

Personalized Comment Reviewing in Active Video Watching: Investigation of Learners' Cognitive Load

Ehsan BOJNORDI*, Antonija MITROVIC, Matthias GALSTER, Sanna MALINEN & Jay HOLLAND

University of Canterbury, Christchurch, New Zealand

*ehsan.bojnordi@pg.canterbury.ac.nz

Abstract: In the context of video-based learning, particularly active video watching, social learning is facilitated by allowing students to review comments written by their peers. In previous studies with AVW-Space, all students had access to the same comments, despite differences in their knowledge. This paper presents a study in which each student received a personalized list of comments to review, based on their student model. The results show that our intervention encouraged learners to think deeper about the comments during reviewing comments.

Keywords: Active Video Watching, Comment Reviewing, Recommendation Systems

1. Introduction

Video-based Learning (VBL) offers flexibility in terms of time and place, and can replace conventional learning approaches (Sablić et al., 2021). However, merely watching videos is a passive activity that leads to limited engagement and low-quality learning (Mitrovic et al., 2019). Hence, Active Video Watching (AVW) was proposed to promote engagement and self-reflective learning (Mitrovic et al., 2017). One way to support engagement in AVW is to write comments on videos. AVW also supports social learning by asking students to review and rate comments written by their peers (Mitrovic et al., 2017), and by allowing students to respond to comments (Bojnordi et al., 2023). However, the task of comment reviewing/rating might be overwhelming and time-consuming for some learners (Mohammadhassan et al., 2022). Therefore, personalizing comments can be beneficial for learners with different skill levels.

We developed a recommender system to select comments adaptively based on the student model. The system first identifies a given learner's areas of weakness, and then selects comments covering those. In this paper, we investigate the impact of our intervention on learners' cognitive load. We used AVW-Space as the platform for our experiments. We invited first-year engineering students at the University of Canterbury to use this platform to improve their presentation skills. Two groups of students from the years 2023 and 2024 were provided with two different versions of the platform. Students first completed Survey 1, which consisted of questions about demographics, training, experiences with presentation skills, and a pre-test of domain knowledge. After that, students watched eight videos and commented on them. In the next phase, students could review, rate, and respond to comments. In 2024, AVW-Space provided personalized comment recommendations, while in 2023 all students received the same set of 238 comments. At the end of the study, students completed Survey 2, which included the NASA Task Load Index (NASA-TLX) questionnaire (Hart, 2006) to analyze the cognitive load while reviewing comments, along with additional open-ended questions about the usefulness of reviewing and rating comments.

2. Methodology

To model the domain (presentation skills), we used the ontology developed by Dimitrova and Mitrovic (2022). This ontology comprises 344 Knowledge Components (KCs) organized in seven levels. The four KCs at the top level are *Delivery*, *Structure*, *Visual Aids*, and *Presentation Attributes*. We implemented the student model as an overlay on top of the ontology. We also prepared a list of phrases (vocabulary) associated with each KC of the ontology. This allows us to monitor the student's knowledge state of each KC. The initialization of the student model is done based on the student's responses to the pre-test in Survey 1 and updated whenever they utilized relevant phrases in their comments/responses. To estimate the score for each KC in the student model, we used the following formulas.

$$S_{i=0}(KC) = \begin{cases} 0.1 & \text{if } KC \text{ appears in pretest,} \\ 0 & \text{otherwise.} \end{cases}$$

$$S_{i+1}(KC) = Q \times 0.1 + S_i(KC)$$

Here, Q represents the quality of the comment/response as proposed by Mohammadhassan et al. (2022). The score of a parent KC node in the student model is calculated as the average score of its children nodes.

The strategy for the Comment Recommendation System (CRS) was to select KCs with the lowest scores in the student model. Each time a student visits a review page, the CRS considers the scores of the four main KCs in the student model and selects the one with the lowest score as the Area of Recommendation (AoR). The CRS then targets the KCs within the AoR in which the student's score is below the threshold of 0.7, designating them as Targeted KCs (TKCs). Next, it searches for comments containing at least one TKC and adds them to a pool. Afterwards, the CRS calculates the fitness function for each comment in the pool, where $C(KC, j)$ represents the condition that KC is in comment j .

$$\forall KC \in TKC, C(KC, j) \rightarrow \text{Fitness}(j) = \sum_{KC} (1 - S(KC))$$

Finally, the CRS sorts comments based on their fitness scores and then publishes them to the student. The top 10 are visible initially, with the rest shown upon request for more recommendations.

