








Characterizing Comment Types and Levels of Engagement in Video-Based Learning as a Basis for Adaptive Nudging

Yassin Taskin¹ , Tobias Hecking¹ , H. Ulrich Hoppe¹ ,
Vania Dimitrova² , and Antonija Mitrovic³ 

¹ University of Duisburg-Essen, Duisburg, Germany
`taskin@collide.info`

² University of Leeds, Leeds, UK

³ University of Canterbury, Christchurch, New Zealand

Abstract. Video is frequently used as a learning medium in a variety of educational settings, including large online courses as well as informal learning scenarios. To foster learner engagement around instructional videos, our learning scenario facilitates interactive note taking and commenting similar to popular social video-sharing platforms. This approach has recently been enriched by introducing nudging mechanisms, which raises questions about ensuing learning effects. To better understand the nature of these effects, we take a closer look at the content of the comments. Our study is based on an ex post analysis of a larger data set from a recent study. As a first step of analysis, video comments are clustered based on a feature set that captures the temporal and semantic alignment of comments with the videos. Based on the ensuing typology of comments, learners are characterized through the types of comments that they have contributed. The results will allow for a better targeting of nudges to improve video-based learning.

Keywords: Learning analytics · Video-based learning · Learner engagement · Adaptive nudging

1 Introduction

Being an integral part of digital learning, videos are utilized to enhance the educational experience and increase student satisfaction [17] in a broad range of educational settings, e.g. MOOCs, flipped classroom, informal learning. Social video-sharing platforms, such as YouTube, are becoming the first source for learners when they want to learn something new. However, watching videos is a passive activity, often resulting in a low level of engagement which hinders the effectiveness of video-based learning [2, 18]. Automatic engagement detection can inform personalized interventions, e.g. motivational messages, questions, reminders, to prevent disengagement and to enhance the learning experience.

In previous work, we developed AVW-Space, a controlled platform for informal video-based learning through note-taking where students watch and comment on videos [21]. Several studies using AVW-Space with undergraduate and postgraduate university students in the context of presentation skills have shown that students who write comments instead of passively watching improve their understanding of presentation skills.

Gathering requirements for the effectiveness of AVW-Space in an earlier study indicated the need to encourage student commenting, which was positively correlated with conceptual learning, while at the same time preserving students' freedom to interact with videos in a way they prefer [21]. This informed the extending of AVW-Space by adding nudges [11] - interventions that influence people's behavior to make beneficial choices (paternalism) in a non-compulsory manner (libertarian). Two forms of nudges have been implemented [22]: signposting (interactive visualizations showing video intervals where past students have commented) and personalized prompts (noting student's commenting behavior or showing example past comments). A user study with the extended AVW-Space [22] indicated a significant increase in comment-writing in the nudging condition, thus providing evidence that the nudges encouraged student commenting. However, this did not lead to significant improvement of the students' conceptual knowledge.

Therefore, there is a need to gain a deeper understanding of the students' cognitive engagement while interacting with videos in order to better assess the impact of nudges in AVW-Space. This is the prime goal of the research presented here. It lays the grounds for more adaptive and selective nudging by addressing the following research questions:

- RQ1:** How can student comments be characterized with regard to cognitive engagement?
- RQ2:** Are there any notable individual differences with regard to commenting and cognitive engagement?

To answer these questions, we first differentiate levels of engagement in learner produced comments. By analyzing the content of the comments, we classify comments by distinguishing "shallow" types of engagement, such as echoing or affirmation, from deeper elaborations that draw associations, summarize, compare or transform the given material. In the second step, we project the comment classification back to the learners, i.e. we characterize learners through their specific set of comment types. This allows us to explore the dependencies between commenting behavior and individual learning characteristics and personal traits.

The paper is structured as follows: We start by providing an overview of related work on engagement detection specially for learning with videos, followed by a brief description of the experimental setting in which the data were collected. The data analysis reported in Sect. 4 includes the initial classification of comments through clustering, as well as the ensuing characterization of learners and their engagement levels. Finally, we reflect the findings in relation to background theories and potential applications.

