Assessing Student Performance Through Pupil Dilation and Problem-Solving Time in Augmented Reality

Jung Hyup Kim[1], Siddarth Monhanty[1], Varun Pulipati[1], Fang Wang[1], Sara Mostowfi[1], Danielle Oprean[1], Yi Wang[1], and Kangwon Seo[1]

Abstract

This research aims to explore the prediction of student learning outcomes in Augmented Reality (AR) educational settings, focusing on engineering education, by analyzing pupil dilation and problem-solving time as key indicators. In this research, we have created an innovative AR learning platform through the incorporation of eye-tracking technology into the Microsoft HoloLens 2. This enhanced learning platform enables the collection of data on pupil dilation and problem-solving duration as students engage in AR-based learning activities. In this study, we hypothesize that pupil dilation and problem-solving time could be significant predictors of student performance in the AR learning environment. The results of our study suggest that problem-solving time may be a critical factor in predicting student learning success for materials involving procedural knowledge at low difficulty levels. Additionally, both pupil dilation and problem-solving time are predictive indicators of student learning outcomes when dealing with predominantly procedural knowledge at high difficulty levels.

Keywords

pupil dilation, problem-solving time, performance, augmented reality

Introduction

Pupil dilation acts as a physiological indicator of cognitive effort (Holmqvist et al., 2011), and solving time provides insight into a student's comprehension of the concepts taught during the lecture (De Boeck & Scalise, 2019). These measures shed light on how students interact with educational content, allowing educators to customize their teaching approaches for enhanced learning results. Despite their potential, there has been limited research exploring the relationship between pupil dilation, solving time, and student learning performance in Augmented Reality (AR) settings.

AR is revolutionizing education, offering dynamic, immersive experiences with tools like Microsoft HoloLens 2. It enriches learning, particularly in complex fields like engineering education, by making content interactive and accessible to various learning styles, thus improving educational results. While numerous studies have highlighted the advantages of AR-based learning, users continue to face challenges such as cognitive overload, triggered by an overload of stimuli, and cyber sickness (Alzahrani, 2020; Wu et al., 2013). Furthermore, it is hard to predict learning performance in such settings (Koumpouros, 2024). Hence, developing innovative methods to precisely assess students' understanding of learning materials in AR environments is a critical area for future AR research.

This research focuses on investigating the use of pupil dilation and problem-solving time as key parameters for predicting student learning success in an AR-based educational environment.

In this research, we have developed an innovative AR learning platform by integrating eye-tracking technology into the Microsoft HoloLens 2. The integration of eye-tracking technology is crucial for the real-time collection and analysis of eye-related metrics, such as pupil dilation, providing insights into student engagement and cognitive load. Such data are invaluable for understanding how students interact with and learn from AR content. The result is an enhanced learning platform that enables the collection of data on pupil dilation and problem-solving duration as students engage in AR-based learning activities. These metrics are then utilized as indicators to assess student learning performance.

[1]University of Missouri, Columbia, MO, USA

Corresponding Author:
Jung Hyup Kim, Department of Industrial & Systems Engineering, University of Missouri, E3437F Thomas & Nell Lafferre Hall, Columbia, MO 65211, USA.
Email: kijung@missouri.edu

Figure 1. Dikablis eye tracking with HoloLens 2 device.

Table 1. The Example of Performance Data for Each Participant.

Q. No	Response	Score	Solving time (s)
1	Musculoskeletal injury	100	21.012
2	Object of mass; distance	100	43.891
3	Zero	100	172.519
4	5 m	100	18.932
5	0.3 m	0	36.798
6	2.16 Nm	100	24.63
7	43.60%	100	30.866

In this study, we hypothesize that pupil dilation and problem-solving time could be significant predictors of student performance in the AR learning environment. For the experiment, we created two AR lectures on biomechanics. The first lecture introduced essential concepts in biomechanics and physics, including fundamental definitions and formulas. Following this, the next lecture built on these foundations by applying the concepts covered earlier, thus increasing the complexity. Lecture #1 was divided into seven modules, whereas Lecture #2 was organized into eight modules.

