INVESTIGATING VIDEO VIEWING BEHAVIORS OF STUDENTS WITH DIFFERENT LEARNING APPROACHES USING VIDEO ANALYTICS

Dr. Gokhan AKCAPINAR
Department of Computer Education & Instructional Technology
Hacettepe University
Ankara, Turkey

Dr. Alper BAYAZIT
Department of Computer Education & Instructional Technology
Yeditepe University
Istanbul, Turkey

ABSTRACT

The deep and surface learning approaches are closely related to the students' interaction with learning content and learning outcomes. While students with a surface approach have a tendency to acquire knowledge without questioning and to try to pass courses with minimum effort, students with a deep learning approach tend to use more skills such as problem-solving, questioning, and research. Studies show that learning approaches of students can change depending on subject, task and time. Therefore, it is important to identify students with a surface learning approach in online learning environments and to plan interventions that encourage them to use deep learning approaches. In this study, video viewing behaviors of students with deep and surface learning approaches are analyzed using video analytics. Video viewing patterns of students with different learning approaches are also compared. For this purpose, students (N=31) are asked to study a 10-minutes-long video material related to Computer Hardware course. Video interactions in this process were also recorded using video player developed by the authors. At the end of the lab session, students were asked to fill in the Learning Approach Scale by taking into account their learning approaches to the course. As a result of the study, it was observed that the students with surface approach made a statistically significant forward seek over to the students used deep learning approach while watching the video. Moreover, an investigation on the time series graphs of two groups revealed that surface learners watched the video more linearly and had fewer interactions with it. These interaction data can be modeled with machine learning techniques to predict students with surface approach and can be used to identify design problems in video materials.

Keywords: Video analytics, educational data mining, learning approach, video learning learning analytics.

INTRODUCTION

The use of video-based learning materials in online learning environments is becoming increasingly across the globe. At the same time, studies showed that students spend more time on video materials than text materials (Seaton, Bergner, Chuang, Mitros, & Pritchard, 2014). One of the key advantages that video-based materials provide is that they contain interactive elements which appeal to both visual and auditory senses. Besides, it allows the students to progress in learning at their own pace and review the sections they want, which can be addressed as other notable advantages (Kim et al, 2014). Despite the benefits and widely use of educational video materials, there is a limited number of studies conducted on the basis of data to investigate the video viewing behavior of students. This

limited-scale research, however, shows that data-driven approaches can provide important information about video-based learning (Kim et al, 2014; PI, Hong, & Yang, 2017; Schiltz, 2015).

In this study, video viewing behaviors of students are discussed in terms of their learning approaches. The difference between the video viewing behaviors of the students with different learning approaches is also analyzed and the correlation between the video metrics and the learning approach scores of the students is examined. At the same time, the video viewing data of deep and surface learning approaches are visualized in a time series graph with an aim to visually compare the video viewing behaviors of the two groups.

Learning Approaches

Learning approaches were first introduced by Marton and Saljo (1976) in 1976. The researchers observed that, in reading-related tasks, some students focus on memorizing the texts to answer the questions while others study towards understanding the underlying meaning of the texts (Marton & Saljo, 1976a, 1976b). Based on this, they inferred that students use either surface or deep learning strategies while carrying out a learning task. The conducted studies point to the fact that surface learning is correlated with low-quality learning outcomes while deep learning is associated with high-quality learning outcomes (Rajaratnam, D'cruz, & Chandrasekhar, 2013). Therefore, in online learning environments, it is crucial that the students who follow the surface approach are identified in time and intervened to ensure that they adopt the deep learning strategies (Ak, 2008).

Video Analytics

Numerous studies are conducted on the collection and analysis of students' interaction data related to their learning processes. The general purpose of these researches, which falls under learning analytics, is to seek solutions to educational problems by analyzing the traces that left behind by students in their learning processes (Chatti, Dyckhoff, Schroeder, & Thüs, 2012). Video analytics can be regarded as a sub-field of learning analytics. The purpose here is to perform data-driven analyses regarding video-based learning by analyzing the students' clickstreams on those videos.