2 Background

The work presented here falls in the broad area of analyzing engagement in digital learning. Generally, engagement analytics approaches utilize the vast amount of data collected while students interact with the system. There is an established research stream on predicting behavior that can have adverse effect on learning, such as quitting in systems that embed free learning tasks (e.g. reading [20] and solving problems [16]), disengagement in MOOC courses [2, 6, 18], and ‘gaming the system’ (i.e. taking advantage of system’s properties to superficially complete the task) [4]. Another stream of work looks at detecting engagement aspects that can be linked to cognition, such as zoning out and mind wandering [5, 12], and information seeking/giving [14]. Thirdly, the affective response to instructions (e.g. frustration [26] and confusion [1]) was also studied. These engagement behaviors, e.g. quitting, mind-wandering, zoning out, capture a rather ‘shallow level’ of engagement which does not show how the learner engages with the educational material. In contrast, our work investigates deeper cognitive levels of engagement by characterizing content (comments) produced by learners.

The prime focus of our work is engagement analytics for improving video-based learning. Existing research analyzes data about the learners’ interaction with and navigation of videos by analyzing play, pause, and seeking actions and which parts of the video are most important [7, 13, 19]. Other works focus on students’ reflections on videos, using their comments to determine students’ conceptual understanding of the specific topics [10, 15]. While the actual content of the video provides valuable information that can be analyzed using text mining methods on the video transcripts [3], the relation of both - student generated content and video content - has not been investigated. We call this relationship “semantic alignment”, and provide computational means for its measurement. This helps us to better understand to what extent knowledge conveyed in the video is taken up by students. Moreover, this enables automatic differentiation (without manual knowledge engineering) between deeper engagement with the course material and shallower student contributions noting points in the videos.

Hence, our work contributes to research on engagement in video-based learning, e.g. [1, 18], with a specific focus on cognitive engagement. We build on the ICAP framework [8] to link cognitive engagement activities to observable behaviors. While ICAP has been used to categorize information seeking in MOOC forums [14, 27], its adoption for analyzing video engagement is novel. In our adoption, we have shifted the focus from behavioral aspects derived from interaction log files (e.g. [9]) to characterizing engagement based on learner-generated content. In our scenario (see next section), the primary unit of study are video comments. In turn, we use the classification of comments (i.e. comment types) as a means to characterize learners and their level of cognitive engagements, which leads to a specific adaptation of the ICAP framework.

The screenshot displays the AVW-Space interface for a video titled "TUTORIAL 4: The five secrets of speaking with confidence" by Christine Clapp. The video player shows a woman speaking, with a progress bar and a volume icon. A "Well done!" nudge is overlaid on the video, stating: "Great job using all aspects to comment on the video!". To the right, a "Your previous comments" panel lists several comments with timestamps and aspects:

- 00:39: The body language, facials and actions should be suitable and match what you are talking about whilst still keeping the audiences attention. Aspect: I did/saw this in the past
- 01:55: avoid tapping and other distracting noises. Aspect: I didn't realize I wasn't doing this
- 03:00: use your voice as a prop. Aspect: I like this point
- 03:46: smile when appropriate. Aspect: I am rather good at this

Below the video player, there is a comment input field with an "Add comment" button and a list of previous comments with their respective aspects.

Fig. 1. Screenshot of AVW-Space, showing a *Diverse Aspects* nudge.

3 Experimental Setting

Educational Context. Our data was collected in the context of a large (1039 students) first-year course on fundamental engineering skills at the University of Canterbury. In this course, students work on a group project, and present their results in the last week of the course. Presentations are marked by two human tutors, who provide two group scores (one for the content of the presentation, the other for visual aids, both with the maximum of 5 marks), as well as an individual mark for each student (with the maximum of 5 marks). Due to an already full curriculum, there was no time in the course to teach students how to give presentations. Instead, the students were directed to AVW-Space as an online resource for presentation skills. The AVW-Space instantiation included eight short videos [22]: four tutorials on how to give presentations, and four videos showing example presentations. We limit analysis in this paper to tutorial videos only, as this is a common form of video content widely used for informal learning in a variety of educational contexts. The learning consisted of students watching and commenting on the videos individually.