Methodology

Twelve engineering students (average age = 20.6) from the University of Missouri participated by completing a survey that gathered demographic information. Following the survey, participants were provided with a comprehensive orientation session to familiarize them with the HoloLens headset and the Dikablis Eye Tracking system (see Figure 1). After the calibration of both devices, the study was structured around two distinct AR lectures, each separated by a minimum interval of 4 hr and a maximum of 48 hr to ensure adequate rest and mental preparation for the participants. The initial lecture covered fundamental concepts in biomechanics, laying the groundwork for the subsequent session. The second lecture, designed to be more challenging, required participants to apply the knowledge acquired from the first lecture to solve advanced problems in the same fields.

During these sessions, participants engaged with a series of modules—seven in the first lecture and eight in the second—each designed to progressively build on their understanding and application of biomechanics. After each lecture concluded, we recorded the performance metrics and problem-solving times for each participant onto the laptop. For example, Table 1 displays data on performance and solving time from the first lecture. To understand the performance data, each module records a solving time alongside a score, with 100 indicating a correct answer and 0 denoting an incorrect one. As participants navigated through the educational content and problem-solving activities, we continuously gathered eye-tracking information, including metrics such as the pupil area for both the left and right eyes, along with the pupil size.

During the experiment, participants were equipped with two devices: a HoloLens for accessing AR learning materials, and Dikablis eye-tracking cameras to monitor pupil activity. Additionally, the NFER system was deployed to trigger the loading of subsequent AR modules based on the participants' physical locations, which included seven specific learning stations for Lecture #1 and eight spots for Lecture #2 (Yu et al., 2023). The NFER tag was placed on the table used by participants during the experiment. This tag tracks the real-time location within a calibrated indoor room. By utilizing this technology, we can accurately display the relevant AR learning contents (see Figures 2 and 3) by minimizing any mismatch issues between physical objects and virtual objects (Guo & Kim, 2021).

To analyze the participant's pupil data, we normalized the pupil area data using equation (1) (Liang et al., 2021) . In this equation, **Pnorm** represents the normalized pupil area, **Pi** stands for each data point in the pupil area ($i = 1, \ldots, n$), **min(P)** is the smallest pupil area observed in the entire dataset of the participant, and **max(P)** indicates the largest pupil area from the same dataset.

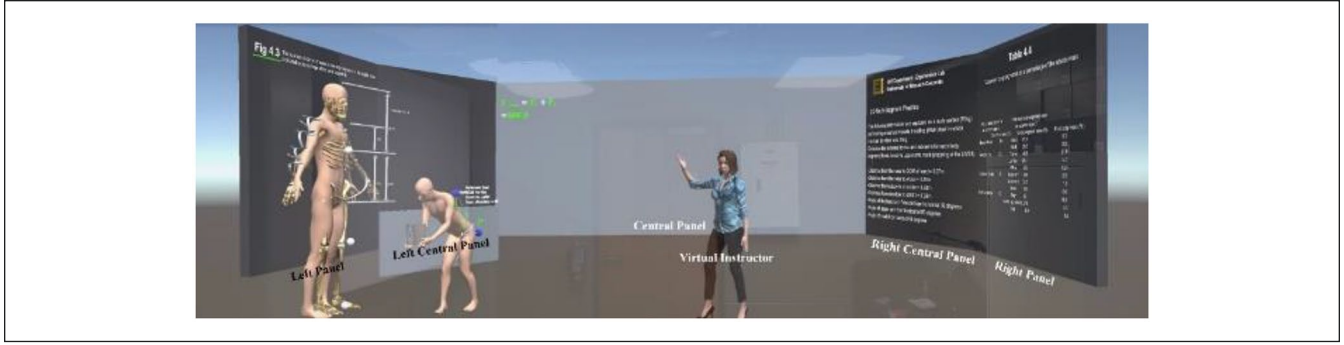$$P_{norm} = \frac{P_i - \min(P)}{\max(P) - \min(P)} \qquad (1)$$

