Proceedings Paper:

Piotrkowicz, A orcid.org/0000-0002-7723-699X, Dimitrova, V orcid.org/0000-0002-7001-0891, Mitrovic, A et al. (1 more author) (2018) Self-Regulation, Knowledge, Experience: Which User Characteristics Are Useful for Predicting Video Engagement? In: Proceedings UMAP '18 Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization. UMAP 2018, 08 Jul 2018, Nanyang Technological University, Singapore. Association for Computing Machinery , pp. 63-68. ISBN 978-1-4503-5784-5

https://doi.org/10.1145/3213586.3226196

Self-Regulation, Knowledge, Experience: Which User Characteristics Are Useful for Predicting Video Engagement?

Alicja Piotrkowicz
University of Leeds, UK
A.Piotrkowicz@leeds.ac.uk

Vania Dimitrova
University of Leeds, UK
V.G.Dimitrova@leeds.ac.uk

Antonija Mitrovic
University of Canterbury, NZ
Tanja.Mitrovic@canterbury.ac.nz

Lydia Lau
University of Leeds, UK
L.M.S.Lau@leeds.ac.uk

ABSTRACT

The use of videos in education has attracted considerable research attention. However, in order to gain the most benefits, learners need to actively engage with videos. It is an important, yet challenging, task to disentangle the relation between engagement with videos and learning, and at the same time to take into account relevant individual differences in order to offer personalised support. In this paper we investigate the question: 'Can user characteristics relating to self-regulation, knowledge, and experience be leveraged for predicting user engagement with videos?'. Our results show that users' domain knowledge and self-regulation abilities can inform overall engagement prediction (inactive, passive and constructive learners), which makes them useful for adaptation and personalisation.

CCS CONCEPTS

• **Applied computing → Interactive learning environments**;
• **Information systems →** *Personalization*;

ACM Reference Format:
Alicja Piotrkowicz, Vania Dimitrova, Antonija Mitrovic, and Lydia Lau. 2018. Self-Regulation, Knowledge, Experience: Which User Characteristics Are Useful for Predicting Video Engagement?. In *UMAP'18 Adjunct: 26th Conference on User Modeling, Adaptation and Personalization Adjunct, July 8–11, 2018, Singapore, Singapore.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3213586.3226196

1 INTRODUCTION

There is a growing area of adaptive learning systems that model learner engagement [5, 11, 18, 26, 35]. Automatic engagement detection can inform personalised interventions, e.g. motivational messages, questions, reminders, to prevent disengagement and enhance learning. Engagement modelling can also be used to evaluate educational systems, as it can point at requirements for improving educational content to keep students active.

We address the engagement detection challenge in a new context: *using videos for informal learning of soft skills*. Videos enable independent self-regulated learning where students familiarise themselves with or revisit key concepts in their own time [30]. A recent survey of a broad range of institutions reports that 99% have teachers who regularly incorporate video into their curricula [19]. Video-based learning is used in a wide spectrum of instructional settings [14, 20–23, 34]. At the same time, social video-sharing platforms are becoming the main means of content production and consumption for millennials. YouTube, which reaches more 18-34-year-olds in the US than any cable network[1], is becoming the first source when they want to learn something. However, watching videos is inherently a passive form of learning [8], often resulting in a low level of engagement [2, 21].

We have developed an active video watching platform (AVW-Space) that allows video use for informal learning [24, 28]. This taps into students' experiences with social video-sharing platforms and integrates interactive note-taking during video watching. The focus of this paper is on the early prediction of user engagement in AVW-Space by making use of the certain user characteristics collected before the user starts interacting with the system. We address the following research question: *Can we predict overall video engagement using data about the user's self-regulation, knowledge, and experience?* These human factors capture the individual differences in the user's learning: their self-regulation abilities, their domain knowledge, and their experience within the domain and with the technology. Leveraging these user characteristics for video engagement prediction would allow to make tailored interventions before the learner actually enters the system, thus potentially preventing disengagement.

The paper is organised as follows: Section 2 presents Related Work, while Section 3 gives an overview of the AVW project. In Section 4 we describe the datasets and in Section 5 the user profile features. We then present the method and results for predicting overall engagement from the user profile (Section 6). We finish with a discussion and conclusions in Section 7.