Platform. The study involved two versions of AVW-Space. The **control condition** included watching and commenting on videos without any intervention from the system. The only support was the offering of reflective micro-scaffolds (aspects) - in addition to entering the text for a comment, students needed to select an aspect indicating the intention of the comment. For the tutorials, aspects (i.e. micro-scaffolds that encourage reflection) were: *I did not realize I was not doing this*, *I am rather good at this*, *I like this point* and *I did/saw this in the past*. The **experimental condition** included an enhanced version of AVW-Space, which additionally provided interactive nudges to enhance engagement, including visualizations and personalized prompts. Interactive visualizations, aimed to support social learning (i.e. learning from peers), are shown below

the video (Fig. 1). The top visualization is the comment timeline; its goal is to provide signposts to the student in terms of previously written comments. Each comment is represented as a colored dot along the horizontal axis, representing the time when the comment was made. The color of the dot depends on the aspect used by the student who wrote that comment. When the mouse is positioned over a particular dot, the student can see the comment (as in Fig. 1). Clicking on a dot begins playing the video from that point. The bottom visualization is the comment histogram visualization; it shows a bar chart representing the number of comments written for various segments of the video. This visualization allows the student to quickly identify important parts of a video, where other students made many comments. These visualizations meet two identified needs: (1) providing social reference points so that students can observe others' comments, and (2) indicating important parts of a video and what kind of content can be expected in those parts, differentiated by aspect colors.

Personalized prompts, which appear next to the video (as in Fig. 1), are designed to encourage students to write comments [22]. For example, reminding the student to make comments when they tend to watch without commenting, encouraging the student to use diverse aspects, or showing examples from past comments to promote attention and stimulate engagement. AVW-Space maintains a profile for each student, used to decide which prompts are appropriate for the learner at a particular time during interaction.

Procedure. All students enrolled in the course were invited to take part in the study. Participants' profile was collected with a *pre-test survey*, including demographic information, background experiences and the Motivated Strategies for Learning Questionnaire (MSLQ) [23]. The survey also contained questions on the participants' knowledge of presentations (conceptual knowledge questions). Students were asked to list as many concepts related to Structure, Delivery and Speech, and Visual Aids as they could. For each of those three questions, students had one minute to write responses. These answers were judged to what extent they covered concepts from an expert-generated domain taxonomy [11]. After a period in which the students interacted with AVW-Space, a *post-test survey* was issued. It included the same questions on knowledge about presentations (to measure change in conceptual knowledge), as well as some usability questions.

Participants. 347 participants have used AVW-Space writing at least one comment, of whom 180 were from the control group (124 males, 55 females, 1 other) and 167 from the experimental group (118 males, 49 females). The majority of participants (79.83%) were native English speakers; most participants (95.39%) were aged 18–23. There were no significant differences between the two groups on their experiences in giving presentations and using YouTube for learning, as well as on MSLQ scales.

Data. The data used for the analysis presented below includes:

- *user-generated data*: for each comment, AVW-Space records the text, selected aspect, the timestamp as well as the cue (i.e. the time in the video when the comment was entered).

- *learner MSLQ profile*: items that are relevant for this study are *intrinsic motivation* (degree of participation in academic activities for reasons of challenge, curiosity, and mastery), *extrinsic motivation* (academic activities mainly for grades and rewards), and *elaboration* (ability to integrate and connect new information with prior knowledge).
- *learning scores*: this includes the presentation scores obtained in the course and the conceptual knowledge (number of concepts named in post- and pre-test surveys)

4 Data Analysis

The first step of the data analysis was to identify different types of student comments. For this purpose, we clustered the comments using a feature set based on the comment content and the time at which a comment was made. This analysis will be presented in more detail in Sect. 4.1. The categorized comments can then be mapped back to the students to characterize these in terms of their commenting profiles. The relation between these profiles and other student variables are investigated in Sect. 4.2.

The data set consists of 1831 student comments. The domain knowledge used in this analysis originated from two sources: (1) a domain taxonomy that was manually created by experts containing key concepts about giving presentations in general, and (2) a set of concepts (per video) based on terms extracted from the videos in a processing chain that involved a speech-to-text transformation followed by term extraction. Using (1), we can determine the number of general domain concepts used in each comment. The video-specific terms (2) allow for a further differentiation: counting the overlap of terms between a comment and the terms extracted from the whole video we get a “global alignment” between the comment and this video. This reflects whether the comment takes up the general theme of the video. Since we know in which parts of the video (on a timeline) the specific terms are used, we can also compute a “local alignment” that only looks at the content of the video around the time the comment was made. This is useful to identify whether the content of a comment reflects the specific focus of the corresponding video section. For this analysis, we used a time window of -30 and $+10$ s around the time of entering the comment.