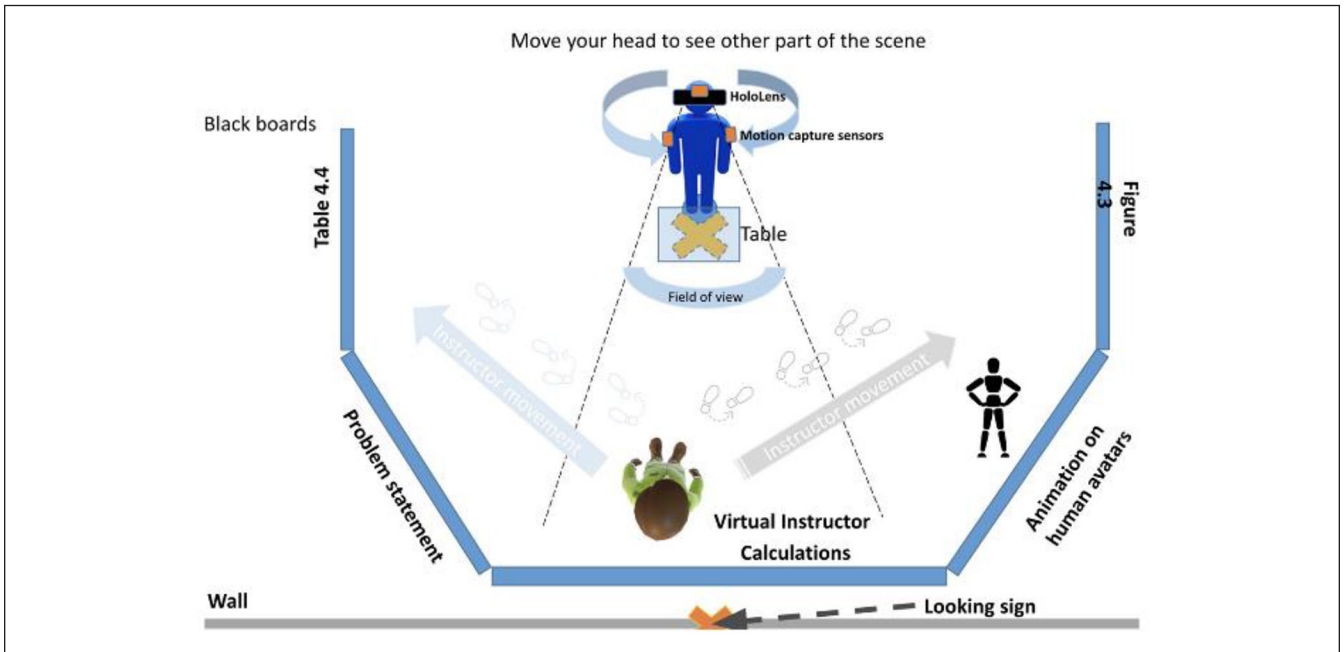**Figure 2.** Labeled 3D scene of an AR module built with unity.

Figure 3. Bird's-eye view of the AR learning platform (Kim et al., 2023).

After completing the normalization process, we divided the pupil dataset into two categories: baseline (**B**) and solving (**S**). The baseline includes the pupil data collected when a participant was waiting for new AR content after moving to the next location spot. In the solving category, pupil data is gathered when the participant addresses a problem associated with the recently learned content through the AR module. To determine the extent of pupil dilation when participants solved problems, the normalized pupil data from the problem-solving category was compared with baseline pupil data, denoted as **B-S**. The absolute value of **B-S** was used to measure its distance from zero, representing no pupil dilation difference between the baseline and problem-solving. This comparison with absolute value occurred across all modules in both Lecture #1 and Lecture #2.

To accurately measure each participant's problem-solving time, we designed an interface accessible via a touchscreen laptop computer (refer to Figure 4). When a participant was

prepared to tackle the question, they would press the "NEXT" button, which triggered the start of the timer. After selecting an answer from multiple choices, the "NEXT" button would activate. The timer stopped as soon as the participant clicked this button. Each problem-solving time is denoted as ST.

In this experiment, participants solved one problem after each AR module. However, module 1 in Lecture 2 did not include a question, resulting in seven ST values for both Lectures #1 and #2. Accordingly, for Lecture #1, seven normalized B-S values corresponded to modules #1 through #7. For Lecture #2, seven normalized B-S values corresponded to modules #2 through #8 for each participant.