In literature, a limited number of studies can be found where the video viewing behaviors of students are analyzed. The results obtained from these studies, however, revealed valuable information about video-based learning and designing educational videos. For example, a study conducted on edX platform by Guo, Kim, and Rubin (2014) revealed that students' interactions have decreased significantly in videos that are longer than 6 minutes. Another study carried out by Kim et al (2014) compared the students' video viewing behaviors in lecture and tutorial videos. This study found that the students watched the course videos more linearly while seeking more frequently in the training videos. Moreover, they have found that there is a frequent replay, especially where an important issue or theory is explained, or where there are screen changes. Based on this result, researchers have indicated that in tutorial videos, the learners often need to replay, therefore putting markers at that points will help their learning.

Chen and Wu (2015) studied how three different video materials (voice-over presentation, picture-in-picture and lecture recording in class) changes in accordance with the cognitive differences and learning styles of the students. The researchers analyzed the students with visual and verbal styles in terms of sustained-attention, cognitive load, emotion and learning performance. The researchers reached the conclusion that all three video types enhanced the learning performance, but the picture-in-picture and lecture recording methods proved more effective compared to voice-over presentation. They concluded that, while visual and verbal learners performed similarly in all three video types, sustained attention and cognitive load values of the students were higher in the video that was prepared using the voice-over presentation method.

Guo et al. (2014) investigated the impact of different video types on students' interaction with the video by the help video analytics. In the study, data from large number of students

were analyzed, the researchers concluded that students showed greater interest in short videos. They further concluded that the talking-head videos produced in an informal manner generated more interactions than the high-quality professional training videos. Kleftodimos and Evangelidis (2016) grouped the video-viewing data of students using cluster analysis. By doing this, they aimed to determine different student profiles. The researchers asserted that those profiles could be used in identifying the students who experience problems with the videos that are viewed particularly for the purpose to carry out a specific task.

From the above-mentioned studies, it is clear that video analytics can be helped to obtain crucial information regarding the effectiveness and design of video courses. None of the studies has compared the video viewing behaviors of learners with deep and surface learning approaches. . Therefore, in this study, the video viewing behaviors of deep and surface learners are aimed to be compared to overcome the following research questions:

- **Is there a statistically significant difference between the video metrics of deep and surface learners (play, pause, seek, etc.)?**
- **Is there a correlation between deep and surface approach scores and video metrics?**
- **What similarities and differences exist between the time series graphs that visualize the video viewing behaviors of deep and surface learners?**

METHOD

The research was carried out in the Computer Education and Instructional Technology department at a state university in Turkey with undergraduate students (N = 31). The data were collected from Computer Hardware course offered by the department. As a video material, a 10-minute-long video about the Hard Disk Drive (HDD) topic was designed for the experiment. Students are asked to study during the course period on this video that they have not seen before. The interactions of the students while watching the video were recorded using a video player developed by the researchers. At the end of the session, students were asked to fill Biggs' The Revised Two-Factor Study Process Questionnaire (R-SPQ-2F). Questionnaire data and interaction data were joined based on students' ID in Moodle. Thus, in the process of data analysis, personal data anonymization technique was applied.

Study Process Questionnaire

The Two-Factor Study Process Questionnaire (R-SPQ-2F), which was developed by Biggs (J. B. Biggs, 1987a, 1987b) and revised in 2001 by J. Biggs, Kember, and Leung (2001), was used in order to evaluate course-oriented learning strategies of students. The Turkish adaptation of the scale's current version was made by Onder and Besoluk (2010). The final version of the scale contains 20 learning-related items, 10 of which were on surface learning and the remaining 10 were on deep learning. In turn, each factor contains in itself the motivation and strategy sub-dimensions, each consisting of five items. In other words, the final version of the scale consists of four sub-dimensions namely, Deep Motivation (DM), Deep Strategy (DS), Surface Motivation (SM), and Surface Strategy (SS); and two factors, Deep Approach (DA) and Surface Approach (SA) (J. Biggs et al., 2001). While the scale scores obtained for each sub-dimension can vary between 5 and 25, the scores that can be obtained regarding the deep and surface approaches vary between 10 and 50.

