Estimation for cognitive load in Video-based learning through Physiological Data and Subjective Measurement by Video Annotation

InHye Lee
Ewha Womans University
22netaurus@gmail.com

ABSTRACT

When designing a video-based learning such as MOOC, it is very important to understand the cognitive aspects of learning and reflect them in the design. Many studies use subjective and physiological data as indicators of cognitive load. To fully understand the cognitive load, we need to understand both of them simultaneously. Therefore, this study is to investigate whether eye data(Mean Pupil Dilation, Mean Fixation Duration) predicts subjective cognitive load during video learning. Furthermore, as a second research question on a broader scale, we examined whether eye data predicts high and low states of subjective cognitive load during video learning. Through this, we expected to find the possibility of Video Annotation and Eye data as a way to measure Cognitive Load during video learning. The experiment was conducted in a controlled laboratory environment with 100 students. In the video learning situation, the learner's eye data was measured using an eye tracker. Immediately afterwards, a video annotation(VA) interview technique was used to put markers according to the cognitive load types such as A(Understanding), B(Easy), C(Complicated), and D(Discomfort). The collected data will be analyzed by Support Vector Machine, a machine learning technique that is considered appropriate for the physiological data.

Keywords

Video-based learning, Physiological data, Eye data, Video Annotation, Eye tracking, Cognitive Load, Support Vector Machine

1. INTRODUCTION

Recently video-based learning has become a common form of learning for both corporate and school education as well as open contents such as MOOCs and Coursera. However, since instructors and learners are separated in time and space in video-based learning, it is difficult to immediately reflect learner's response to the instructional design. In addition, universal instructional design does not reflect the characteristics of each learner. For this reason, universal instructional design in video-based learning tends to result in learner neglect or dropout, as can be seen in MOOC's high dropout rate. Therefore, instructional design considering the learner's learning process is important in video-based learning.

Cognitive load is the one of the most remarkable factors in human learning process. In many studies, including Moreno's work[8], we have accumulated evidence that cognitive load is a reliable factor for effective video-based learning design. In addition, various data left by learners during video-based learning are important resources for instructional design considering individual cognitive load. However, most pedagogical studies measure the learner's internal processes using a psychometric scale. This approach has problem to be solved that memory distortion may occur because it is usually measured after learning. Hence, physiological data are used as an objective measurement index. Especially, eye data can be measured in real time. Moreover, in case of pupillary reflex, it is under control of autonomous nervous system and cannot be voluntary controlled by the subject. However, despite its advantages, it is sensitive to environmental variations such as luminance [3][5][18]. Therefore, using both psychometric subjective scales and eye data can complement each other.

In this study, we will examine how the physiological data predicts the subjective measurements of learners using Support Vector Machine which is machine learning techniques. Also, in case of subjective measurement, video annotation is used to prevent memory loss after learning. This study proceeded with Video Annotation right after eye tracking experiment. This study aims to discover the possibility of using both of physiological data and Video Annotation to measure cognitive load reliably. Through this study, we expect that indicators of cognitive load will be used as more reliable resources for video-based instructional design.

2. LITERATURE REVIEW

2.1 Cognitive Load Theory

According to the Cognitive Load Theory proposed by Sweller[9], Cognitive Load is defined as the sum of mental activities imposed on human working memory when processing new information. There are three types of cognitive loads, and each type of cognitive load is additive. First, external cognitive load is a cognitive load imposed by inappropriate instructional design, which can be controlled by efficiently structured and designed learning environment and tasks[11].

Since external cognitive load is not a desirable load for learning, it needs to be minimized through instructional design. Second, the intrinsic cognitive load is the cognitive load caused by the element interactivity of the learning contents. As the task becomes more complicated, the intrinsic cognitive load felt by the learner increases. Finally, the germane load is the cognitive load used to handle schema acquisition and automation in learning. Germane

In-Hye Lee "Estimation for cognitive load in Video-based learning through Physiological Data and Subjective Measurement by Video Annotation" In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 785 - 789

load is a very important load to facilitate learning. If these three types of cognitive loads exceed working memory, information processing, including learning, will be threatened. Therefore, the purpose of instructional design is to minimize the external cognitive load and increase the intrinsic load within the learner's memory capacity. Instructional design considering each type of cognitive load presented by Cognitive Load Theory can provide more customized learning for individual students[19].