Results

The findings indicate that the mean pupil dilation difference between baseline and problem-solving is 0.11 with a standard deviation of 0.077 for Lecture #1 and 0.09 with a
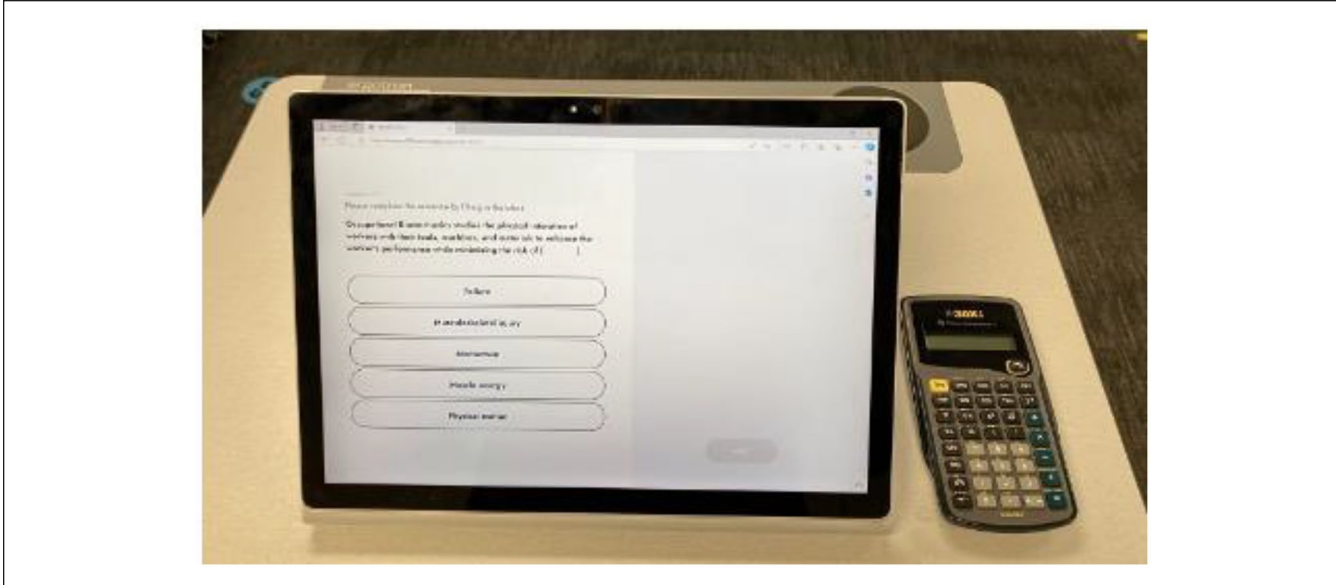
Figure 4. The touchscreen laptop computer to collect problem-solving time.

standard deviation of 0.079 for Lecture #2 (see Table 2). This suggests a 9% to 11% fluctuation in pupil size while participants were engaged in problem-solving compared to when no visual stimuli were presented. For the problem-solving time, the mean solving time is 41.61 s with a standard deviation of 42.32 for Lecture #1 and 79.5 s with a standard deviation of 74.11 for Lecture #2 (see Table 3). This implies that participants took longer to solve the problems in Lecture #2, possibly perceiving them as more difficult than those in Lecture #1. This interpretation could be supported by a comparison of performances between the two lectures: Lecture #1 had an average performance score of 90.42 with a standard deviation of 9.31. In contrast, Lecture #2 had an average score of 68.84 with a standard deviation of 18.43. The performance in Lecture #2 was significantly lower than in Lecture #1 (see Figure 5).

For Lecture #1, the most effective regression model considers problem-solving time and pupil dilation as independent variables, with the sum of each module's performance as the dependent variable, including only the solving times for modules #4, #5, and #6. This is evidenced by an R-squared value of 0.86 (see Figure 6 (a)). It means that the solving times for those modules, called ST-4, ST-5, and ST-6, indicate their critical role in determining student performance for Lecture #1 contents. In Lecture #2, which contains more complex biomechanics contents, the model's R-squared value rises to 0.97 (see Figure 6 (b)). This indicates a significantly enhanced predictive accuracy. The solving times for modules #4 (ST-4) and #8 (ST-8) and pupil dilations for modules #2 (B-S 2), #5 (B-S 5), #7(B-S 7), #8 (B-S 8) are significant predictors of performance. Based on the regression coefficient findings from Lecture #1, as detailed in Table 4, it is evident that there is a positive correlation

Table 2. Descriptive Statistics for Pupil Dilation Differences with Absolute Value Between Baseline and Problem-Solving.