Our study only relies on the tutorial videos. We had to exclude the example videos, since the presentations in those videos covered various areas (such as medicine or chemistry), and therefore cannot be matched to our general knowledge domain (related to presentation skills). The tutorial videos deal with presentation techniques and do not show this discrepancy. This corpus comprises 1144 comments overall.

4.1 Classification of Comments

We used the K-Means algorithm to cluster the tutorial comments. The features used for the clustering were global and local video alignment, number of domain

specific concepts and the relative time at which a comment was made. All features were normalized to values between 0 and 1. Different cluster counts were explored to spot meaningful differences in clusters which would allow differentiating between comment types. The chosen number of clusters was 7, which identified more distinct cluster types, compared to lower number of clusters. Clustering quality was also assessed based on silhouette analysis [24], which compares the distance of a sample to its own cluster to the distance of the sample to the nearest other cluster, to calculate whether clusters are well separated. Clustering quality improved as we increased the number of clusters to seven clusters. Higher numbers of clusters did not reveal any additional significant differences or increase silhouette scores. The results of the clustering can be seen in Fig. 2. For the first four clusters, comments mention on average a little more than two domain concepts. The local video alignment is low, especially close to the beginning of a video. This is to be expected, since there is not yet much content from the video to compare them to. Global video alignment tracks domain concepts. For the comments in these clusters, both domain concepts and global video alignment generally extract the same concepts with the global video alignment having some false positives. Each of these clusters has around 200 comments, which means these types of comments are made very frequently.

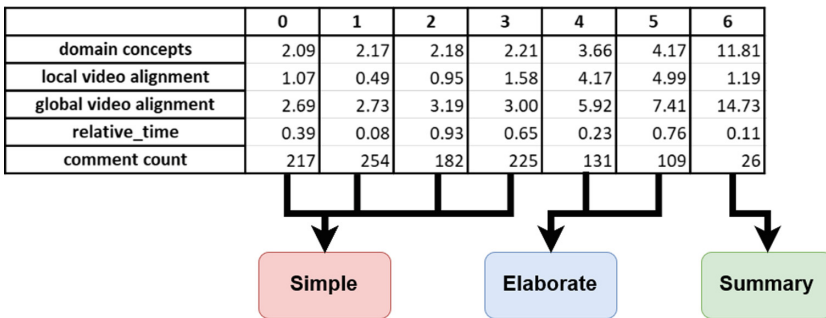


Fig. 2. Comment clusters and mapping to comment types.

Clusters 4 and 5 are different from clusters 0–3 in that they have a higher number of domain concepts. They also have a high local video alignment, showing that students who made these comments strongly engaged with the video content at that particular time. Cluster 5 has somewhat more domain concepts and higher local video alignment than cluster 4. Comments from this cluster are made at about 3/4 of the video time. Around that time, generally the last major point of the video is made, and students have already received a substantial amount of information on the topic. This might make students more confident to talk about the video content and relate it to previous information. This is also indicated by the high global video alignment. In contrast, the comments from cluster 4 are made earlier in the video, roughly at 1/4 of the video time, where students cannot relate the content as much to previously seen information.

Comments from cluster 6 have the highest number of domain concepts and the highest global video alignment. However, their local video alignment is very low, and they are made close to the start of the video. These comments seem to be summaries of the video content. We analyzed the text of comments in different clusters and identified three types of comments.

Simple Comments take up a single point made in the video and generally contain two domain concepts. Example:

“each slide to one idea//people cant read slides and listen.”

Elaborate Comments take up multiple points and elaborate on them, rather than simply repeating the content of the video. They contain a high number of domain concepts and have a high local video alignment. Example:

“I always try to stand up straight, with good posture and little movement as it gives the impression of confidence. I think that by keeping the hands by your side and using them for hand gestures when needed it creates the feeling that the presenter is at ease in the situation.”