2 RELATED WORK

Engagement detection spans several research streams. There is an established research stream on predicting behaviour that can have adverse effect on learning, such as quitting in systems that embed free learning tasks (e.g. reading [26] and solving problems [18]), disengagement in MOOCs [2, 7, 21, 35], and 'gaming the system' (i.e. taking advantage of system's properties to superficially complete the task, rather than by attempting and thinking about the task) [3]. Another stream of work looks at detecting engagement aspects that can be linked to cognition, such as zoning out [11], mind wandering [5], and information seeking/giving [16, 33]. Affective response to instruction, e.g. frustration [32] and confusion [1], was also studied.

Our work contributes to research in user modelling for video engagement, e.g. [1, 21]. We present a novel prediction framework for

[1] https://www.youtube.com/intl/en-GB/yt/about/press/

active video watching with the following distinctive features. To inform the overall video engagement classes we *adopt the ICAP framework* [8] that links cognitive engagement activities to behaviours. While ICAP has been used to categorise information seeking in MOOC forums [16, 33], its adoption here for video engagement prediction is unique.

3 THE AVW PROJECT

AVW-Space supports active learning from videos via note-taking, which fulfills two major functions: to record information and to aid reflection [6]. It serves as an external cognitive aid for increasing learning [25]. We chose presentation skills as the learning domain and carefully selected eight videos from YouTube, e.g. each video is not longer than 10 minutes and a balanced coverage of male and female speakers [28]. Four videos were tutorials (T1-4) on how to give presentations, and the other four were examples of presentations (E1-4). To encourage reflection, we defined four aspects for tutorials (*I didn't realize I wasn't doing it*), self-awareness (*I am rather good at this*), relate to past experience (*I did/saw this in the past*), and note useful points (*I like this point*). For example presentations, we defined four aspects *Delivery*, *Speech*, *Structure*, and *Visual aids*, which allow the student to critique the presentations.

Initially students watch and comment on videos individually in the Personal Space (Figure 1), using aspects to tag their comments. AVW-Space records data about all actions the student performs on the platform, including time-stamped comments. The student can watch videos multiple times, including rewinding or skipping parts. Once the teacher approves comments for sharing, anonymised comments are available in the Social Space, in which students can rate comments. The teacher defines options for rating to promote deeper reflections. In addition to reading/rating the comments, the students can watch the part of the video associated with a comment.

We conducted several studies with the AVW-Space. All the studies, including the two presented in this paper, used the same materials (videos, aspects and rating categories) and were based on the same method. Different types of cohorts were recruited as users.

Method. In addition to the interaction data logged by AVW-Space, we designed surveys to collect data for profiling participants, their knowledge of presentation skills, and opinions on AVW-Space. Survey 1 was administered online at the beginning of the study, and contained question about demographics, background experiences, and the Motivated Strategies for Learning Questionnaire (MSLQ) [29]. Then, participants had one minute per question to write phrases they associated with (i) structure, (ii) delivery and speech, and (iii) visual aids. After interactions with AVW-Space, participants filled in Survey 2, which included the questions about knowledge of presentations, the NASA-TLX instrument [15] to check participants' perception of cognitive load during interaction, and the Technology Acceptance Model (TAM) [9] to check participants' perceived usefulness and usability of AVW-Space.

AVW engagement categories. The first study was conducted in March 2016 with postgraduate students [28]. Analyzing the collected data, we identified three categories of behaviours, based on the ICAP framework [8]. ICAP classifies overt learner behaviours into four types of learning modes, which correspond to different levels of cognitive engagement: Interactive, Constructive, Active and Passive. Due to the specifics of AVW-Space, we categorise users into three categories. **Inactive** users are those who filled out Survey 1, but did not log into AVW-Space. **Passive** users completed Survey 1 and watched some videos, but did not make comments. Finally, **Constructive** users are those who filled out Survey 1, watched videos, and made some comments.

Domain vocabulary. In order to analyse the comments and students' answers on conceptual understanding of presentation skills, we developed a taxonomy of domain keywords. The taxonomy was derived using a semi-automatic ontology engineering process, described in [17]. The taxonomy contains 645 domain keywords, organised into three main categories: *Structure*, *Delivery and Speech*, and *Visual Aids*.