Cognitive Load Measurement Methods

The development of a method to measure cognitive load more effectively plays an important role in the study of cognitive loads [10]. Researches measuring cognitive load are largely divided into subjective and objective measurement methods. First, the subjective measurement method is measured through rating scales. This subjective item tool has the advantage of being able to measure each element of cognitive load separately. It is also widely used to measure cognitive load in many studies because it can be measured by learners relatively easily[13][14]. Cognitive Load mainly uses subject ratings of mental effort and task difficulty as indicators of cognitive load[12]. However, the subjective scale has a disadvantage in that the memory is distorted because it is influenced by the learner's bias and occurs after learning. In addition, there are objective measurement methods measuring through physiological data using EEG, eye data and skin sensitivity. In particular, the eye data measured by the eye tracker has attracted attention in recent studies[18]. Since eye data is measured at the moment of learning, it can be measured without affecting learning process. However it has a problem to be solved that it is sensitive to the external environment.

The preceding studies comparing subjective data and eye data are as follows. Korbach, Brünken and Park[6] set up three groups that distinguish external, intrinsic, and intrinsic cognitive loads, and then identified the differences through rhythm methods, subjective ratings, and eye data. As a result, both objective and subjective measures significantly distinguished the differences between groups. Most of the studies that classify cognitive load use subjective measurement methods or use both objective and subjective measurements together. However, studies that deal with both subjective and objective data should be preceded more by classifying cognitive load types in various learning contexts.

2.2 Eye data

This study utilizes Pupilary data and Fixation Duration among eye data. For Pupil data, Mean Pupil Dialation(MPD) is used. MPD has been used as a reliable indicator of cognitive load. In most cases, MPD expands in according to increasing cognitive load[1]. However, as mentioned earlier, it needs to be careful when collecting because it responds to not only psychological changes but also visual stimuli caused by environmental changes. Marquart & Winter[7] measured cognitive loads while the driver was driving using blinking eyes, eye fixation, and pupil dilation indicators. As a result, the expansion of the pupil size was observed statistically when the workload occurred during operation. Fixation Duration also can be used to measure the attention that individuals have paid to stimuli which means that Fixation Duration can be one of factors that increase cognitive load[16].

2.3 Video Annotation

Video Annotation is a kind of retrospective technique. Retrospective technique is a follow-up observation method that uses visual or auditory clues to access a subject's memory and recall the thoughts and strategies that occurred while performing a specific action or task[2]. The video annotation presents the learning video to the learner immediately after the learning ends for recall. Learners stop the video where they want it and record their thoughts[4]. It is a promising approach to facilitating video retrieval but also it can avoid the intensive labor cost of pure manual annotation[17].

The study used Techsmith's Morae software for video annotation. Notes settings were A(Understanding) which indicates germane load, B(Easy) and C(Complicated) which indicate low and high intrinsic load and D(Discomfort) for extraneous cognitive load.

3. METHODS

3.1 Research Question

The research questions of this study are as follows.

1. Does eye data (pupil response, eye fixation duration) predict subjective cognitive load during video-based learning?

2. Does eye data (pupil response, eye fixation duration) predict the total amount of subjective cognitive load(high and low) during video-based learning?

3.2 Experiment Settings

This study was approved according to the Institutional Review Board (IRB) Institutional Review Board (EBH), and was conducted on 100 male and female college students.

3.2.1 Procedure

This study was conducted in the Edutech Convergence laboratory at Ewha Womans University in order to provide an optimal environment in consideration of illuminance and noise that affect measurement data. The window in the laboratory was covered so that the laboratory was not affected by light intensity. The height of the chair and the pedestal were adjusted to each subject just before measurement, so that the environment of the participant's pupils was accurately tracked. The experiment time for each subject lasted about 100 minutes and one person at the same time.

