Effect of Short-Form Videos on Student Attention

1st Mohsen Mahmoudi
ECE Department
University of Tehran
Tehran, Iran
Mohsen.Mahmoudi@ut.ac.ir

2nd Fattaneh Taghiyareh
ECE Department
University of Tehran
Tehran, Iran
ftaghiyar@ut.ac.ir

3rd Kianosh Arian
ECE Department
University of Tehran
Tehran, Iran
kianosh.arian@ut.ac.ir

***Abstract*— Short-form videos (SFVs) have recently emerged as one of the most prominent forms of entertainment content on social media, attracting huge audiences across platforms such as TikTok, Instagram, and YouTube. This study explores the potential effects of short-form entertainment content on cognitive functions, especially learning-related ones. We conducted an experimental study to investigate the immediate effects of SFV consumption on student learning. During the experiment, volunteers engaged in two learning tasks; they avoided watching SFVs for three hours before the first learning task and watched 15 minutes of SFVs before doing the second task. The SFV content viewed by subjects was personalized and recommended to them by their own Instagram algorithm to ensure real-world relevance. Subjects wore a Pupil Core eye-tracking device during the experiment. Subjects' eye measures were monitored and analyzed in pre- and post-experiments to gain insight into the immediate neurophysiological effects of using Instagram on learning performance. Our findings revealed that watching SFVs negatively affects the subject's pupil diameter, which implies reduced attention and processing load. This negative effect gradually decreases over time, and its dissipation can be modeled with an autoregression and space-state model. The study highlights that watching short entertainment videos before study sessions can impair the learning process. It could be beneficial for students to incorporate guidelines that regulate short-form video consumption before learning activities.**

Keywords— Short-Form Videos, Learning Performance, Eye-Tracking, Social Media, User Modeling

I. Introduction

Short-form videos are one of the latest trends in entertainment content available on social media. Despite being relatively new, this type of content has become one of the most popular forms of entertainment and is consumed by people all around the world for millions of hours every day. However, many believe that SFVs have potential negative impacts on people's lives, particularly on the human brain and cognitive functions, especially those related to learning. Understanding the effect of consuming such content on the human mind is crucial for planning a healthier lifestyle and preventing mental health problems.

Due to the increasing presence of SFVs in people's everyday lives, researchers have started studying their effect on the human mind and cognitive constructs. For example, a study shows how getting exposed to short-form animal videos leads to reduced attention and increased reaction time in the Stroop task [1]. Another study suggests that automatic use of TikTok before sleep causes daytime fatigue due to higher levels of cognitive arousal [2].

Even though most SFV consumers are students, we lack sufficient information about the impact of consuming these videos on their learning abilities. In this research, we aim to answer the question of whether consuming SFVs affects the learning process by examining students' neurophysiological measures, such as pupil diameter while learning before and after consuming SFVs.

Although there are growing research efforts on studying the potential effects of short-form video consumption, significant gaps remain in understanding how this matter influences learning abilities. Current literature often focuses on the general impacts of social media addiction or digital distractions on academic performance, but it does not adequately explore the immediate effects of SFV consumption on learning processes. Moreover, its neurophysiological and cognitive consequences in educational environments remain underexplored, making it essential to investigate these aspects to understand the full implications of SFV consumption on students' learning.

To address the gaps mentioned, we conducted a controlled experiment to measure the effect of SFV consumption on neurophysiological responses during learning tasks. By employing a pretest-posttest design, our study captures the immediate impacts of consuming algorithmically recommended short-form videos on Instagram. Through eye-tracking technology, we monitored key neurophysiological indicators, such as pupil diameter, to assess changes in cognitive constructs before and after exposure to SFVs. This approach allows us to draw connections between the consumption of short-form content and subsequent learning performance, providing a better understanding of how digital media consumption may influence students' ability to retain and process educational material in real-time.

