was driven by a decrease during the high-intrinsic/high-extraneous condition $b = -0.210$, $SE = 0.064$, 95% CI [−0.335, −0.085], $t(161) = -3.29$, $p = .001$.

Integration & Execution at Fz. A strong Phase effect was found of greater engagement during application than training $F(1, 161) = 42.81$, $p < .001$, $\eta^2 = .21$, with a marginal interaction, $F(3, 161) = 2.58$, $p = .055$, $\eta^2 = .05$. The Phase contrast was highly significant: $b = 0.485$, $SE = 0.092$, 95% CI [0.305, 0.664], $t(161) = 5.28$, $p < .001$. The significant interaction revealed that workload affected this indicator more during application than during training. Significant contrasts between the phases appeared when intrinsic load was high and extraneous load was both low ($b = -0.321$, $SE = 0.130$, 95% CI [−0.575, −0.066], $t = -2.47$, $p = .015$) and high ($b = -0.2845$, $SE = 0.130$, 95% CI [−0.539, −0.030], $t = -2.19$, $p = .030$).

Integration & Execution at Cz. A clear Phase effect, $F(1, 161) = 17.32$, $p < .001$, $\eta_p^2 = .10$; Condition × Phase was non-significant, $F(3, 161) = 0.76$, $p = .517$, $\eta_p^2 = .01$, although these appeared to be in the same direction as observed at Fz. The Phase contrast confirmed the effect: $b = 0.253$, $SE = 0.097$, 95% CI [0.062, 0.443], $t(161) = 2.60$, $p = .01$.

Integration & Execution at Pz. The Phase effect was modest, $F(1, 161) = 6.87$, $p = .010$, $\eta_p^2 = .04$; no interaction was found, $F(3, 161) = 0.99$, $p = .398$, $\eta_p^2 = .02$. The Phase contrast concurred: $b = 0.238$, $SE = 0.106$, 95% CI [0.030, 0.447], $t(161) = 2.24$, $p = .026$. Again, these trended in the same direction as observed at Fz, reflecting greater engagement during applications than during training.

Discussion

Our results demonstrate a clear shift in neural engagement when learners move from guided training to unguided application in a VR assembly task. Both Sustained Attention and Integration and Execution EEG engagement indicators were significantly lower during the **training** phase compared to **application.** For the Sustained Attention indicator, this finding was in contradiction of H1. There was no significant effect on the Cognitive Effort indicator which did not follow the prediction in H2. The Integration and Execution Index, however, did show significantly higher activation during application in support of H2. Finally, two indicators revealed that the cognitive workload condition did indeed interact with task phase partially supporting H3 (i.e., Sustained Attention and Integration and Execution).

While the Sustained Attention and the Integration and Execution engagement indicators were reliably stronger in the application phase, there were two instances where this was not true. First, a significant interaction between condition and phase signaled higher Sustained Attention during training in the single condition where both intrinsic and extraneous load were high. Second, while not significant, the Cognitive Effort indicator trended toward this switch to more effort during training than application also when intrinsic and extraneous load were high. Taken together with the significant Sustained Attention results, this may indicate the workload manipulations were exerting high demands that drove competition for cognitive resources and attention was dominant. This lends support for H3, that changes in engagement would occur differently depending on the type of load which is supported by prior literature on the relationship between workload and engagement.

The Integration and Execution at Fz indicator results suggest it was the most sensitive to task phase and workload condition. Notably, the indicator had a large increase moving from training to application when both intrinsic load and extraneous load were low. This increase potentially highlights the cognitive adaptation to a steep increase in task demands. In the low workload condition, the training instructions were straightforward and the working memory demands were low resulting in low engagement. In contrast, the application phase, although matched in objective complexity across conditions, demanded *relatively* more mobilization of integration and execution resources. This is perhaps due to the Contrast Effect, a cognitive bias that refers to the tendency to evaluate or respond based on prior context (Baumeister, 2007). This pattern was not observed at Cz or Pz, suggesting potential spatial specificity to the prefrontal cortex which typically hosts these functions and warrants further investigation through source localization. Future work should also investigate if this relative shift in engagement has impacts on behavioral performance during application.

These findings align with Cognitive Load Theory, which distinguishes between intrinsic, extraneous, and germane load (Sweller, 1988). Germane load refers to the mental effort expended during learning (e.g., developing schemas, forming connections, bolstering long-term retention). The Integration and Execution indicator's pronounced increases during application may reflect heightened germane load; as predicted, learners must internally integrate and sequence multiple task elements without external cues. By contrast, the more modest boost in Sustained Attention suggests that filtering demands imposed by verbose instructions (extraneous load) exerted their maximal effect during training. Cognitive Effort under the highest combined loads showed increased effort during training and increased engagement during application, following expected trends from this theory.

Practically, these neural markers could power task phase-aware adaptive training systems. For example, real-time monitoring of γ-band power could signal when to withdraw scaffolding or provide richer guidance: a steep rise in γ might indicate that a learner is successfully engaging germane processes, whereas stagnation could trigger additional support. Similarly, sustained-attention indices could help identify when instructional materials become too verbose, prompting user interface adjustments. Integrating these signals into VR platforms may enable dynamic calibration of both task complexity and instructional clarity to optimize learning transfer.