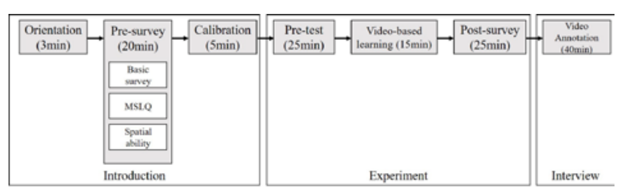

Figure 1. Experiment procedure

As shown in Figure 1, the study purpose, the duration of the experiment, and the precautions related to the experiment were announced before entering the controlled experiment site. Subsequently, the basic personal information and learning motivation strategy which is called MSLQ[15]. They were collected through the 5 Likert questionnaire. After the survey, the subject enters a laboratory where light and noise are controlled. The controlled laboratory is shown in Figure 2.

Figure 2. Environment setting for eye tracking

Figure 3. Environment setting for video annotation

According to Tobbi eye tracker's manual, it was recommended to keep the distance between the study subject and the measuring device constant for accurate measurement. Chin pedestal was used to maintain constant distance. Concerned about an increase in the subject's fatigue, they were guided to rest for 30 seconds at the end of each step. In each step (pretest, video study, posttest), the sequence was set to stare at the front for 10 seconds to find the baseline of the pupil.

As shown in Figure 3, video annotation was performed in the room prepared for the interview. To implement a special interview method called video annotation, we used TechSmith's Morae program.

3.2.2 Participants

This study estimates the cognitive load in the video-based learning situation and recruits the most accessible adults (college students) in the video-based learning context such as MOOCs and Coursera. For accurate measurement, only those who do not have eye-related diseases and who can replace glasses when wearing lenses were allowed to participate in the experiment. In addition, the gender and major categories were selected to be evenly distributed. Even if there were no eye-related diseases, if the eye tracker did not track the eye during calibration, the ear was taken because no data could be collected. In addition, subjects whose data exceeded the recommended range for eye tracking during pretreatment or missing more than 50% of the data due to missing values based on pupil range outliers were excluded from the analysis.

Of the total 100 participants, 96 participated endlessly without returning home halfway. Among them, 82 were studied except for missing values. Thus, 18% of the participants were excluded from the analysis. The demographic information of the study subjects used for analysis in this study is as follows.

Table 1. Demographic Information of Subjects(n=82)

Gender	Male	39
	Female	43
Major	Liberal arts	42
	Science and Engineering	40

3.2.3 Stimuli

The stimuli given to the subject during eye tracking are as follows.

Pre-test and post-test

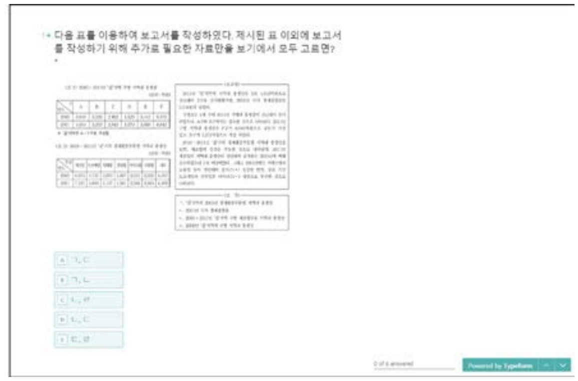

Figure 3. A screen shot of pre-test

The pre- and post-examination tests consisted of PSAT questions, a test to select South Korean civil servants. The problem is that both the pre- and post-test have a total of six questions and the time limit is 25 minutes.

Video-based learning

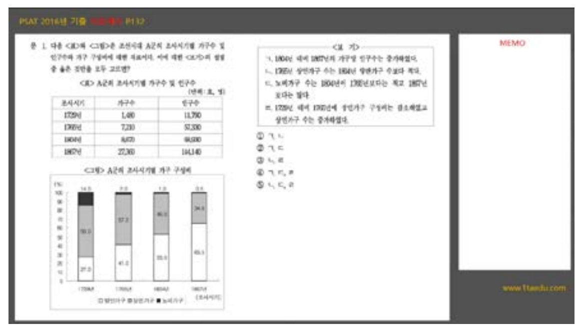

Figure 4. An example of video screen

As shown in Figure 5, video-based learning is a learning video that teaches PSAT problem solving strategies. In order not to distract the learners' attention other than the contents of the study, the lecturer was selected as a video in which the lecture was conducted only by voice.