3. Results

We requested feedback from students regarding the reviewing and rating of comments across four cognitive load factors: *Mental Demand*, *Effort*, *Frustration* and *Performance*, using a NASA-TLX instrument. The scores were based on a Likert scale from 1 (very easy) to 20 (very hard) for the first three factors, and from 1 (very poor) to 20 (very good) for the last factor.

Table 1 shows the means and standard deviations of scores received from two groups of students. The results imply that reviewing and rating comments were more mentally challenging for students in 2024, suggesting that personalized comment recommendations led learners to think deeper about what they reviewed and invest greater effort into understanding their peers' points of view.

We also asked students to respond to the following question: '*What would be the usefulness of reviewing comments on the videos made by others?*'. Out of 142 responses obtained from participants exposed to the intervention, 97 (68.31%) indicated that it helped deepen their understanding, identify missing aspects of the videos, and expand their learning. In 22 (15.50%) cases, students stated that it prepared them for self-evaluation, validation, and reflection on their own interpretations. Ten students (7.04%) pointed out that it encouraged them to read more comments and think critically about the comments and responses of their

peers. There were only 13 responses (9.15%) expressing participants' dissatisfaction with the comment reviewing task.

Table 1. Comparison of Cognitive Load Factors – Mean (Std)

	2023 (n=501)	2024 (n=167)	Significance		
			t	p-value	Cohen's d
Mental Demand	5.30 (4.38)	6.96 (3.99)	-4.54	< .001	.39
Effort	5.01 (4.15)	6.26 (3.75)	-3.65	< .001	.26
Frustration	5.60 (4.99)	6.90 (4.89)	-2.98	< .01	.26
Performance	11.61 (4.76)	11.07 (4.06)	1.45	.15	.12

4. Conclusions

We developed a comment recommendation system for the task of reviewing comments in AVW, utilizing an overlay ontology-based learner model. Our intervention prompted learners to consider points and aspects they had not previously thought of or might have missed while watching the videos, encouraging them to think more deeply during reviewing comments. This impact was clearly reflected not only in students' cognitive load factors but also in the explicit feedback we gathered from learners.

There are several limitations of our study. First, generalizability of our results might be limited due to the specific target population (first-year engineering students) and the relatively small sample size in the 2024 group. Second, the study relied on self-reported measures (NASA-TLX), which, while validated, may introduce bias due to participants' subjective interpretations of their experiences. Third, the potential impact of external factors such as learners' prior knowledge or varying levels of motivation on cognitive load should be considered.

For future work, we aim to add explanations to clarify the reasoning behind comment recommendations based on the learner model and ontology, and explore to what extent this intervention can impact learners' cognitive load.

References

- Bojnordi, E., Mitrovic, A., Galster, M., Malinen, S., & Holland, J. (2023). Adding Interactive Mode to Active Video Watching. *Proc. 31st Int. Conf. Computers in Education*, Vol. II, pp. 1042-1044. Asia-Pacific Society for Computers in Education.
- Dimitrova, V., & Mitrovic, A. (2022). Choice architecture for nudges to support constructive learning in active video watching. *Int. Journal of Artificial Intelligence in Education*, 32(4), 892-930.
- Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proc. Human Factors and Ergonomics Society Annual Meeting*, 50, 904 - 908.
- Mitrovic, A., Gordon, M., Piotrkowicz, A., & Dimitrova, V. (2019). Investigating the effect of adding nudges to increase engagement in active video watching. *Proc. 20th Int. Conf. Artificial Intelligence in Education* pp. 320-333.
- Mitrovic, A., Gostonski, P., Herritsch, A., & Dimitrova, V. (2017). Improving presentation skills of first-year engineering students using active video watching. *Proc. 28th Annual Conf. of the Australasian Association for Engineering Education*, pp809-816.
- Mohammadhassan, N., Mitrovic, A., & Neshatian, K. (2022). Investigating the effect of nudges for improving comment quality in active video watching. *Computers & Education*, 176, 104340.
- Sablić, M., Miroslavljević, A., & Škugor, A. (2021). Video-based learning (VBL)—past, present and future: An overview of the research published from 2008 to 2019. *Technology, Knowledge and Learning*, 26(4), 1061-1077.