Previous findings. Overall, the previous studies show that only constructive students improved their conceptual understanding of presentations skills [10]. One AVW study compared an experimental condition with reflection mini-scaffolds (i.e. aspects) to the control condition that followed free YouTube-like video watching [27]. We found a significant increase in conceptual knowledge in the experimental group where participants used aspects and exhibited constructive learning behaviour (i.e. made comments during video watching and rated comments written by others). There was no significant increase in conceptual knowledge in the control group.

4 DATASETS

In 2017 we conducted two studies (referred to as Study A and Study B) at the University of Canterbury, based on the method explained above. In both courses the students had to give a short presentation, which was marked by human tutors, but with no formal training on presentation skills. AVW-Space was provided as an online training resource on a voluntary basis. **Study A** was conducted in a mandatory course for all first-year Engineering students. The participation in the study was worth 1% of the final grade. Of the 904 students enrolled in the course, 463 completed Survey 1 (333 male, 128 female, 2 other). There were 150 constructive, 153 passive, and 160 inactive students. The majority of participants were aged 18 to 23 (96.54%), and were native English speakers (83.15%). The constructive students made 1,129 comments, with the average of 7.53 comments per student (sd = 9.37, range = [1, 75]). **Study B** was conducted with Business students in their second semester of study. The participation in the study was worth 2% of the final grade. Of 400 students enrolled in the course, 204 completed Survey 1 (104 male, 100 female). The majority of participants were aged 18 to 23 (89.7%), and were native English speakers (74.51%). There were 62 constructive, 26 inactive, and 116 passive students. The constructive students made 713 comments, with the average of 11.5 comments (sd = 11.81, range = [1, 56]). In the following section, we describe the explicit user profile features from Survey 1.

5 USER CHARACTERISTICS

We investigate three types of human factors using the explicit user profile: self-regulation, knowledge, and experience. These factors allow the investigation of individual differences in the user's learning: their ability to self-regulate their learning, their knowledge of the domain (making presentations), and their experience in making presentations and in using video platforms. The self-regulation,

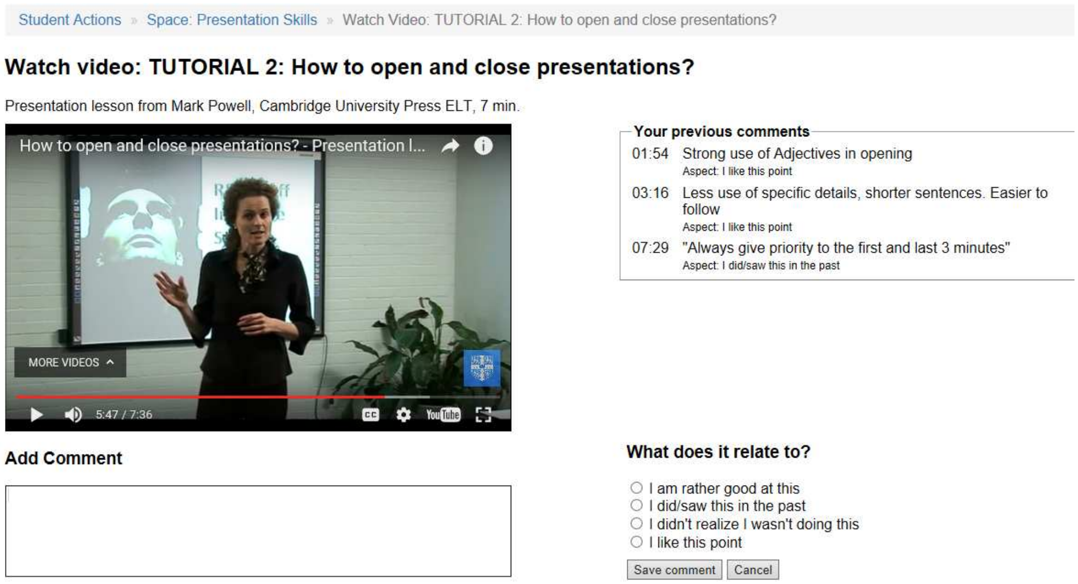

Figure 1: Adding comments in Private Space of AVW-Space

Table 1: Feature values comparison between the two studies (features with significant differences at $p < 0.01$ are indicated in bold; calculated using the Mann-Whitney test). In the range column, n indicates any number.

