



## Machine Learning-Based Cognitive Load Prediction from Multimodal Neurophysiological Data

Dr. V. Maniraj<sup>1\*</sup>, N. Mahalaksmi<sup>2</sup>

<sup>1</sup>Principal (i/c) Associate Professor, Department of Computer Science, A.V.V.M.Sri Pushpam College(Autonomous), Poondi, Thanjavur(Dt), Affiliated to Bharathidasan University, Thiruchirappalli, Tamilnadu

<sup>2</sup>Department of Computer Applications, A.V.V.M.Sri Pushpam College(Autonomous),Poondi, Thanjavur(Dt), Affiliated to Bharathidasan University, Thiruchirappalli, Tamilnadu

\*Corresponding author, manirajv61@gmail.com

DOI: <https://doi.org/10.63680/ijate0326124.095>

### Abstract

A better understanding of students' cognitive load is essential to optimize instruction. However, it could be challenging in the classroom, as it is difficult for teachers to observe each student continuously when dozens of students co-exist. The rise of multimodal learning analytics now allows for predictions of individuals' cognitive load using neurophysiological features. Therefore, this research aimed to equip teachers with insights into students' cognitive load in high school math classes by leveraging neurophysiological data. Results highlighted that the teacher's cognitive load assessments for the overall class do not generalize well to individual students, especially on mental effort. In contrast, multimodal neurophysiological models demonstrated a notable correlation, accounting for 15.3% of perceived task difficulty and 25.9% of mental effort variance ( $R^2$ ), providing additional information for teacher assessments. Our study demonstrates that multimodal neural correlates may help teachers gain a deeper understanding of students' cognitive states, enhancing instructional design accordingly.

**Keywords:** Cognitive load; multimodal learning analytics; neurophysiological data; mental effort assessment; instructional optimization

### Introduction

Cognitive load, defined as "the load imposed on an individual's working memory by a specific learning task" (van Gog & Paas, 2012), is a crucial psychological state in the learning process. Due to the limited capacity of working memory, students with a high level of cognitive load may struggle to retain all necessary materials in mind, leading to poor performance. Therefore, evaluating students' cognitive load is important for teachers, as they can thereby optimize instructional design to improve learning efficiency (Paas et al., 2003).

In practice, experienced teachers often evaluate whether students are overloaded by observing their facial expressions and behaviors. Nevertheless, as there are usually dozens of students in the classroom, teachers

may focus more on class-level rather than individualized states (Young, 2020). These observations may not well capture the cognitive load dynamics of individual students, which may vary greatly depending on their knowledge base. Moreover, cognitive load is implicit in human cognitive architecture, and may not always be explicitly revealed in behaviors (Paas & Van Merriënboer, 1994). This is likely more common in certain cultural contexts, where students are usually shy to express feeling difficult when they experience high cognitive load.

With the advancements in portable biosensors, multimodal learning analytics provide a new lens to overcome the limitations of classroom observations (Blikstein, 2013). These sensors can capture individualized dynamics through continuous neurophysiological recordings. Additionally, this approach is capable of detecting implicit activities underlying cognitive processes. Substantial evidence has indicated that cognitive load changes can lead to changing physiological responses in central and peripheral nervous systems (Ayres et al., 2021; Callister et al., 1992), which manifest through various biosignals including electroencephalograph (EEG) recordings (Curtis & D'Esposito, 2003), electrodermal activity (EDA) (Setz et al., 2009), and photoplethysmography (PPG) signals (Lyu et al., 2015). Therefore, it has become promising to predict cognitive load through these neural correlates. The predicted values, thereby, could potentially provide additional, valuable information for teachers to know whether the instructional design is overloaded for students.

Based on the issues discussed, we have formulated the following research questions: RQ1: Do the teacher's assessments of class-level cognitive load align with students' individualized states? RQ2: Can multimodal neurophysiological data assist teachers in understanding their students' cognitive load? Therefore, we collected data on cognitive load assessments during actual math classroom sessions from both students and the teacher, along with neurophysiological recordings, to investigate the aforementioned questions.