This paper is organized as follows: Section 2 provides a comprehensive review of the literature, highlighting existing studies on the impact of SFV consumption on the human brain, cognitive constructs, and learning. Furthermore, we delve into the neurophysiological measurements, especially eye-tracking methods, employed in such research. Section 3 outlines the research methodology, detailing the experimental design, data collection processes, and analytical techniques for investigating the neurophysiological data. Section 4 presents the study's results, analyzing the neurophysiological data and its implications. Finally, Section 5 discusses the findings in the context of the broader literature, draws conclusions, and suggests potential avenues for future research.

II. Literature Review

Short-form entertainment content are videos typically lasting from a few seconds to a few minutes, with varying subjects ranging from comedic sketches and musical performances to animal videos and everyday event storytelling. In recent years, short-form video platforms have experienced substantial growth, and many established social media platforms (like Instagram and YouTube) have integrated short-form video sections into their platforms. One of the main factors of this growth can be attributed to advanced recommendation algorithms that these platforms use, which recommend personalized content aimed to minimize the user's choosing cost so the audience just needs to slide the video up or down [3] and therefore makes the platform easier to use.

A. Behavioral Studies

Due to the increasing presence of SFVs in people's everyday lives, researchers have started studying the effects of their consumption on the human mind and cognitive constructs. For instance, studies have linked TikTok use disorder (TTUD) to memory loss, anxiety, and depression [4]. Other studies have found that SFV consumption causes a significant degradation in prospective memory's performance, person's ability to Retain Intentions, while other kinds of social media usage have very small to no effect on the prospective memory [5], which highlights the different nature of this new type of content and shows the importance of studying its effect on the human mind.

B. Brain Studies

Considering the consequences mentioned above, it becomes evident that SFVs' impact extends beyond mere entertainment consumption. The root of these negative effects can be found by studying the effect of SFVs on the brain. SFV platforms like TikTok use highly personalized and complex recommendation algorithms to maximize user engagement with their content, which is shown to increase the activity of the Default Mode Network (DMN) and Ventral Tegmental Area (VTA) parts in the brain [6]. The DMN is known for its involvement in self-referential thinking and mind wandering, and the VTA is a crucial component of the brain's reward system, which is responsible for producing dopamine and linked with addictive behavior [7], which explains the addictive and distractive qualities of SFVs.

C. About Learning

With a high percentage of SFV users being of school or college age and SFVs having so many side effects on the human mind, their potential impact on learning outcomes demands careful examination. By studying how SFVs affect cognitive processes and behavioral patterns, educators can gain insights into the specific challenges students may face in the learning environment and develop strategies to overcome these challenges.

The current studies done on the effects of SFV on student's academic performance show that students addicted to SFV platforms are more likely to procrastinate on their tasks [8], have less learning motivation [9] and lowered happiness and perseverance [10]. In another study, Audrey Mekler (2021) shows that the more time students spend on TikTok each day, the more likely they are to become distracted by TikTok when in the classroom or completing schoolwork [11].

D. Eye-tracking

Eye-tracking is a method of collecting neurophysiological measures from people and is commonly used in cognitive research. Attributes like pupil diameter, fixation count and duration, and saccadic eye movements are some of the metrics that can be obtained by eye-tracking. These metrics give us information that can be used to analyze people's mental state and cognitive constructs.

Eye-tracking methods are used to study many learning processes. The advantage to this method is that it causes minimal distractions to learners and records the information in real-time throughout the whole learning process. For example, in a study, eye-tracking was used to assess students' learning styles by analyzing the way users look at learning material [12]. The possibility to replay records of eye movements can also contribute to improving the design of instructional materials or procedures [13].

Eye measures, such as pupil diameter, fixation count, and average fixation duration are some of the neurophysiological measures that can be obtained by eye-tracking methods. These measures can then be used to assess subjects' mental state and cognitive constructs. For example, a study showed that participants' pupils dilate more when studying high-valued than low-valued words, and these changes were associated with better memory for high-valued words [14]. Pupil diameter has been shown to correlate with multiple cognitive functions, namely processing load, such that as processing load increases, pupil diameter also increases [15].