Feature	Range	Study A (463)	Study B (204)
Self-Regulation			
MSLQ-I	[1-5]	3.68	3.61
MSLQ-E	[1-5]	4.07	4.01
MSLQ-TV	[1-5]	3.89	3.87
MSLQ-C	[1-5]	4.11	4.17
MSLQ-SE	[1-5]	3.59	3.63
MSLQ-MSR	[1-5]	3.22	3.28
MSLQ-R	[1-5]	**3.08**	**3.51**
MSLQ-El	[1-5]	3.59	3.64
MSLQ-O	[1-5]	**3.11**	**3.36**
MSLQ-ER	[1-5]	3.45	3.44
Knowledge			
Conceptual Knowledge	[0-n]	12.49	11.79
Experience			
Training	[1-5]	1.67	1.74
Experience	[1-5]	2.19	2.29
YouTube	[1-5]	4.11	4.01
YouTube for Learning	[1-5]	**3.15**	**2.66**

knowledge, and experience factors are captured in 15 features obtained from Survey 1.

The user's self-regulation is captured using ten features which are the aggregated scores of the MSLQ questions [29]:

MSLQ-I intrinsic motivation, representing the degree to which the student participates in academic tasks for reasons linked to challenge, curiosity and mastery;

MSLQ-E extrinsic motivation, the degree to which the student participates in academic tasks for reasons such as grades and rewards;

MSLQ-TV Task Value, which refers to the student's perceptions of academic studies in terms of interest, importance and utility;

MSLQ-C Control, indicating whether the learner feels in control of his/her own performance;

MSLQ-SE Self Efficacy, the student's confidence in having skills to perform academic tasks;

MSLQ-MSR Metacognitive Self-Regulation;

MSLQ-R Rehearsal

MSLQ-El Elaboration

MSLQ-O Organisation

MSLQ-ER Effort Regulation

The user's knowledge about the domain is captured using **Conceptual Knowledge (CK)**. The CK score is calculated from the user's textual answers to conceptual knowledge questions, which were annotated with the domain vocabulary (cf. Section 3). The total number of domain concepts the student has named in Survey 1 is used as a proxy for the student's conceptual knowledge.

Finally, the last four features capture the user's previous experiences relevant to the domain and video watching. They were collected using a 5-point Likert scale.

Training on giving presentations;

Experience in giving presentations;

YouTube frequency of using YouTube;

YouTube for Learning the extent to which the user has used YouTube for learning.

Table 2: Significant differences between the three categories (In. - Inactive; Pass. - Passive and Con. - Constructive) (calculated using Kruskal-Wallis test).

	In.	Pass.	Con.	p	Pairwise diff.
Study A					
# users	(160)	(153)	(150)		
CK	11.18	12.14	14.25	**	In-Con; Pass-Con
MSLQ-E	3.99	4.03	4.19	*	In-Con
MSLQ-SE	3.48	3.61	3.68	*	In-Con
MSLQ-MSR	3.14	3.25	3.29	*	In-Con
MSLQ-ER	3.25	3.52	3.58	***	In-Con; In-Pass
Study B					
# users	(26)	(116)	(62)		
Experience	2.65	2.16	2.4	*	In-Pass
CK	9.35	11.81	12.79	*	In-Con
MSLQ-C	3.87	4.16	4.32	**	In-Con

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

In Table 1 we compare the profile features for the two studies we conducted. Only three features had statistically significant differences between Studies A and B: YouTube for Learning is higher for Study A, both MSLQ-Rehearsal and MSLQ-Organisation are higher for Study B.

6 PREDICTING OVERALL ENGAGEMENT

We now turn to investigating the combined predictive power of the profile features to predict engagement. This prediction can be carried out before the user even starts interacting with the system, which would allow early personalised interventions to influence the user's behaviour.

6.1 Model building

Task. Our task is to predict the overall engagement using just the explicit user profile. We consider three task settings:

(1) Simply predicting whether the student would be inactive, passive, or constructive (three-class classification).
(2) Using binary classification, whereby for each category we build a binary classifier of the format One-class-vs-Others (e.g. Inactive vs. Passive + Constructive), which allows zooming into specific categories. Table 2 shows the significant differences between the three categories (= prediction classes).
(3) Repeating the binary classification with only the most predictive features selected through recursive feature elimination. This way we could limit the amount of data we need to gather from the users (i.e. collect only the data for the most predictive features).