## **Methods**

### **Dataset**

Twenty-two students (9 females and 13 males, aged 15-16) and one teacher (female) in a high school were recruited for the data collection. Data collection spanned three weeks in March 2023, involving thirteen 40-minute high school math lessons. The experience sampling method (ESM) (Dykstra & Paul, 2018) was implemented by administering three surveys in each lesson, dividing it into three segments, each lasting approximately 10-15 minutes. In the ESM self-report survey, students were asked to rate the degree of perceived task difficulty and mental effort (Ayres, 2006; Camp et al., 2001; Paas, 1992) over the past 10 minutes, using the 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). These measures are frequently used as indicators of cognitive load in prior research. The survey also requested the teacher to assess the cognitive load experienced by most students using similar items, as detailed in Table 1. Considering the implicit nature of mental effort, the teacher's assessment was grounded in a more observable concept: engagement (Finn & Zimmer, 2012).

**Table 1**

*Items of the Survey (5-point Likert scale)*

Items	The teacher's evaluation	The students' evaluation
Task difficulty	<i>The learning material was difficult for the majority of students.</i>	<i>The learning material was difficult for me</i>
Mental effort	<i>The majority of students were well engaged in the classroom learning.</i>	<i>I invested high mental effort.</i>

**Neural correlates of cognitive load**

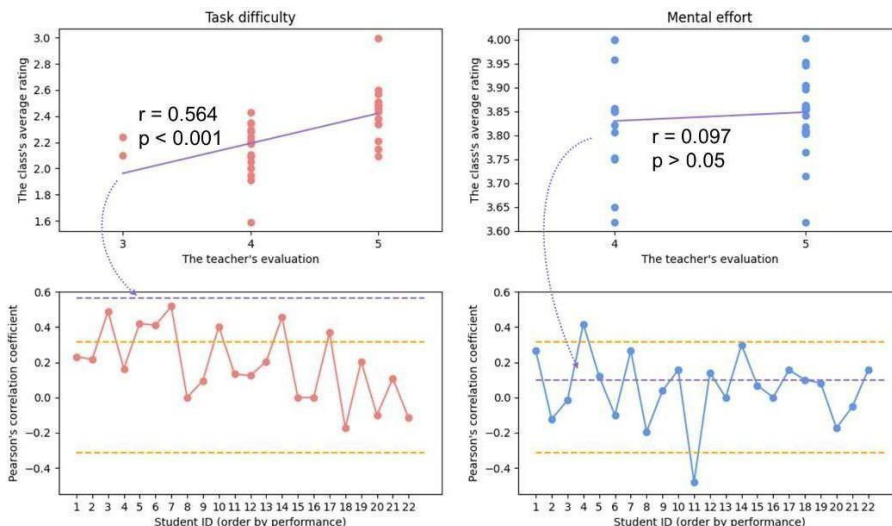
Throughout the lessons, all participants wore head belts and wristwatches to concurrently collect EEG (from the Fp1 and Fp2 channels), EDA, and PPG signals. For the EEG data, the absolute and relative spectral band powers of delta (1 Hz – 4 Hz), theta (4 Hz – 8 Hz), alpha (8 Hz – 13 Hz), beta (13 Hz – 30 Hz), and gamma (30 Hz – 45 Hz) bands were calculated after signal preprocessing (Koley & Dey, 2012). Regarding the EDA signal, means and standard deviations of the skin conductance response (SCR) and skin conductance level (SCL) were extracted, representing the rapidly changing phasic component and the slowly changing tonic component of the EDA signal, respectively (Boucsein, 2012). Heart rate (HR) and pulse rate variability (PRV) were estimated using the PPG signal (Qi et al., 2020). The extracted features were normalized and subsequently applied in a support vector regression model to predict cognitive load. To evaluate the model's performance, 10-fold cross-validation was employed, and metrics including mean absolute error (MAE), root mean squared error (RMSE), and R2 were calculated.

**Results**

RQ1: Do the teacher's assessments of class-level cognitive load align with students' individualized states?