III. METHODOLOGY

A. Setting of the Study

This study aims to measure the impact of consuming short-form entertainment content on students' learning and neurophysiological measures. The experiment was conducted in a controlled environment within a laboratory setting to ensure consistency and to minimize external variables that could influence the results. The duration of the experiment was about 60 minutes per subject and subjects were required to wear Pupil Core eye-tracking glasses to have their neurophysiological eye measures recorded for later analysis. Participants were university students aged between 20 and 30 years and were required to be regular users of Instagram so that the platform's algorithm could recommend personalized short videos to them. The subjects were told to avoid using social media platforms for 3 hours before the experiment to prevent any interference with the experiment.

B. Experiment design

This study employs a pretest–posttest design during which subjects first engage in a learning task. In this task, they first watch a 15-minute educational video and then answer a questionnaire about the content of the video for about 5 minutes. After completing this task, they watch SFVs on Instagram for 15 minutes. The videos are algorithmically recommended to the subjects on their personal Instagram account. After this, they get five minutes of unstructured break time before starting the second learning task. The second learning task has a similar structure and length to the first learning task. The educational content in both learning tasks is of STEM fields, and to avoid any task-specific biases the two learning tasks used for pre- and post-tests were randomly switched for subjects, resulting in each of the tasks being used as pre-test content for half of the subjects and post-test content for the other half. The results of this experiment are analyzed for potential performative and neurophysiological effects.

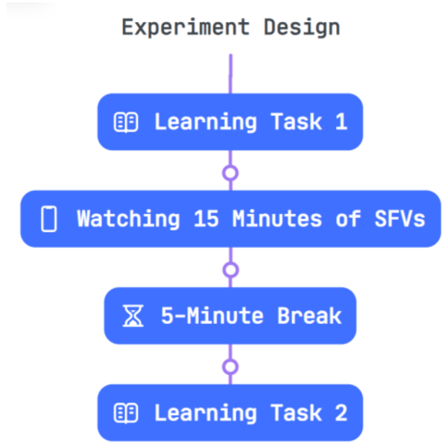

Fig. 1. Experiment Design

C. Participants and Recruitment

This study was conducted on 20 university students recruited through announcements and flyers. Participants were required to fill out a survey about their health status, medication use, and average daily Instagram usage time. Eligible participants who spent at least 30 minutes on the app daily and did not use any medication that could affect brain activity were selected and contacted for the experiment.

D. Data Collection and Analysis

In the preprocessing stage, we first removed data points with less than 85% confidence, as assessed by the Pupil Core device. This ensures the reliability of the results and removes device errors. Then, we used the Interquartile Range (IQR) method to detect and remove outlier data from the set. Subjects answered the assessment questions at the end of each learning task with different speeds, resulting in different session durations for each subject. Since answering questions requires different mental processes from watching educational videos, we decided to remove the eye-tracking data points relating to answering the questionnaire and only focus on watching the educational content for our data analysis process.

After preprocessing, we averaged the subject results for each of the pre- and post-experiment data, resulting in two datasets, one for post-experiment data and the other for pre-experiment data. We analyzed pupil diameter in these two datasets.

This study examined Autoregression (AR) and State-Space models to predict how the pupil's size changes over time. These models were chosen for their ability to capture temporal dependencies in time series data and their robustness in handling evolving data, especially physiological measurements like pupil diameter, which may exhibit complex, time-dependent behavior after exposure to stimuli such as Instagram reels.

1) Autoregression (AR) Model

The Autoregression (AR) model is a straightforward yet powerful method for modeling time series data. The AR model assumes that the current value of the time series can be expressed as a linear combination of its previous values. Specifically, in an AR model of order ppp, the current value of the time series is a function of the previous ppp values plus a noise term. Mathematically, this can be expressed as:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

Where y_t is the value of the time series at time t, c is a constant, $\phi_1, \phi_2, \ldots$ are the model's coefficients, and ϵ_t is the noise term.