Parameters	Lecture 1		Lecture 2	
	Mean	Std Dev	Mean	Std Dev
B-S 1	0.15	0.076	-	-
B-S 2	0.11	0.058	0.09	0.076
B-S 3	0.10	0.087	0.10	0.088
B-S 4	0.12	0.091	0.11	0.078
B-S 5	0.11	0.064	0.10	0.074
B-S 6	0.10	0.075	0.08	0.058
B-S 7	0.11	0.090	0.10	0.095
B-S 8	-	-	0.07	0.081

Table 3. Descriptive Statistics for Problem-Solving Time.

Parameters	Lecture 1		Lecture 2	
	Mean	Std Dev	Mean	Std Dev
ST 1	17.08	8.46	-	-
ST 2	41.47	32.25	31.5	66.98
ST 3	37.64	58.38	62.53	56.82
ST 4	38.96	30.32	42.25	41.16
ST 5	74.75	90.77	53.05	23.29
ST 6	38.43	24.52	26.44	15.95
ST 7	42.97	51.59	56.51	112.63
ST 8	-	-	284.22	201.97

between student performance and the time spent solving problems in modules #4 and #5. Conversely, a negative correlation was observed between student performance and
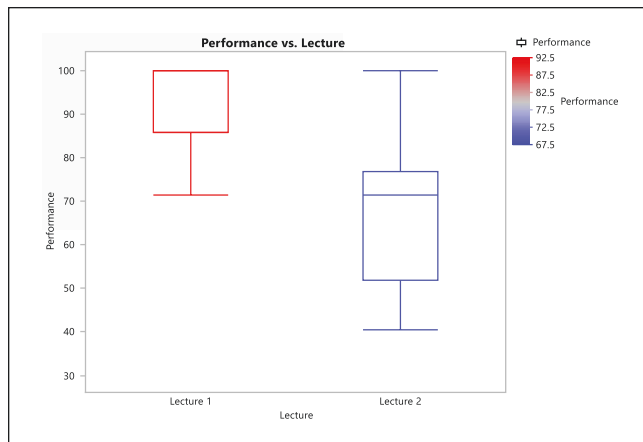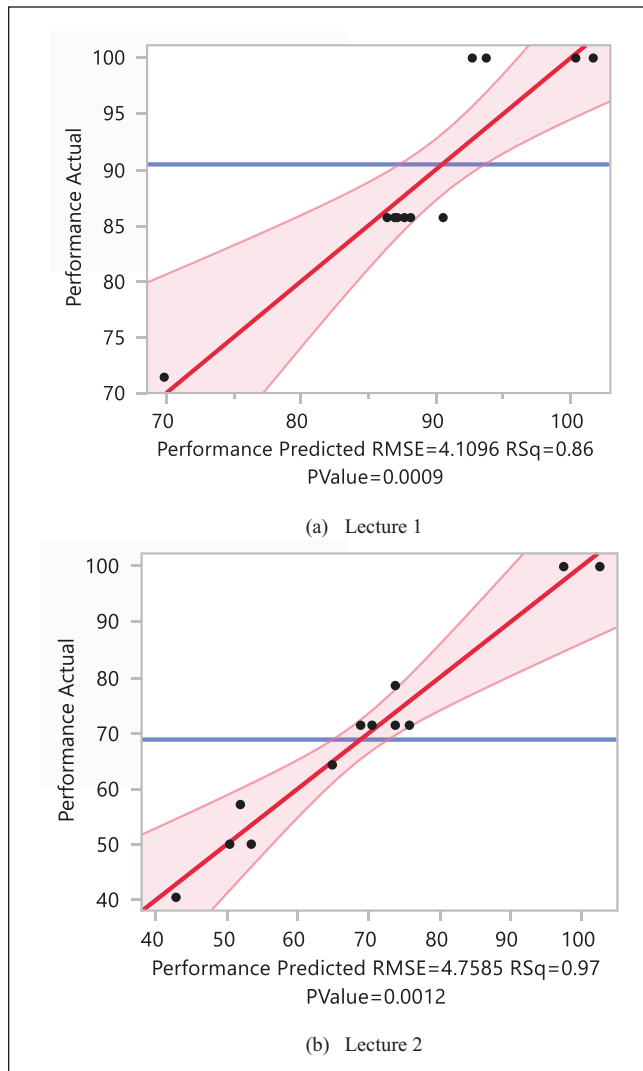
Figure 5. Performance results.