Video Interaction Data

Video interaction data collected by the video player developed by the authors (Bayazıt & Akcapinar, 2018). This video player was integrated into the Moodle Learning Management System (LMS). Thus it enables the students' video viewing data to be recorded by linking them to the user IDs on the Moodle platform. Each interaction made by students on video player has recorded in the database as a row. These records contain student ID, session ID, date-time, type of interaction (pause, play, seek, etc.) and descriptions of action. When the

session is done, hundreds of rows of the records are created for each student. A pre-processing tool has also been developed so that these data can be used for analysis. This tool removes duplicate and incorrect records of the student interactions and generates features for further analysis. These features and their descriptions are provided in Table 1.

Table 1. Features Obtained from Video Interactions and Their Descriptions

No	Feature	Description
1	Total Interaction	The number of total interactions on the video player
2	Duration	Total video viewing duration
3	Video Page Visit	Total number of video page visit
4	Video Playing	The number of times the play button on the video player is clicked
5	Video Pausing	The number of times the pause button on the video player is clicked
6	Video Completion	The number of times the video is completed
7	Video Seeking	The number of clicks on the time-line for seeking purposes
8	Forward Seeking	The total number of forward seeking
9	Backward Seeking	The total number of backward seeking

The 40-minutes session consisting of 31 students resulted 1793 records regarding the student interactions in the database.

Data Analysis

Study Process Questionnaire generates scores between 10 and 50 using the sum of each learning approach in accordance with its sub-scales. Different approaches are used in previous studies to tell whether the students are deep or surface learners based on the scale scores (Beheshitha, Gasevic, & Hatala, 2015; Hamm, 2009). This study aimed to group the students who obtained similar scores from the questionnaire by using cluster analysis. Cluster analysis was conducted by using the X-Means clustering algorithm. The process visualized in Figure 1 is an illustration from RapidMiner software. Other statistical analyses were conducted using SPSSS software. The time series graphs to visualize the students' video viewing data were generated using the R statistic software.

For cluster analysis, at first, the scores obtained by the students from the sub-dimensions of the scale were normalized by converting them into their z-scores. Next, the number of optimal clusters was determined by the X-Means algorithm. After that, whether there were meaningful differences between the obtained clusters in terms of scale scores was analyzed using independent samples t-test. Since the features regarding the video analytics do not distribute normally, the difference analyses within the scope of the first research problem were conducted using the Mann Whitney - U test. Finally, the correlations between the features within the scope of the second research problem were examined using the Spearman's Rho correlation analysis.

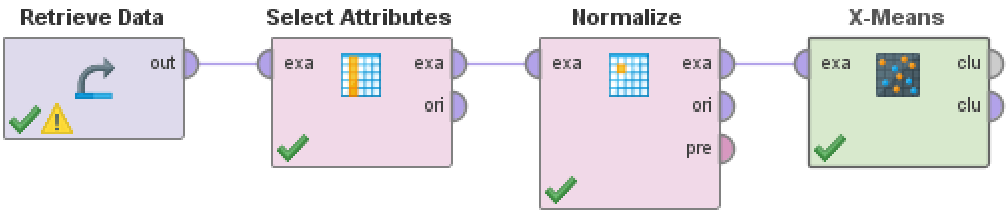

Figure 1. Cluster Analysis Process

FINDINGS

Cluster Analysis

Based on the cluster analysis result, the data were divided into two groups. Table 2 shows the analysis results of the independent samples t-test, which was conducted to test whether there is a difference between the two clusters in terms of scale scores. The cluster means were considered to name these clusters (see Table 2), and the students in Cluster 1 were named Deep Learner (n = 15) and the students in Cluster 0 were named Surface Learner (n = 16). Independent samples t-test analysis result revealed that there is a significant difference between the two groups in the scores of deep and surface learning approaches.