The AR model is particularly well-suited for the data in this study because it assumes that past pupil diameter measurements can provide valuable information for predicting future values. This assumption is valid because previous stimuli often influence physiological responses such as pupil dilation and may not change instantaneously. Using the AR model, we can effectively capture the time-dependent patterns in pupil diameter, including gradual changes, sudden shifts, or cyclic behavior.

Moreover, the AR model's simplicity makes it a good starting point for understanding the temporal structure of the data. It allows for easy interpretation of the relationship between current and past values, which is critical when the goal is to explain how past exposure to stimuli like Instagram reels affects pupil diameter over time.

2) State-Space Model

The State-Space model is a more flexible and general approach to modeling time series data, especially when the underlying process is complex and potentially driven by unobserved factors. State-space models describe the observed time series as being generated by a hidden (latent) process that evolves over time. The model consists of two fundamental equations: the state and observation equations.

State Equation: This equation describes the evolution of the hidden state over time. It captures the dynamics of the underlying process that governs the observed data.

Observation Equation: This equation links the hidden state to the observed data. It explains how the hidden state translates into the measurements we observe.

The State-Space model is particularly advantageous when the time series data exhibit characteristics such as non-stationarity, where the statistical properties of the series change over time, or when the process driving the data is assumed to be influenced by factors that are not directly observable. In the context of this study, the State-Space model can accommodate the idea that pupil diameter might be influenced by both observed factors (such as time elapsed since watching Instagram reels) and unobserved factors (such as cognitive load or emotional response).

The flexibility of the State-Space model allows it to capture a wide range of temporal dynamics, including trends, cycles, and irregular fluctuations. This makes it a powerful tool for modeling pupil diameter's complex and potentially nonlinear behavior over time.

3) *Fitting the Models*

Both the AR and State-Space models were fitted to the pupil diameter data using standard statistical techniques. For the AR model, the order p was selected based on the Akaike Information Criterion (AIC), which balances model complexity and goodness of fit. The coefficients were estimated using least squares, which minimizes the sum of squared differences between the observed and predicted values.

For the State-Space model, the parameters were estimated using maximum likelihood estimation. This technique finds the parameter values that maximize the likelihood of the observed data given the model. This process involves adjusting the model parameters until the best fit is achieved.

In summary, the AR model provides a simple yet effective way to capture the temporal dependencies in the pupil diameter data. In contrast, the State-Space model offers a more flexible framework that can accommodate complex, time-varying dynamics. Both models are well-suited for understanding and predicting the evolution of pupil diameter over time in response to stimuli like Instagram reels.

E. *Ethical Considerations*

Participants provided informed consent before the experiment, with the assurance that their anonymity and confidentiality would be maintained throughout the study. They were informed of their right to withdraw from the study at any point, and the study was approved by the TUMS Research Ethics Committee.

F. *Experiment's Intrusiveness*

Eye-tracking tools are generally considered to be among the least intrusive devices [16]. However, they can still cause some limitations for users. In our experiment, subjects' heads were placed on chin rests to prevent excessive head movement, which is unlike the natural environment of subjects. Additionally, some individuals may feel self-conscious about their gaze being tracked, which could potentially affect the naturalness of their responses.

IV. RESULTS

In this section, we analyze the results of the pre- and post-experiment eye-tracking data. We investigate the average pupil diameter of the subjects while doing the two learning tasks before and after consuming SFVs. We then try to fit the difference of the pre- and post-experiment diameters into a function to understand the decay of the impact of watching SFVs when learning.

A. *Pupil Diameter Comparison*

As seen in Fig. 2, the average pupil diameter results in the pre- and post-experiment data show that subjects had lower pupil diameter when learning after watching SFVs compared to before watching SFVs, which implies lower processing load [15] and worse memory encoding [17]. The observation that subjects exhibit a lower processing load with the same learning material after viewing SFVs could suggest reduced learning quality.