**Figure 1**

*Correlation between the Teacher's Evaluation and Students' Cognitive Load*



Students rated their perceived task difficulty with an average score of  $2.27 \pm 0.95$ , and their mental effort at  $3.84 \pm 0.81$  ( $n = 813$ ). In contrast, the teacher evaluated the task difficulty at  $4.36 \pm 0.58$ , and the students' mental effort at  $4.64 \pm 0.49$  ( $n = 39$ ). Beyond the numerical discrepancies, we also explored the correlations between these ratings. As shown in Figure 1, there was a significant correlation between the teacher's assessment of task difficulty and the average ratings of the students (Pearson's  $r = 0.564$ ,  $p < 0.001$ ,  $n = 39$ ), whereas it was not significant for mental effort (Pearson's  $r = 0.097$ ,  $p > 0.05$ ,  $n = 39$ ). We further examined the correlation between the teacher's ratings and those of each student. Only 7 out of 22 students demonstrated a significant correlation with the teacher regarding their ratings of task difficulty. Regarding mental effort, the teacher's assessments did not align with those of nearly all students. These findings suggest that the teacher's observations might not precisely capture the students' implicit cognitive load.

RQ2: Can multimodal neurophysiological data assist teachers in understanding their students' cognitive load?

Subsequently, a machine learning approach was utilized to predict students' self-report cognitive load by recognizing patterns within the neurophysiological features. The results, as shown in Table 2, indicate that our model accounts for 15.3% of the variability in students' perceived task difficulty and 25.9% of the variability in their mental effort ( $R^2$ ). However, when trained exclusively with data from either the head belt (EEG) or wristwatch (EDA + PPG), the model's performance showed a noticeable decline, underscoring the importance of multimodal fusion.

Due to individual differences among students, the teacher's estimation was not effective in predicting individualized cognitive load. Nevertheless, as demonstrated in Table 2, the integration of neurophysiological data and the teacher's evaluation enhanced the prediction accuracy of task difficulty for each student. When combining the two sources of data, was able to explain 16.4% of the variation in students' perceived task difficulty ( $R^2$ ), the value of information gain through data fusion.

**Table 2**

*Results of Predicting Cognitive Load with Different Modalities of Data ( $n = 813$ )*

Modality	Predicting "task difficulty"			Predicting "mental effort"		
	MAE	RMSE	$R^2$	MAE	RMSE	$R^2$
EEG	0.742	0.977	0.037	0.713	0.882	0.113
EDA+PPG	0.737	0.978	0.035	0.740	0.870	0.123
EEG+EDA+PPG	0.687	0.916	0.153	<b>0.651</b>	<b>0.858</b>	<b>0.259</b>
$T^\dagger$	0.796	0.990	0.019	0.823	1.000	0.000
EEG+T	0.731	0.969	0.050	0.718	0.888	0.107
EDA+PPG+T	0.742	0.974	0.041	0.742	0.872	0.121
EEG+EDA+PPG+T	<b>0.685</b>	<b>0.910</b>	<b>0.164</b>	0.661	0.862	0.253

*Note: MAE: mean absolute error; RMSE: root mean squared error; T: the teacher's evaluation  
 $\dagger$ : Estimated through ordinary least squares*

**Discussion**

Analysis of RQ1: Do teachers' assessment of class-level cognitive load align with students' individualized states? The disparity between teachers' assessments and students' subjective perceptions has long been an

issue in the field of education. In this study, we found a significant correlation between the teacher's and students' assessments of difficulty, demonstrating that the teacher has a relatively accurate understanding of the learning materials. However, there is a negative trend in the Pearson's correlation with regards to students' academic performance rankings, suggesting that the teacher was more focused on high-performing students when evaluating the perceived difficulty of students. For mental effort, a more implicit component of cognitive load, the teacher's assessment was much less effective. These results can be explained within the theoretical framework of cognitive load, which constructs it with factors related to tasks, subjects, and subject-task interactions (Paas & Van Merriënboer, 1994). While teacher observation may capture the factors related to tasks, the other factors are implicit and individualized, which are hard to be directly observed.