Summary Comments are made at the start of the video and summarize the points made in the video. They therefore have a high amount of domain concepts and low local video alignment. Example:

“confident speaker - stance - hip width, stand tall, neutral position, gestures -sound - projection, lower the shoulders, slow understandable pace, tones, loud and quiet, dynamics . . . -sight - eye contact, establish with all parts”

For each cluster, we looked at a sample of the comments and labelled them according to which comment type best fit the pattern of the comments in the sample. This process resulted in clusters 0–3 being labelled as simple comments, clusters 4 and 5 as elaborate comments, and cluster 6 as summary comments.

To analyze the quality of this mapping and compute decision rules which further solidify the differences between the comment types, we trained a decision tree to predict the comment type based on the same features used in the clustering. The resulting decision tree can be seen in Fig. 3. To create the decision tree, we used the CART algorithm [25] with the Gini Impurity as a criterion for deciding how to split samples at each node. The Gini Impurity is 1 for an equal distribution of the classes and 0 if there is only one class represented in a sample. The class distribution of the sample at each node is given in the form [simple, elaborate, summary]. The first split is made on the local video alignment being lower than 3.5. If that is the case, the sample is further split on the number of domain concepts. All comments with less than 8.5 domain concepts are classified as simple comments and comments with a value higher than that are summary comments. The classification works well for simple and elaborate comments, but

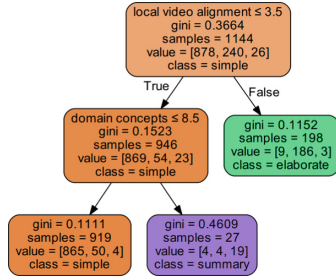


Fig. 3. Decision tree for classifying comments.

has problems separating out the summary comments. This indicates that the decision boundaries between summary and other types of comments are not that clear. Using 10-fold cross validation, the model has an average accuracy of 0.93 which supports that the mapping of the clusters works well.

4.2 Characterizing Learners and Learner Engagement

To characterize learners and their engagement, we count the number of comments of each type that a student made to create student profiles. These profiles are then related to the data collected about students in the form of MSLQ scores, conceptual knowledge pre- and post-tests (c.f. Section 3) and scores for a group presentation the students gave after the learning activity. For the experimental group, we also included the number of nudges a student received. Statistical measures were reported as significant for $p \leq .05$. The number of different comments in the different categories and the number of learners who posted such comments are summarized in Table 1.

Table 1. Statistics of comments in different classes.

Comment type	Comments	Students	Comments/student
Simple	878	179	4.90
Elaborate	240	80	3.00
Summary	26	18	1.44

The majority of comments were of the simple type. There were 193 students in total, of which 179 wrote at least one simple comment. The number of elaborate comments were about 1/4 of simple comments and only 80 users made at least one. Summary comments were made by only 18 students. Thus, this type of comment indicates divergent behavior of a small subset of students. System logs showed that these students fully watched a video then restarted it at which point they made the summary comment, explaining this phenomenon.

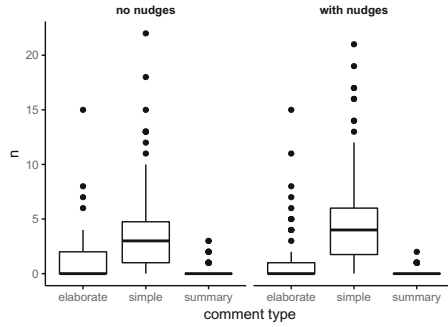


Fig. 4. Distribution of different comment types in both conditions.

Figure 4 shows the distribution of comments for the two groups. The average number of simple and elaborate comments between the two groups is very similar. However, the experimental group has a higher number of simple comments.

Relation Between Nudging and Commenting Behavior. A linear regression for the students in the experimental condition shows that the number of nudges is a significant predictor for the amount of simple comments, ($b = .12, t(106) = 4.67, p < .01$). The predicted number of simple comments is equal to $(3.30 + 0.12 * n_{nudges})$. The regression also significantly explained a portion of the variance ($F(1, 106) = 21.86, p < .01, R^2 = .17$). However, statistically significant relationships between the number of nudges and elaborate or summary comments could not be observed. These results suggest that the implemented nudges animate students to post comments, but often do not trigger deeper reflection on the video content.