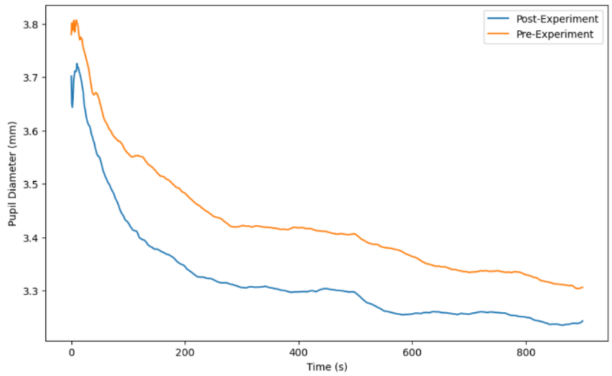

Fig. 2. Average Pupil Diameter in Time

B. *Diameter difference function*

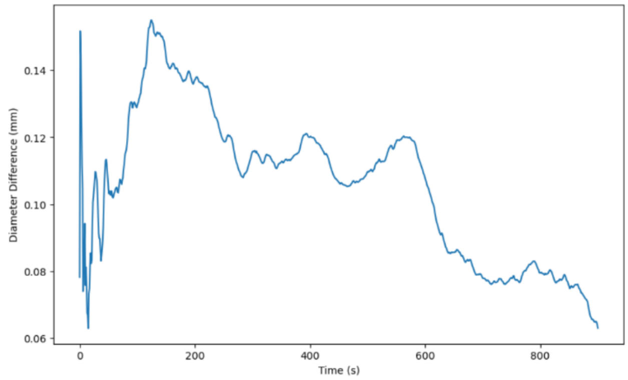

Fig. 3. Difference in Pre and Post Pupil Diameters

Fig. 3 shows the difference in pupil diameter for pre- and post-experiments. This difference is highest near the second minute of the learning tasks, after which it gradually decreases. By the end of the 20-minute learning task, this difference has decayed by 57%, implying users' relative recovery to their baseline psychological state before the consumption of SFVs.

C. *Fitting's Results*

After fitting the Autoregression (AR) and State-Space models to the pupil diameter data, we conducted a detailed analysis of their performance, focusing on the model hyperparameters, prediction accuracy, and visualization of the fitted models against the actual data.

1) *Autoregression (AR) Model Results*

The Autoregression (AR) model was fitted with a lag of 2, meaning that the current pupil diameter value was predicted using the previous two-time points. The specific hyperparameters obtained from the model fitting were:

- Constant (Intercept): 0.000352
- The coefficient for Lag 1: 1.038872
- The coefficient for Lag 2: -0.043127

These coefficients suggest that the pupil diameter at a given time is strongly influenced by the immediately preceding measurement (Lag 1), with a slight negative adjustment based on the second preceding measurement (Lag 2). The constant term is small, indicating that the baseline change in pupil diameter is minimal when no other factors are considered.

2) Accuracy Measures for AR Model

To evaluate the accuracy of the AR model, we calculated the Mean Absolute Error (MAE) and Mean Squared Error (MSE):

- Mean Absolute Error (MAE): 0.000632
- Mean Squared Error (MSE): 3.396×10^{-6}

The low error values suggest that the AR model effectively represents the short-term connections in the pupil diameter data. The model accurately anticipates the slight changes in pupil diameter over time, producing predictions that closely align with the actual data.

3) State-Space Model Results

The State-Space model was fitted using the SARIMAX framework with the following hyperparameters:

- AR Coefficient for Lag 1: 1.040255
- AR Coefficient for Lag 2: -0.041322
- Variance of the Error (σ^2): 0.000009

Like the AR model, the State-Space model shows that the pupil diameter is primarily driven by the first lag, with a minor adjustment based on the second lag. The variance parameter indicates that the model has captured a tiny amount of unexplained variability, reflecting the model's precision.