3.2.4 Instruments for measurement

Eye tracker

In this study, Tobii Pro X2-30 (30Hz) eye tracker was used to measure 30 frames per second to measure pupil response and eye fixation duration. As can be seen in Figure 2, pupils can be measured non-intrusively simply by keeping the 50 ~ 70cm distance between the measuring device and the subject without additional wearing. Therefore, data can be collected without pre- and post-testing and video-based learning. The collected data was extracted in the form of csv data that can be analyzed using Tobii Pro Studio program, Excel and R Studio program.

Morae for video annotation

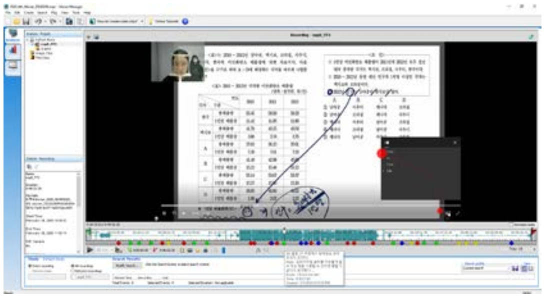

Figure 5. A screenshot of Morea for video annotation

Figure 6 shows an example of TechSmith's Morae program. As you can see from the enlarged figure, Morae program screen is replayed by the learner at the time. In the upper left corner, the face image of the learner was taken in the previous experiment. The screen shows the points and paths that the subject looked at then with red dots and lines..

The learner entered the thoughts and feelings he had heard at the moment while looking at the screen, his gaze, and facial expressions that he stared at in the previous experiment. Press Ctrl + M to select one of A (Understanding), B (Easy), C (Complex), or D (Discomfort), and then type the sentence directly for the reason. During the direct entry by the subject, the researchers looked at the screen together and helped when confused about which markers to choose.

3.3 Data Analysis

3.3.1 Datasets

The variables selected for the research question are shown in Table 2 below.

Table 2. Variables(Eye data and video annotation)

Variables		Description
Eye data	LocalTime Stamp	Timestamp counted from the start of the recording
	GazeEven tType	Type of eye movement event classified by the fixation filter settings
	GazeEvent Duration	Duration of an eye movement event
	Distance Right/Left	Distance between eyes and the eye tracker
	Pupil Right/Left	Estimated size of the right(left) eye pupil
	Validity Right/Left	Indicates the confidence level that eyes have been correctly identified
Video Annot ation	Elapsed Time	Timestamp when subjected noted markers
	Details	Indicate markers which subjects noted

3.3.2 Data pre-processing

As mentioned earlier, even if the experiment was completed, the subjects whose data were more than 50% that could not be used for analysis due to missing values in the preprocessing process were excluded from the analysis. Eye data was excluded from the analysis when the recommended distance was out of 50 ~ 70cm. In the case of the pupil size, the difference between the pupil size of the left and the right is more than 0.4 mm and was also determined as the pupil portion.

In the case of the pupil size, the average of the baseline measured in each section was calculated, and then derived by subtracting the baseline from the measured pupil size. Therefore, because the pupil size measured every second is the baseline minus, the pupil expansion indicators are positive when the pupil is larger than the baseline, but may be negative when the pupil is reduced.

After that, the LocalTimeStamp and Elapsed Time variables of the eye data were changed to the same time expression, and then preprocessed by matching the note and eye data shown by time zone. Therefore, the variables used for the actual predictive analysis are as follows: Markers from Video Annotation (VA), Mean Pupil Dilation (MPD), Mean Fixation Duration (MFD)

The first research question analyzed SVM classification of preprocessed data. The second study divided the A (Understanding), C (Complicated), and D (Discomfort) groups with high cognitive loads and the B (Easy) markers with low cognitive loads. Since the high group of the three markers combined had more than three times the number of data, we randomized and set the same number as the low group. After that, we checked whether the high and low groups were predicted.