(a) Lecture 1

(b) Lecture 2

Figure 6. A Scatter plot comparing actual performance and predicted performance: (a) Lecture 1. (b) Lecture 2.

Table 4. The Most Effective Regression Model of Problem-Solving Times Correlated with Performance in Lecture 1.

Term	Estimate	Std error	*t* Ratio	Prob>|t|
Intercept	95.203509	2.597573	36.65	<.0001*
ST-4	0.1597777	0.045701	3.50	.0081*
ST-5	0.0593085	0.015113	3.92	.0044*
ST-6	−0.400484	0.058719	−6.82	.0001*

Table 5. The Most Effective Regression Model of Problem-Solving Times and Pupil Dilation Correlated with Performance in Lecture 2.

Term	Estimate	Std error	*t* ratio	Prob>|t|
Intercept	80.111508	7.989319	10.03	.0002*
B-S 2	−230.3605	32.65722	−7.05	.0009*
B-S 5	−182.5271	36.89973	−4.95	.0043*
B-S 7	−90.87113	35.19589	−2.58	.0493*
B-S 8	212.11168	51.96491	4.08	.0095*
ST 4	0.2144843	0.045473	4.72	.0053*
ST 8	0.0489495	0.010133	4.83	.0048*

the time spent on problem-solving in module #6. In Lecture #2, as indicated in Table 5, positive correlations were identified between student performance and the time spent solving problems in modules #4 and #8, as well as the pupil dilation observed in module #8. Additionally, negative correlations were noted between student performance and pupil dilation in modules #2, #5, and #7.

Discussion

According to the results, we observed that when students engaged with learning materials primarily composed of procedural knowledge at low difficulty levels (Lecture #1 in modules #4 to #6), their performance demonstrated a significant correlation with the time spent on problem-solving but showed no correlation with pupil dilation. As previously mentioned, Lecture #1 introduced fundamental concepts and simple examples in biomechanics. Specifically, modules #1 to #3 covered the definition of biomechanics, forces and moments, and static equilibrium using free body diagrams. Modules #4 to #6 offered straightforward examples of calculating forces and moments on body segments. Module #7 detailed additional information for solving more complex multi-segment biomechanics problems, which will be explored further in Lecture #2. In other words, modules #1, #2, #3, and #7 are structured to impart declarative knowledge related to biomechanics, while modules #4 to #6 focus on teaching procedural knowledge, such as calculating forces and moments on the body. the results from the regression model indicate that the time spent solving problems in modules #4 to #6 is a significant predictor of

student learning success in Lecture #1. It is interesting to observe how problem-solving time can predict performance. According to Table 4, spending more time on modules 4 and 5 but less time on module 6 results in better performance. This could be because mastering modules 4 and 5 provides a significant advantage in understanding and efficiently completing module 6. In other words, a student who spends more time solving problems correctly in modules 4 and 5 spend less time on module 6 due to their high level of understanding of calculating forces and moments in biomechanics.

However, when students interacted with learning materials that predominantly featured procedural knowledge at a high difficulty level (Lecture #2), their performance significantly correlated with the time they dedicated to problem-solving and pupil dilation. Aside from module #1, which focused on summarizing the content from Lecture #1, all other modules were dedicated to addressing multi-segment biomechanics problems to assess potential ergonomic risks during manual material handling tasks. This indicates that for procedural learning at high difficulty level, both the effort invested in solving problems and the physiological responses, as reflected by pupil dilation, are important predictors of student success.

Prior research has established pupil dilation as a reliable measure of cognitive load (Gavas et al., 2017; Kim & Yang, 2017; Sibley et al., 2011). This phenomenon becomes particularly evident when students tackle new content, such as calculating forces and moments affecting individual body segments. Such topics demand considerable cognitive effort to grasp the sequence and methodology involved. Participation in these procedural tasks at a high difficulty level can increase cognitive load and student engagement, as indicated by changes in pupil dilation. Additionally, the time spent on problem-solving acts as a direct measure of student engagement with a task.

Our regression model supports these observations. According to Table 5, the variables of problem-solving time and pupil dilation in modules #3 and #6 are not significant predictors of performance in Lecture #2. This may be due to the lack of new information or equations in these modules. Specifically, in module #3, the virtual instructor repeated the method for calculating forces and moments, previously covered, but applied it to a different posture. Similarly, in module #6, the same procedure used in module #5 for calculating forces and moments on the upper arm was repeated.