Relation Between Commenting Behavior and Student Variables. After the learning activity with AVW-Space, students gave group presentations in the last week of the course. We found a statistically significant correlation between the presentation score and the number of elaborate comments, ($r(191) = .3, p < .01$). Correlations between the presentation score and simple or summary comments were not significant. Extrinsic motivation was negatively correlated with number of elaborate comments, ($r(192) = -.18, p < .01$). In particular, students who are just compliant to meet the external requirements do not tend to invest high effort in commenting. Conversely, if students are motivated less by grades or rewards, they tend to write more elaborate comments. Surprisingly, there was no significant correlation between writing elaborate comments and the MSLQ score for elaboration (as a learning strategy). This suggests that motivational state during the learning activities is a more decisive factor for higher engagement in the commenting task than the personality trait.

With respect to the gain in conceptual knowledge measured as the difference in scores between the post- and the pre-test, it cannot be said that those who

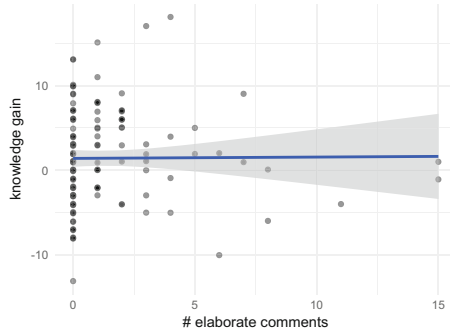


Fig. 5. The total number of elaborate comments of a user does not indicate an increase in conceptual knowledge.

post more elaborate comments have higher gain, as it can be seen from Fig. 5. Only the post-test score correlates significantly with the number of elaborate comments, ($r(144) = .19, p = .05$). However, the figure shows that among the learners who do not write any elaborate comments there is a tendency towards no or a lower increase in conceptual knowledge. This is also depicted more explicitly in Fig. 6. On average, learners who wrote at least one elaborate comment had a higher gain in conceptual knowledge ($M = 2.95, SD = 5.52$) compared to other learners ($M = 0.49, SD = 5.19$). A t-test reports that this difference is significant, $t(144) = -2.71, p = .007$.

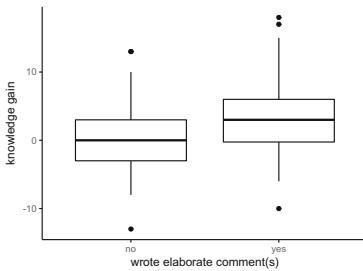


Fig. 6. Comparison of the conceptual knowledge gain of learners with and without elaborated comments.

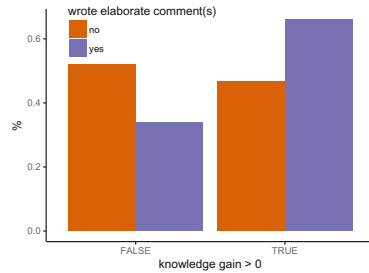


Fig. 7. Learner writing elaborate comments are more likely to increase their conceptual knowledge.

A limitation of these result is that some students did not perform as well in the post-test as in the pre-test, which unreasonably leads to negative values in knowledge gain. This was the case for 52 of the 146 participants who completed the post-test. Reasons can be a ceiling effect (already high pre-knowledge) or a lack of motivation to participate in another test. For these reasons, we have

compared the commenting behavior of learners with positive or no gain separately from the one of learners with the negative knowledge gain. Figure 7 shows the fraction of participants who had increased conceptual knowledge after active video watching among the 56 participants who wrote elaborate comments and among the 90 remaining participants respectively.

Evidently, there is a significant difference in the number of students with an increase in conceptual knowledge in the two groups of participants with and without elaborate comments ($\chi^2(1) = 4.48, p = .034$). The majority (66%) of the 56 students who posted elaborate comments showed increased conceptual knowledge reflecting on concrete issues mentioned in the video while this is only the case for 47% of the participants without such comments. There are two possible explanations for the observation that the learners with elaborate comments did better in the post test than others. Either these findings can be attributed to the overall higher level of engagement of these students during and after video watching, or a deeper reflection of the video content enables them to be also more elaborate in their answers in the post-test.