3.3.3 Data Analysis

The analysis was performed using R Studio, a statistical analysis program. The preprocessed data is analyzed by Support Vector Machine (SVM) technique. The reason for analysis by SVM method is as follows. First, because of experimental data characteristics. Since eye data is measured at 30 frames per second for 82 subjects, a very large amount of data is collected. In addition, due to the nature of the physiological data, even a laboratory set up is susceptible to microenvironmental influences. Therefore, we chose SVM that is less affected by outliers and has higher accuracy. Second, SVM has the advantage of less overfitting than other neural network techniques. Finally, it is suitable for the markers (A, B, C, D) variables obtained through video annotation because they can be classified and predicted simultaneously.

The first research question analyzed SVM classification of preprocessed data. The second study divided two groups with high cognitive loads which consists of A (Understanding), C (Complicated), and D (Discomfort) and low cognitive loads which consists of B (Easy) markers with low cognitive loads. Since the high group of the three markers combined had more than three times the number of data, we randomized and set the same number as the low group. After that, we checked whether the high and low groups were predicted.

5. REFERENCES

[1] Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of Psychophysiology*, *2*, 142-162.

[2] Colasante, M., & Douglas, K. (2016). Prepare-Participate-Connect: Active learning with video annotation. Australasian Journal of Educational Technology, 32(4). DOI= https://doi.org/10.14742/ajet.2123

[3] Di Stasi LL, Antolí A, Cañas JJ. 2011. Main sequence: an index for detecting mental workload variation in complex component analysis study. *International Journal of Psychophysiology*, *77*(1), 1-7. DOI= https://doi.org/10.1016/j.ijpsycho.2010.03.008

[4] Ethel, R. G., & McMeniman, M. M. 2000. Unlocking the knowledge in action of an expert practitioner. Journal of Teacher Education, 51(2), 87-101. DOI= https://doi.org/10.1177/002248710005100203

[5] Jainta, S., & Baccino, T. 2010. Analyzing the pupil response due to increased cognitive demand: An independent tasks. Appl Ergonomics 42. 807–813. DOI= https://doi.org/10.1016/j.apergo.2011.01.003

[6] Korbach, A., Brünken, R., & Park, B. 2018. Differentiating different types of cognitive load: A comparison of different measures. *Educational Psychology Review, 30*(2), 503-529.

[7] Marquart, G., Cabrall, C., & de Winter, J. 2015. Review of eye-related measures of drivers' mental workload. Procedia Manufacturing, 3, 2854-2861.

[8] Moreno, R. 2010. *Cognitive load theory: Historical development and relation to other theories*. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), Cognitive load theory, 9–28. Cambridge University Press. DOI= https://doi.org/10.1017/CBO9780511844744.003

[9] Sweller, J. 1988. Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257-285.

[10] Sweller, J. 2018. Measuring cognitive load. *Perspectives on medical education*. 7(1). 1-2. DOI= https://doi.org/10.1007/s40037-017-0395-4

[11] Sweller, J., Ayres, P., & Kalyuga, S. 2011. Cognitive Load Theory. New York, NY, US: Springer.

[12] 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. DOI= https://doi.org/10.1037/0022-0663.84.4.429

[13] Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. 1994. Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79(1), 419-430.

[14] Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. 2010. Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63-71. DOI= https://doi.org/10.1207/S15326985EP3801_8

[15] Paul R. Pintrich, Davide A. F. Smith, Teresa Garcia, and Wilbert J. McKeachie. 1991. A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)

[16] Vertegaal, R., & Ding, Y. (2002, November). Explaining effects of eye gaze on mediated group conversations: amount or synchronization?. In Proceedings of the 2002 ACM conference on Computer supported cooperative work. 41-48.

[17] Wang, Q., Yang, S., Liu, M., Cao, Z., & Ma, Q. 2014. An eye-tracking study of website complexity from cognitive load perspective. Decision support systems, 62, 1-10.

[18] Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. 2018. Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in hearing*, 22. 1-32. DOI= https://doi.org/10.1177/2331216518800869

[19] Young, J. Q., Van Merrienboer, J., Durning, S., & Ten Cate, O. 2014. Cognitive load theory: implications for medical education: AMEE Guide No. 86. Medical teacher, 36(5), 371-384.