4) Accuracy Measures for State-Space Model

The accuracy of the State-Space model was evaluated using the same error metrics:

- Mean Absolute Error (MAE): 0.000628
- Mean Squared Error (MSE): 3.401×10^{-6}

The accuracy measures for the State-Space model are very close to those of the AR model, with a slightly lower MAE, indicating that it provides an almost equally precise fit to the data. The similar performance between the two models suggests that the State-Space model, while more complex, offers a comparable level of accuracy in this specific context.

To visually assess the fit of both models, we plotted the predicted values against the actual pupil diameter measurements (Fig 4). The plots show that the AR and State-Space model predictions closely follow the observed data, with only minor deviations. This strong visual alignment confirms that both models accurately capture the temporal dynamics of pupil diameter changes, with the State-Space model also effectively accounting for underlying factors that may influence the measurements.

The AR and State-Space models accurately predicted pupil diameter over time, with very close MAE and MSE values. While the AR model offers simplicity and ease of interpretation, the State-Space model provides a more flexible approach that can potentially accommodate more complex dynamics. However, in this particular case, the added complexity of the State-Space model did not significantly improve prediction accuracy over the AR model.

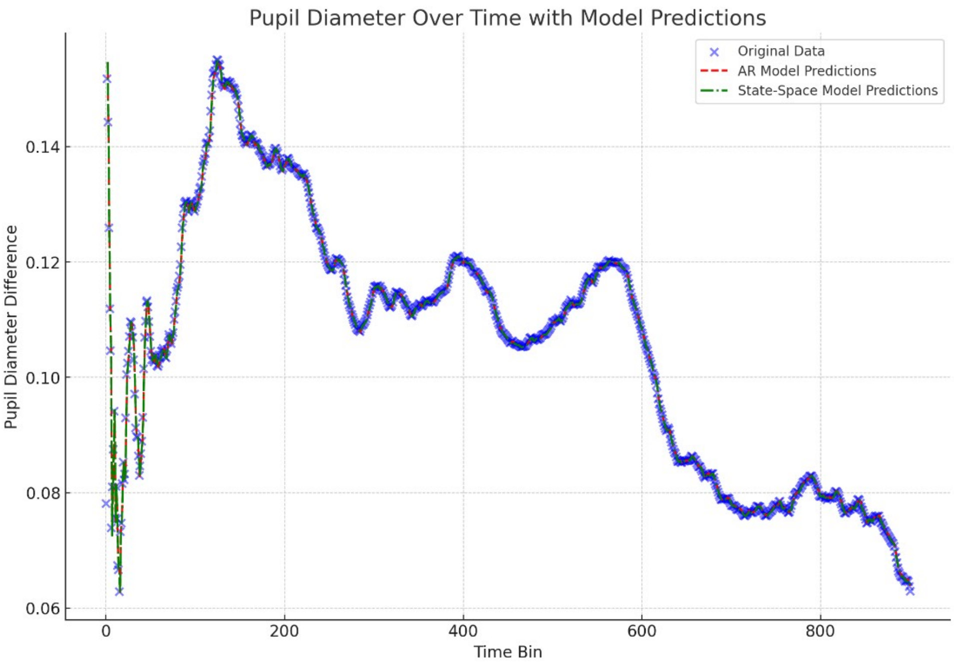

Fig 4 - Pupil Diameter Over Time with Model Predictions

V. CONCLUSION

This study investigated the immediate effects of short-form video (SFV) consumption on learning performance, focusing on neurophysiological responses using eye-tracking. The results indicate that viewing SFVs before learning tasks decreases pupil diameter, implying reduced processing load and attention, hindering learning. However, this effect diminishes over time, with recovery observed during the learning task. Although a 5-minute break was provided for participants between the pre and post-experiment phases, fatigue and state changes resulting from participating in the experiment for a prolonged period may have affected the subjects' behavior.

Future investigations can focus on the long-term effects of SFVs on learning and potential mitigation strategies. They can address the limitations of this study by including a larger subject pool and examining additional neurophysiological indicators like fixation count and duration to better understand SFVs' effects on cognitive processes, especially those related to learning.