Data pre-processing. We use all the user characteristics described in Section 5 as predictors, and the categorisation relevant for each task setting as the target variable. The features are first preprocessed by removing near-zero variance predictors, and scaling the remaining ones between 0 and 1. This normalises the data, which helps in case of some classifiers (e.g. Support Vector Machines). Since in nearly all task settings (with the exception of three-class prediction for Study A) there was a class imbalance, we used upsampling (i.e. repeating the instances from the minority class) to balance out the class distribution and prevent the classifier from always predicting the majority class.

Table 3: Classification results using user profile data. Metrics include: accuracy (Acc.), precision (Prec.), recall (Rec.). RFE selected features are presented in Table 4.

	Study A				**Study B**		
Three class classification							
Acc.	0.41			Acc.	0.63		
	In.	Pass.	Con.		In.	Pass.	Con.
Prec.	0.5	0.55	0.53	Prec.	0.88	0.87	0.81
Rec.	0.71	0.29	0.55	Rec.	0.98	0.69	0.88
Binary classification (one-class vs. others)							
	In.	Pass.	Con.		In.	Pass.	Con.
Acc.	0.66	0.62	0.62	Acc.	0.86	0.46	0.7
Prec.	0.77	0.77	0.77	Prec.	0.93	0.71	0.85
Rec.	0.79	0.82	0.79	Rec.	0.93	0.66	0.82
Binary classification with RFE selected features							
	In.	Pass.	Con.		In.	Pass.	Con.
Acc.	0.61	0.64	0.65	Acc.	0.73	0.6	0.65
Prec.	0.66	0.7	0.75	Prec.	0.86	0.73	0.74
Rec.	0.74	0.79	0.62	Rec.	0.78	0.74	0.76

Training. The classifiers are trained separately for each study. We use a Leave-One-Out cross-validation (LOOCV), whereby the data for all but one user is used for training, and the performance for each iteration is evaluated on that one user. The process is repeated for each user (i.e. the number of times equal to the number of users in the dataset). Cross-validation allows to utilise all of the data for training, while also keeping training and test data separate in each iteration. We experimented with a number of classifiers (including generalised linear models, decision trees, and support vector machines) and found support vector machines (SVM) with radial basis function (RBF) kernel to yield best results. Random forests produced similarly high results, however took significantly longer to train (especially when using LOOCV).

Recursive feature elimination. For each task setting we also ran recursive feature elimination (RFE) with resampling in order to identify the best predictors. RFE ranks all predictors according to their importance to the model, then all subset sizes of ranked predictors (from N to 1, where N is the number of predictors) are used to train the classifier which is evaluated according to a standard metric (accuracy in case of classification). The RFE function returns the best subset size, i.e. the one that yields the best accuracy score. In order to reduce overfitting in the RFE model, this process is repeated in a 10-fold cross-validation.

Evaluation. To assess the model, we use the accuracy averaged over the iterations of the Leave-One-Out cross-validation. We also report precision and recall, which we calculate using the predictions from the final model fitted through cross-validation.

Table 4: Profile features selected through recursive feature elimination for each of the classifiers. A/B refers to the study.

	Three classes		In.		Pass.		Con.	
	A	B	A	B	A	B	A	B
CK	+	+	+	+	+	+	+	+
MSLQ-I				+		+		
MSLQ-E	+					+		+
MSLQ-SE	+	+	+	+	+		+	
MSLQ-R			+		+		+	
MSLQ-MSR	+	+	+		+	+	+	+
MSLQ-ER	+	+	+	+	+		+	+
MSLQ-C		+		+				+

6.2 Results

We report the classification results in Table 3. In the three-class prediction setting, the results were noticeably higher for Study B (Acc. = 0.63), compared to Study A (Acc. = 0.41). That was also the case when looking at the precision and recall for each of the categories. Although better than a random choice (which would have an accuracy of approx. 0.33), the results of the three-category prediction for Study A are not reliable enough to predict user engagement.

The results for the binary classifiers are overall higher. The only binary classifier that suffers a drop in performance compared to the three-class prediction is the Passive classifier for Study B. In particular, the results for Study A show an improvement. Overall, the fairly high (at least 0.7) precision and recall values mean that the prediction model can be reliably used to predict user engagement.