5 Discussion and Conclusion

Following ICAP [8] as a theoretical framework, we would identify simple commenting with “active learning” (A) and elaborate and summary type of commenting with “constructive learning” (C). The learning activities facilitated and analyzed in our study did indeed not foresee co-construction, so there was no “interactive learning” (I) in terms of ICAP. Different from the original activity-based specification of engagement levels, our analysis is based on artifacts in the form of textual comments. This content-analytic approach is particularly well-suited for dealing learner-generated content and can possibly be extended to other learning scenarios beyond video-based learning.

Distinguishing between learners who have written at least one elaborate comment ($E > 1$) and those who have not ($E = 0$), we have seen that $E > 1$ goes along with higher average knowledge gain, although we could not back the assumption that “more is better”. We have also seen from this and previous analyses that nudging increases the number of comments written, yet not particularly in the “elaborate” category. The finding that extrinsic motivation is negatively correlated with elaborate commenting corroborates the assumption that mere “compliance”, also in response to nudges, does not lead to the desired types of higher-level contributions. Elaborate commenting as a higher form a cognitive engagement appears to be very much driven by motivational state during learning activities rather than by more stable personality traits.

Although our findings indicate that the students’ writing of elaborate comments (interpreted as higher cognitive engagement) goes along with better learning results, we should refrain from interpreting this empirical coincidence as a causality. Yet, the we can refer to the ICAP framework as a theoretical underpinning for the assumption that increasing the degree of elaboration in the students’ commenting behavior would be beneficial for learning. Certainly, this would be

beneficial for the richness and the ensuing affordances of the learning environment. Accordingly, the introduction of nudges should not just stimulate activity but should also support elaboration. The semantic features used to classify comments (including the “alignment” relationship) can serve as conditions for “adaptive nudging” that takes into consideration the learner’s engagement with the video.

The immediate future challenge is to validate the findings, i.e. the comment types and prediction model, and to investigate their generalizability in similar learning contexts. We will apply the analytics approach on another AVW-Space dataset (from another course). The prospect is to enhance the nudges by improving the student modeling mechanism to take into account not just the fact that a student is commenting but also the student’s cognitive engagement as evidenced in the comments.

The work presented here has potentially broader application in digital learning contexts with learner-generated content. Measures of semantic alignment with lecture materials applied to learner-created artifacts (such as notes, comments, forum posts, summaries) can indicate different levels of cognitive engagement and elaboration and thus offer insights to better inform interventions to promote deeper learning. Evidently, this can also be applied to other types of learning materials beyond videos (e.g., textbooks or slide presentations). Our approach can also be extended towards using content analysis to indicate the quality of reflections or critical discussions, thus increasing the system awareness of the students’ learning achievements and needs as a basis for adaptive scaffolding.

References

1. Agrawal, A., Venkatraman, J., Leonard, S., Paepcke, A.: YouEDU: addressing confusion in MOOC discussion forums by recommending instructional video clips (2015)
2. Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J.: Engaging with massive online courses. In: Proceedings of the 23rd International Conference on World Wide Web, pp. 687–698. ACM (2014)
3. Atapattu, T., Falkner, K.: Impact of lecturer’s discourse for student video interactions: video learning analytics case study of MOOCs. *J. Learn. Anal.* **5**(3), 182–197 (2018)
4. Baker, R.S.J., Mitrović, A., Mathews, M.: Detecting gaming the system in constraint-based tutors. In: De Bra, P., Kobsa, A., Chin, D. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 267–278. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13470-8_25
5. Bixler, R., D’Mello, S.: Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Model. User-Adap. Inter.* **26**(1), 33–68 (2016)
6. Brooks, C., Thompson, C., Teasley, S.: A time series interaction analysis method for building predictive models of learners using log data. In: Proceedings of 5th Conference on Learning Analytics and Knowledge, pp. 126–135. ACM (2015)