Several limitations invite future work. Our low-density EEG montage (Fz/Cz/Pz), while advantageous for field applications, may miss lateral or deeper sources; expanded coverage and source analyses could clarify regional contributions. The object-assembly paradigm emphasizes visuospatial integration, and other tasks (e.g., verbal reasoning) may yield different engagement profiles. The null effect for Cognitive Effort suggests floor/ceiling concerns or ratio insensitivity—alternative θ or α metrics or machine-learning classifiers might capture subtler effort shifts. Exploring individual differences (e.g., working-memory capacity) and combining EEG with other neuroimaging or eye-tracking could refine adaptive systems' responsiveness. By pinpointing when and how engagement subprocesses change, we can design smarter training environments that support learners through both practice and application.

Conclusion

The current study sought to characterize changes in cognitive engagement as individuals complete training and then apply their new learned skill using EEG. Additionally, we explored how cognitive workload interacted with these task phases. We used four spectral bands in different formulations established by prior work as indicators for engagement. Our findings suggest that different EEG indicators capture distinct dimensions of engagement as learners progress through structured practice to autonomous performance. In most cases, engagement was higher during the application of a newly learned procedure than during training. There were interactions between intrinsic and extraneous workload and task phase highlighting important implications for adaptive training and evaluation design.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Natalie Steinhauser and the Office of Naval Research conducted under grant numbers ONR N00014-24-1-2214 and ONR N00014-23-1-2298.

ORCID iDs

Cara A. Spencer https://orcid.org/0000-0002-8743-6960
Joanna E. Lewis https://orcid.org/0000-0003-0974-2230

References

Apicella, A., Arpaia, P., Frosolone, M., Improta, G., Moccaldi, N., & Pollastro, A. (2022). EEG-based measurement system for monitoring student engagement in learning 4.0. *Scientific Reports*, *12*(1), Article 5857.

Baumeister, R. F. (2007). *Encyclopedia of social psychology*. Sage.

Beauchemin, N., Charland, P., Karran, A., Boasen, J., Tadson, B., Sénécal, S., & Léger, P. M. (2024). Enhancing learning experiences: EEG-based passive BCI system adapts learning speed to cognitive load in real-time, with motivation as catalyst. *Frontiers in Human Neuroscience*, *18*, Article 1416683.

Bueno-Vesga, J. A., Xu, X., & He, H. (2021, March). *The effects of cognitive load on engagement in a virtual reality learning environment* [Conference session]. IEEE Virtual Reality and 3D User Interfaces (VR), Lisboa, Portugal (pp. 645–652). IEEE.

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. *Frontiers in Neuroscience*, *14*, Article 519228.

Halderman, L. K., Finn, B., Lockwood, J. R., Long, N. M., & Kahana, M. J. (2021). EEG correlates of engagement during assessment. *ETS Research Report Series*, *2021*(1), 1–17.

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.), *Human mental workload*. (Vol. 52, pp. 139–183) North Holland Press.

Landsberg, C. R., Astwood, R. S., Jr., Van Buskirk, W. L., Townsend, L. N., Steinhauser, N. B., & Mercado, A. D. (2012). Review of adaptive training system techniques. *Military Psychology*, *24*(2), 96–113.

Larson, E., Gramfort, A., Engemann, D. A., Leppakangas, J., Brodbeck, C., Jas, M., Brooks, T. L., Sassenhagen, J., McCloy, D., Luessi, M., King, J.-R., Höchenberger, R., Goj, R., Brunner, C., Favelier, G., van Vliet, M., Wronkiewicz, M., Rockhill, A., Holdgraf, C., . . . luzpaz. (2024). *MNE-Python (v1.9.0)*. Zenodo.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236.

McDonnell, A. S., Simmons, T. G., Erickson, G. G., Lohani, M., Cooper, J. M., & Strayer, D. L. (2023). This is your brain on autopilot: Neural indices of driver workload and engagement during partial vehicle automation. *Human Factors*, *65*(7), 1435–1450.

Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., & Stiers, P. (2010). Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. *Psychophysiology*, *47*(6), 1142–1150.

Newman, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. M. Newman (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11–39). Teachers College Press.

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

Rahman, M., Karwowski, W., Fafrowicz, M., & Hancock, P. A. (2019). Neuroergonomics applications of electroencephalography in physical activities: A systematic review. *Frontiers in human neuroscience*, *13*, Article 182.

Sumathi, K., Dhanavarsha, N., Gayathri, S. V., & Pavithra, N. (2023, December 14–15). *Ocular artifact removal using independent component analysis* [Conference session]. Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), Chennai, India.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.

Wickens, C. D., Hutchins, S., Carolan, T., & Cumming, J. (2012). Effectiveness of part-task training and increasing-difficulty training strategies. *Human Factors*, *55*(2), 461–470.