We want to explore whether we can reduce the number of survey questions we ask users to complete. The binary classification results which use RFE selected features indicate whether a feature set reduced to only the top predictors achieves comparable performance to the full feature model. The sets of RFE selected features for each classifier are reported in Table 4. The top selected feature across all classifiers is Conceptual Knowledge (CK), followed by Meta-Self-Regulation (MSLQ-MSR; seven classifiers) and Effort Regulation (MSLQ-ER, seven classifiers). With the exception of Task Value and Elaboration, all MSLQ features were selected by at least two classifiers, which indicates that this data (as well the conceptual knowledge) should be collected from the users.

Overall, looking at the binary classification with RFE selected features the performance drops with the reduced number of features. However, the results for the Passive classification for Study B are improved by using only the top predictors. This brings them to a similar level to the binary classification with the full feature set, which gives us good results for all classifiers.

7 DISCUSSION

Key findings. We found that by using user profile relating to their self-regulation, knowledge, and experience, we can predict whether a given user will display Inactive, Passive, or Constructive engagement behaviour with a fair degree of accuracy. The user profile features which are the best predictors across the different prediction models were conceptual knowledge and MSLQ features. Our work gives supporting evidence that self-regulation abilities are important user characteristics to consider in adaptive video learning. Hence, collecting information on human factors such as self-regulation through surveys like MSLQ offer valuable insights for personalisation and adaptation.

Explicit vs. implicit user profile. There are advantages to using both explicit and implicit user profiles. We have shown that knowledge and self-regulation features from the user profile can be reasonably used as predictors for video engagement. This allows the adaptation before the user begins interacting with the system. However, in order to derive the explicit profile, users are asked to complete lengthy questionnaires. An investigation such as ours helps to identify which aspects of the explicit profile actually contribute to engagement profiling and interventions, which may limit the number of questions presented to the user. Moreover, it has been suggested that psychological questionnaires might suffer from a bias due to the self-assessment [12]. On the other hand, implicit profiling does not place any burden on the user, and derives the relevant information from user interactions. However, implicit profiling (and the consequent adaptation) can only happen after some period of interaction, which could lead to some users already disengaging. Secondly, deriving complex user characteristics, such as self-regulation, from user interaction is an open research task. This research strand has mostly focused on one aspect of self-regulation – reflection – through the use of reflective writing analytics [13, 31]. There is also some initial work at modelling learners' self-regulation using process mining [4]. In order to leverage the benefits of both explicit and implicit profiling, a possible solution is an adaptation model which initially rely on the explicit profile, and then shift to using implicit profiling once enough data has been collected. Further studies are needed to investigate the predictive power of explicit and implicit profile features and the interplay between them.

Applications of the prediction model. The model we developed using the explicit user profile can be applied in several personalisation and adaptation contexts. This includes *planning interventions* – implementing nudges is one of the primary goals for the AVW project [10]. The presented prediction model can help identify *whom* to target before the users start watching videos, and thus start targeted interventions that would aim to prevent disengagement.

8 CONCLUSIONS

Our goal was to leverage self-regulation, knowledge, and experience from the user profile to predict user engagement with the view of implementing early interventions to support beneficial video engagement. We found that the three types of human factors we considered can be used for early prediction of user engagement with a reasonable degree of accuracy. In particular, knowledge and self-regulation feature were found to be strong predictors. Engagement modelling approaches, as the one presented here, will enable adapting the video learning experience to individual differences in a broad range of contexts. We plan to explore active video watching for other soft skill training contexts which address pressing needs of the modern society, e.g. unconscious bias and cultural awareness.

Acknowledgements. Support by EU-FP7-ICT-257184 ImREAL, Ako Aotearoa and teaching development grants from the University of Canterbury. We thank Amali Weerasinghe for her work on the experiment design, as well as the study participants.

REFERENCES

[1] Akshay Agrawal, Jagadish Venkatraman, Shane Leonard, and Andreas Paepcke. 2015. YouEDU: Addressing Confusion in MOOC Discussion Forums by Recommending Instructional Video Clips. (2015), 297–304.