7. Chatti, M.A., et al.: Video annotation and analytics in coursemapper. *Smart Learn. Environ.* **3**(1), 10 (2016)
8. Chi, M.T., Wylie, R.: The ICAP framework: linking cognitive engagement to active learning outcomes. *Educ. Psychol.* **49**(4), 219–243 (2014)
9. Cocea, M., Weibelzahl, S.: Disengagement detection in online learning: validation studies and perspectives. *IEEE Trans. Learn. Technol.* **4**(2), 114–124 (2011)
10. Daems, O., Erkens, M., Malzahn, N., Hoppe, H.U.: Using content analysis and domain ontologies to check learners' understanding of science concepts. *J. Comput. Educ.* **1**(2), 113–131 (2014)
11. Dimitrova, V., Mitrovic, A., Piotrkowicz, A., Lau, L., Weerasinghe, A.: Using learning analytics to devise interactive personalised nudges for active video watching. In: *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, pp. 22–31. ACM (2017)
12. Drummond, J., Litman, D.: In the zone: towards detecting student zoning out using supervised machine learning. In: Alevan, V., Kay, J., Mostow, J. (eds.) *ITS 2010. LNCS*, vol. 6095, pp. 306–308. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13437-1_53
13. Giannakos, M.N., Chorianopoulos, K., Chrisochoides, N.: Making sense of video analytics: lessons learned from clickstream interactions, attitudes, and learning outcome in a video-assisted course. *Int. Rev. Res. Open Distrib. Learn.* **16**(1), 260–283 (2015)
14. Hecking, T., Chounta, I.A., Hoppe, H.U.: Role modelling in MOOC discussion forums. *J. Learn. Anal.* **4**(1), 85–116 (2017)
15. Hecking, T., Dimitrova, V., Mitrovic, A., Ulrich Hoppe, U.: Using network-text analysis to characterise learner engagement in active video watching. In: *ICCE 2017 Main Conference Proceedings*, pp. 326–335. Asia-Pacific Society for Computers in Education (2017)
16. Hong, J.K., Mitrovic, A., Neshatian, K.: Predicting quitting behaviour in SQL-tutor. In: *Proceedings of the 23th International Conference on Computers in Education*, pp. 37–45. APSCE (2015)
17. Kaltura Inc.: *The state of video in education* (2017)
18. Koedinger, K.R., Kim, J., Jia, J.Z., McLaughlin, E.A., Bier, N.L.: Learning is not a spectator sport: doing is better than watching for learning from a MOOC. In: *Proceedings of 2nd ACM Conference on Learning@Scale*, pp. 111–120 (2015)
19. Kovacs, G.: Effects of in-video quizzes on MOOC lecture viewing. In: *Proceedings of the 3rd ACM Conference on Learning@ Scale*, pp. 31–40. ACM (2016)
20. Mills, C., Bosch, N., Graesser, A., D'Mello, S.: To quit or not to quit: predicting future behavioral disengagement from reading patterns. In: Trausan-Matu, S., Boyer, K.E., Crosby, M., Panourgia, K. (eds.) *ITS 2014. LNCS*, vol. 8474, pp. 19–28. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07221-0_3
21. Mitrovic, A., Dimitrova, V., Lau, L., Weerasinghe, A., Mathews, M.: Supporting constructive video-based learning: requirements elicitation from exploratory studies. In: André, E., Baker, R., Hu, X., Rodrigo, M.M.T., du Boulay, B. (eds.) *AIED 2017. LNCS (LNAI)*, vol. 10331, pp. 224–237. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61425-0_19
22. Mitrovic, A., Gordon, M., Piotrkowicz, A., Dimitrova, V.: Investigating the effect of adding nudges to increase engagement in active video watching. In: Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R. (eds.) *Artificial Intelligence in Education. LNCS*, pp. 320–332. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23204-7_27

23. Pintrich, P.R., et al.: A manual for the use of the motivated strategies for learning questionnaire (MSLQ). (1991)
24. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987)
25. Steinberg, D., Colla, P.: CART: classification and regression trees. *Top Ten Algorithms Data Min.* **9**, 179 (2009)
26. Vail, A.K., Wiggins, J.B., Grafsgaard, J.F., Boyer, K.E., Wiebe, E.N., Lester, J.C.: The affective impact of tutor questions: predicting frustration and engagement. In: *International Educational Data Mining Society* (2016)
27. Wang, X., Wen, M., Rosé, C.P.: Towards triggering higher-order thinking behaviors in MOOCs. In: *Proceedings of the 6th International Conference on Learning Analytics & Knowledge*, pp. 398–407. ACM (2016)