Analysis of RQ2: Can multimodal neurophysiological data assist teachers in understanding their students' cognitive load? Neurophysiological models showed remarkable higher accuracy than teacher observation. Notably, the model's performance on mental effort surpassed that on task difficulty, indicating the potential to capture implicit aspects of cognitive load. With a comparable metrics in the classroom setting (Gao et al., 2020), our model has demonstrated its feasibility in predicting students' cognitive load. Additionally, when considering different data modalities, we observed that constructing the model solely using data from wristwatches or head belts resulted in lower metrics compared to using all modalities. This indicates that certain combinations of neural correlates may provide a more accurate and sensitive prediction of cognitive load.

Despite the promising results, the study is subject to several limitations. First, the sample size was relatively small, particularly considering that the data was gathered from only one teacher, which could potentially limit the generalizability of our conclusions. Second, in real-world learning scenarios, various psychological activities coexist, which may result in similar neurophysiological responses. Therefore, the predicted values in a data-driven model may be closely associated with other similar constructs, like stress or fatigue. Subsequent research should control these covariates to achieve a more accurate prediction of cognitive load.

### **Declaration of Conflicting Interests**

The authors declare no potential conflicts of interest with respect to the research, authorship and publication of this article.

### **Funding**

The author received no financial support for the research, authorship and publication of this article.

### **References**

- Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(3), 287-298.
- Ayres, P., Lee, J. Y., Paas, F., & van Merriënboer, J. J. (2021). The validity of physiological measures to identify differences in intrinsic cognitive load. *Frontiers in Psychology*, 12, 702538.
- Blikstein, P. (2013). Multimodal learning analytics Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium. <https://doi.org/10.1145/2460296.2460316>

- Boucsein, W. (2012). *Electrodermal activity*. Springer Science & Business Media.
- Callister, R., Suwarno, N. O., & Seals, D. R. (1992). Sympathetic activity is influenced by task difficulty and stress perception during mental challenge in humans. *The Journal of physiology*, 454(1), 373-387.
- Camp, G., Paas, F., Rikers, R., & van Merriënboer, J. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort and mental efficiency. *Computers in Human Behavior*, 17(5-6), 575-595.
- Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in cognitive sciences*, 7(9), 415-423.
- Dykstra, J., & Paul, C. L. (2018). Cyber Operations Stress Survey (COSS): Studying fatigue, frustration, and cognitive workload in cybersecurity operations. 11th {USENIX} Workshop on Cyber Security Experimentation and Test ({CSET} 18).
- Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In *Handbook of research on student engagement* (pp. 97-131). Springer.
- Gao, N., Shao, W., Rahaman, M. S., & Salim, F. D. (2020). n-Gage: Predicting in-class Emotional, Behavioural and Cognitive Engagement in the Wild. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(3), Article 79. <https://doi.org/10.1145/3411813>
- Koley, B., & Dey, D. (2012). An ensemble system for automatic sleep stage classification using single channel EEG signal. *Computers in biology and medicine*, 42(12), 1186-1195.
- Lyu, Y., Luo, X., Zhou, J., Yu, C., Miao, C., Wang, T., Shi, Y., & Kameyama, K.-i. (2015). Measuring photoplethysmogram-based stress-induced vascular response index to assess cognitive load and stress. *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 857-866.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1), 1-4.
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *Journal of educational psychology*, 84(4), 429.
- Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational psychology review*, 6, 351-371.
- Qi, X., Wu, J., Shen, J., Ji, X., Lyu, Y., & Zhang, Y. (2020). Assessing cognitive load in adolescent and adult students using photoplethysmogram morphometrics. *Cognitive Neurodynamics*, 14(5), 709-721.
- Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., & Ehlert, U. (2009). Discriminating stress from cognitive load using a wearable EDA device. *IEEE Transactions on information technology in biomedicine*, 14(2), 410-417.
- van Gog, T., & Paas, F. (2012). Cognitive Load Measurement. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 599-601). Springer US. [https://doi.org/10.1007/978-1-4419-1428-6\\_412](https://doi.org/10.1007/978-1-4419-1428-6_412)
- Young, N. A. (2020). Getting the teacher's attention: Parent-teacher contact and teachers' behavior in the classroom. *Social Forces*, 99(2), 560-589.