[2] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2014. Engaging with Massive Online Courses. In *Proc. 23rd Int. Conf. World wide web*. ACM, 687–698.

[3] Ryan S. J. d. Baker, Antonija Mitrovic, and Moffat Mathews. 2010. Detecting Gaming the System in Constraint-Based Tutors. In *User Modeling, Adaptation, and Personalization*, Paul De Bra, Alfred Kobsa, and David Chin (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 267–278.

[4] Maria Bannert, Peter Reimann, and Christoph Sonnenberg. 2014. Process mining techniques for analysing patterns and strategies in studentsâĂŹ self-regulated learning. *Metacognition and learning* 9, 2 (2014), 161–185.

[5] Robert Bixler and Sidney D'Mello. 2016. Automatic Gaze-based User-independent Detection of Mind Wandering During Computerized Reading. *User Modeling and User-Adapted Interaction* 26, 1 (March 2016), 33–68. https://doi.org/10.1007/s11257-015-9167-1

[6] Francoise Boch and Annie Piolat. 2005. Note Taking and Learning: A Summary of Research. *The WAC Journal* 16 (2005), 101–113.

[7] Christopher Brooks, Craig Thompson, and Stephanie Teasley. 2015. A Time Series Interaction Analysis Method for Building Predictive Models of Learners Using Log Data. In *Proc. 5th Int. Conf. Learning Analytics And Knowledge (LAK '15)*. ACM, New York, NY, USA, 126–135. https://doi.org/10.1145/2723576.2723581

[8] Michelene TH Chi and Ruth Wylie. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist* 49, 4 (2014), 219–243.

[9] Fred D Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS quarterly* (1989), 319–340.

[10] Vania Dimitrova, Antonija Mitrovic, Alicja Piotrkowicz, Lydia Lau, and Amali Weerasinghe. 2017. Using Learning Analytics to Devise Interactive Personalised Nudges for Active Video Watching. In *Proc.25th Conf. User Modeling, Adaptation and Personalization*. ACM.

[11] Joanna Drummond and Diane Litman. 2010. In the Zone: Towards Detecting Student Zoning Out Using Supervised Machine Learning. In *Int. Conf. Intelligent Tutoring Systems*. Springer, 306–308.

[12] David Dunning, Chip Heath, and Jerry M Suls. 2004. Flawed self-assessment: Implications for health, education, and the workplace. *Psychological science in the public interest* 5, 3 (2004), 69–106.

[13] Andrew Gibson, Adam Aitken, Ágnes Sándor, Simon Buckingham Shum, Cherie Tsingos-Lucas, and Simon Knight. 2017. Reflective writing analytics for actionable feedback. In *Proc. 7th Int. Learning Analytics & Knowledge Conference*.

[14] Philip J Guo, Juho Kim, and Rob Rubin. 2014. How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. In *Proc. 1st ACM conference on Learning@ scale conference*. ACM, 41–50.

[15] Sandra G Hart. 2006. NASA-Task Load Index (NASA-TLX); 20 Years Later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage Publications Sage CA: Los Angeles, CA, 904–908.

[16] Chounta Irene-Angelica Hecking, Tobias and Ulrich Hoppe. [n. d.]. Role Modelling in MOOC Discussion Forums. *Learning Analytics* 4, 1 ([n. d.]), 85–116.

[17] Tobias Hecking, Vania Dimitrova, Antonija Mitrovic, and H Ulrich Hoppe. 2017. Using Network-Text Analysis to Characterise Learner Engagement in Active Video Watching. In *Proc. ICCE 2017*. Asia-Pacific Society for Computers in Education, 326–335.

[18] Jin Kwang Hong, Antonija Mitrovic, and Kourosh Neshatian. 2015. Predicting Quitting Behaviour in SQL-Tutor. In *ICCE 2015 Proceedings*. Asia-Pacific Society for Computers in Education, 37–45.

[19] Kaltura Inc. 2017. *The State of Video in Education 2017*. Technical Report. Kaltura Inc. https://corp.kaltura.com/sites/default/files/The_State_Of_Video_in_Education_2017.pdf

[20] Sethela June, Aizan Yaacob, and Yeoh Khar Kheng. 2014. Assessing the Use of YouTube Videos and Interactive Activities as a Critical Thinking Stimulator for Tertiary Students: An Action Research. *International Education Studies* 7, 8 (2014), 56–67.