REFERENCES

[1] Chen, Y., Li, M., Guo, F., & Wang, X. (2023). The effect of short-form video addiction on users' attention. *Behaviour and Information Technology*, *42*(16), 2893–2910. https://doi.org/10.1080/0144929X.2022.2151512

[2] Wang, K., & Scherr, S. (2022). Dance the Night Away: How Automatic TikTok Use Creates Pre-Sleep Cognitive Arousal and Daytime Fatigue. *Mobile Media and Communication*, *10*(2), 316–336. https://doi.org/10.1177/20501579211056116

[3] Wu, J. (2021). Study of a Video-sharing Platform: The Global Rise of TikTok. Diss. Massachusetts Institute of Technology, 2021.

[4] Sha, P., & Dong, X. (2021). Research on adolescents regarding the indirect effect of depression, anxiety, and stress between tiktok use disorder and memory loss. *International Journal of Environmental Research and Public Health*, *18*(16). https://doi.org/10.3390/ijerph18168820

[5] Chiossi, F., Haliburton, L., Ou, C., Butz, A. M., & Schmidt, A. (2023, April 19). Short-Form Videos Degrade Our Capacity to Retain Intentions: Effect of Context Switching On Prospective Memory. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3544548.3580778

[6] Su, C., Zhou, H., Gong, L., Teng, B., Geng, F., & Hu, Y. (2021). Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area. *NeuroImage*, *237*. https://doi.org/10.1016/j.neuroimage.2021.118136

[7] Polter, A. M., & Kauer, J. A. (2014). Stress and VTA synapses: Implications for addiction and depression. *European Journal of Neuroscience*, *39*(7), 1179–1188. https://doi.org/10.1111/ejn.12490

[8] Xie, J., Xu, X., Zhang, Y., Tan, Y., Wu, D., Shi, M., & Huang, H. (2023). The effect of short-form video addiction on undergraduates' academic procrastination: a moderated mediation model. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1298361

[9] Ye, J. H., Wu, Y. T., Wu, Y. F., Chen, M. Y., & Ye, J. N. (2022). Effects of Short Video Addiction on the Motivation and Well-Being of Chinese Vocational College Students. *Frontiers in Public Health*, *10*. https://doi.org/10.3389/fpubh.2022.847672

[10] Mu, H., Jiang, Q., Xu, J., & Chen, S. (2022). Drivers and Consequences of Short-Form Video (SFV) Addiction amongst Adolescents in China: Stress-Coping Theory Perspective. *International Journal of Environmental Research and Public Health*, *19*(21). https://doi.org/10.3390/ijerph192114173

[11] Mekler, A. (2021). The Effects of TikTok Use on College Student Learning. In *Undergraduate Review* (Vol. 16). https://vc.bridgew.edu/undergrad_rev/vol16/iss1/19

[12] Nugrahaningsih, N., Porta, M., & Klašnja-Milićević, A. (2021). Assessing learning styles through eye tracking for e-learning applications. *Computer Science and Information Systems*, *18*(4), 1287–1309. https://doi.org/10.2298/CSIS201201035N

[13] van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. In *Learning and Instruction* (Vol. 20, Issue 2, pp. 95–99). https://doi.org/10.1016/j.learninstruc.2009.02.009

[14] Ariel, R., & Castel, A. D. (2014). Eyes wide open: Enhanced pupil dilation when selectively studying important information. *Experimental Brain Research*, *232*(1), 337–344. https://doi.org/10.1007/s00221-013-3744-5

[15] Beatty, J. (1982). Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. In *Psychological Bulletin* (Vol. 91, Issue 1).

[16] Darvishi, A., Khosravi, H., Sadiq, S., & Weber, B. (2022). Neurophysiological Measurements in Higher Education: A Systematic Literature Review. *International Journal of Artificial Intelligence in Education*, *32*(2), 413–453. https://doi.org/10.1007/s40593-021-00256-0

[17] Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly journal of experimental psychology (2006)*, *64*(10), 1971–1989. https://doi.org/10.1080/17470218.2011.588335