The results of our study indicate that learning modules focused on procedural knowledge significantly impact student learning performance. This is because adequate knowledge about the subject is essential for solving problems that require procedural understanding. An interesting observation from the study is that at low difficulty levels, only problem-solving time is relevant for predicting performance. However, at high difficulty levels, both pupil dilation and problem-solving time emerge as significant predictors in the model.

Conclusion

This study seeks to predict student learning outcomes in Augmented Reality (AR) educational environments by examining pupil dilation and problem-solving time as primary indicators. The results of our study suggest that both pupil dilation and problem-solving time serve as predictive indicators of student learning outcomes when the material is predominantly procedural knowledge. This study shows that AR's impact could be amplified through pupil dilation and solving time to understand student learning in the AR environment. Pupil dilation reflects the mental effort in AR interactions, while solving time measures how students comprehend concepts. These metrics could provide real-time insight into student engagement and cognitive challenges. Educators can utilize these insights to customize content, ensuring it matches student capability, and thereby fostering personalized learning. The current study explores AR's role in education, focusing on how pupil dilation and solving time as performance metrics can enhance AR-based engineering education. This research contributes to digital education, highlighting eye-tracking's potential to improve AR methodologies. Conclusively, by leveraging these findings, we will be able to create an innovative AR learning platform that tailors itself to the learner's cognitive state and performance metrics.

Future studies should aim to increase the sample size and diversity of participants to improve the generalizability of the results. Additionally, it is important to expand the range of AR content to encompass various educational and training contexts. Furthermore, incorporating additional physiological sensors could enhance the accuracy and reliability of performance prediction models in AR settings.

Declaration of Conflicting Interests

Funding

References

Alzahrani, N. M. (2020). Augmented reality: A systematic review of its benefits and challenges in e-learning contexts. *Applied Sciences*, *10*(16), 5660.

De Boeck, P., & Scalise, K. (2019). Collaborative problem solving: Processing actions, time, and performance. *Frontiers in Psychology*, *10*, 1280.

Gavas, R., Chatterjee, D., & Sinha, A. (2017). Estimation of cognitive load based on the pupil size dilation. In *Paper presented at the 2017 IEEE international conference on systems, man, and cybernetics (SMC)*, 1499–1504.

Guo, W., & Kim, J. H. (2021). Designing augmented reality learning systems with real-time tracking sensors. In *Paper presented at the Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2021 Virtual Conferences on Neuroergonomics and Cognitive Engineering, Industrial Cognitive Ergonomics and Engineering Psychology, and Cognitive Computing and Internet of Things*, July 25–29, 2021, USA.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Kim, J. H., & Yang, X. (2017). Applying fractal analysis to pupil dilation for measuring complexity in a process monitoring task. *Applied Ergonomics*, *65*, 61–69.

Kim, J. H., Yu, C. Y., Seo, K., Wang, F., & Oprean, D. (2023). The effect of virtual instructor and metacognition on workload in a location-based augmented reality learning environment. In *Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 67, No, 1, pp. 1550–1555). Sage CA: Los Angeles, CA: SAGE Pulications.

Koumpouros, Y. (2024). Revealing the true potential and prospects of augmented reality in education. *Smart Learning Environments*, *11*(1), 2.

Liang, N., Yang, J., Yu, D., Prakah-Asante, K. O., Curry, R., Blommer, M., Swaminathan, R., & Pitts, B. J. (2021). Using eye-tracking to investigate the effects of pre-takeover visual engagement on situation awareness during automated driving. *Accident Analysis and Prevention*, *157*, 106143.

Sibley, C., Coyne, J., & Baldwin, C. (2011). Pupil dilation as an index of learning. In Paper presented at the proceedings of the human factors and ergonomics society annual meeting (Vol. 55, No. 1, pp. 237–241). Sage CA: Los Angeles, CA: SAGE Publications.

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, *62*, 41–49.

Yu, C. Y., Kim, J. H., Mostowfi, S., Wang, F., Oprean, D., & Seo, K. (2023). Developing an augmented reality-based interactive learning system with real-time location and motion tracking.In *Paper presented at the international conference on human-computer interaction*, 227–238. Cham: Springer Nature Switzerland.