[21] Kenneth R Koedinger, Jihee Kim, Julianna Zhuxin Jia, Elizabeth A McLaughlin, and Norman L Bier. 2015. Learning is not a Spectator Sport: Doing is Better than Watching for Learning from a MOOC. In *Proc. 2nd ACM Conf. Learning at scale*. ACM, 111–120.

[22] Miloš Kravčík, Petru Nicolaescu, Aarij Siddiqui, and Ralf Klamma. 2017. Adaptive Video Techniques for Informal Learning Support in Workplace Environments. In *Emerging Technologies for Education*, Ting-Ting Wu, Rosella Gennari, Yueh-Min Huang, Haoran Xie, and Yiwei Cao (Eds.). Springer International Publishing, Cham, 533–543.

[23] Gila Kurtz, Alexandr Tsimerman, and Orna Steiner-Lavi. 2014. The Flipped-Classroom Approach: The Answer to Future Learning? *European J. Open, Distance and E-learning* 17, 2 (2014), 172–182.

[24] Lydia Lau, Antonija Mitrovic, Amali Weerasinghe, and Vania Dimitrova. 2016. Usability of an Active Video Watching System for Soft Skills Training. In *Proc. 1st Int. Workshop on Intelligent Mentoring Systems (ITS)*.

[25] Tamas Makany, Jonathan Kemp, and Itiel E Dror. 2009. Optimising the Use of Note-taking as an External Cognitive Aid for Increasing Learning. *British Journal of Educational Technology* 40, 4 (2009), 619–635.

[26] Caitlin Mills, Nigel Bosch, Art Graesser, and Sidney D'Mello. 2014. To Quit or not to Quit: Predicting Future Behavioral Disengagement from Reading Patterns. In *Int. Conf. Intelligent Tutoring Systems*. Springer, 19–28.

[27] Antonija Mitrovic, Vania Dimitrova, Lydia Lau, Amali Weerasinghe, and Moffat Mathews. 2017. Supporting Constructive Video-based Learning: Requirements Elicitation from Exploratory Studies. In *18th Int. Conf. Artificial Intelligence in Education*. 224–237.

[28] Antonija Mitrovic, Vania Dimitrova, Amali Weerasinghe, and Lydia Lau. [n. d.]. Reflective Experiential Learning: Using Active Video Watching for Soft Skills Training, In ICCE 2016. *Proc. 24th Int. Conf. Computers in Education*, 192–201.

[29] Paul R Pintrich et al. 1991. A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ). (1991).

[30] Daniel L. Schacter and Karl K. Szpunar. [n. d.]. Enhancing Attention and Memory During Video-recorded Lectures. *Scholarship of Teaching and Learning in Psychology* 1, 1 ([n. d.]), 60–71.

[31] Thomas Daniel Ullmann. 2017. Reflective Writing Analytics-Empirically Determined Keywords of Written Reflection. In *Proc. 7th Int. Learning Analytics & Knowledge Conference*.

[32] Alexandria Katarina Vail, Joseph B. Wiggins, Joseph F. Grafsgaard, Kristy Elizabeth Boyer, Eric N. Wiebe, and James C. Lester. 2016. The Affective Impact of Tutor Questions: Predicting Frustration and Engagement. In *EDM*.

[33] Xu Wang, Miaomiao Wen, and Carolyn P. Rosé. 2016. Towards Triggering Higher-order Thinking Behaviors in MOOCs. In *Proc. 6th Int. Conf. Learning Analytics & Knowledge (LAK '16)*. ACM, New York, NY, USA, 398–407. https://doi.org/10.1145/2883851.2883964

[34] Ahmed Mohamed Fahmy Yousef, Mohamed Amine Chatti, and Ulrik Schroeder. 2014. The State of Video-based Learning: A Review and Future Perspectives. *Int. J. Adv. Life Sci* 6, 3/4 (2014), 122–135.

[35] Yue Zhao, Dan Davis, Guanliang Chen, Christoph Lofi, Claudia Hauff, and Geert-Jan Houben. 2017. Certificate Achievement Unlocked: How Does MOOC Learners' Behaviour Change?. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP '17)*. ACM, New York, NY, USA, 83–88. https://doi.org/10.1145/3099023.3099063