Table 2. Independent Samples T-Test Analysis Result

	Cluster_0 (n = 16)		Cluster_1 (n = 15)		t-test
	$\bar{X}$	SD	$\bar{X}$	SD	t
Deep Approach	28,81	4,66	36,67	4,10	-4,96*
Surface Approach	31,69	3,72	24,40	3,31	5,75*

Note. *p<.001

Difference Analyses

To analyze whether there are differences in deep and surface learners' video metrics, Mann Whitney-U test was conducted. The test results are shown in Table 3. When the test results are analyzed (see Table 3), except forward seek, there were no statistically significant differences observed between the two groups. According to this, it can be seen that the surface learners (Mean Rank = 19.06) watched the video by doing more forward seek than the deep learners (Mean Rank = 12.73) (U = 71, p = 0.048).

Table 3. Mann Whitney - U Test Results Related to the Video Metrics

Feature	Group	n	Mean Rank	U	Z	P
Total Interaction	Cluster_0	16	15,41	110	-0,376	0,707
	Cluster_1	15	16,63			
Duration	Cluster_0	16	15,22	107	-0,503	0,615
	Cluster_1	15	16,83			
Video Page Visit	Cluster_0	16	17,28	99	-1,111	0,267
	Cluster_1	15	14,63			
Video Playing	Cluster_0	16	14,06	89	-1,227	0,22
	Cluster_1	15	18,07			
Video Pausing	Cluster_0	16	14,22	91	-1,13	0,258
	Cluster_1	15	17,90			
Video Completion	Cluster_0	16	14,72	99	-0,879	0,379
	Cluster_1	15	17,37			
Video Seeking	Cluster_0	16	17,53	95	-0,971	0,332
	Cluster_1	15	14,37			
Forward Seeking	Cluster_0	16	19,06	71	-1,976	0,048
	Cluster_1	15	12,73			
Backward Seeking	Cluster_0	16	16,81	107	-0,516	0,606
	Cluster_1	15	15,13			

Correlation Analysis

Within the scope of the second research problem, Spearman's Rho correlation analysis is used to examine the correlation between students' scale scores and their video metrics. According to the results of the analysis, only the correlation between students' surface approach scores and the number of forward seek was significant ($r = 0.44$, $p < 0.05$). In other words, there appears to be a moderately positive correlation between the students' surface approach scores and the number of forward seeks.

Table 4. Results of Spearman's Rho Correlation Analysis between the Video Metrics and Scale Scores

	Deep Approach	Surface Approach
Total Interaction	0,095	-0,027
Duration	0,141	0,069
Video Page Visit	-0,068	-0,032
Video Playing	0,229	-0,173
Video Pausing	0,225	-0,154
Video Completion	0,245	-0,188
Video Seeking	-0,083	0,198
Forward Seeking	-0,255	0,444*
Backward Seeking	-0,009	0,142

Note. *p<.05

Time Series Graphs

To advent the third research problem, video viewing behaviors of deep and surface learners were visualized in a time series graph so that students' video viewing behaviors can be visually compared. Time series graphs are used in video analytics studies to visualize students' interactions in video viewing processes (Giannakos, Krogstie, & Aalberg, 2016; Kim et al., 2014). In general, these graphs make it possible to see in which parts of the video students interacted the less and the most (the parts where they play, pause, seek etc.), and to acquire information on their video viewing behaviors.

Figure 2 shows the time series graphs reflecting the video viewing behaviors of students with surface and deep learning approaches. The graph was drawn using the data of all students in the related cluster. In the graph, the y-axis represents the frequency values regarding the play, pause and seek events. On the other hand, the x-axis represents the timelines of the video. An investigation on the graph reveals that surface learners watched the video more linearly and had fewer interactions with it. On the contrary, while viewing the video, deep learners were seen to play-pause the video more frequently and focus on the certain areas of the video (peaks).

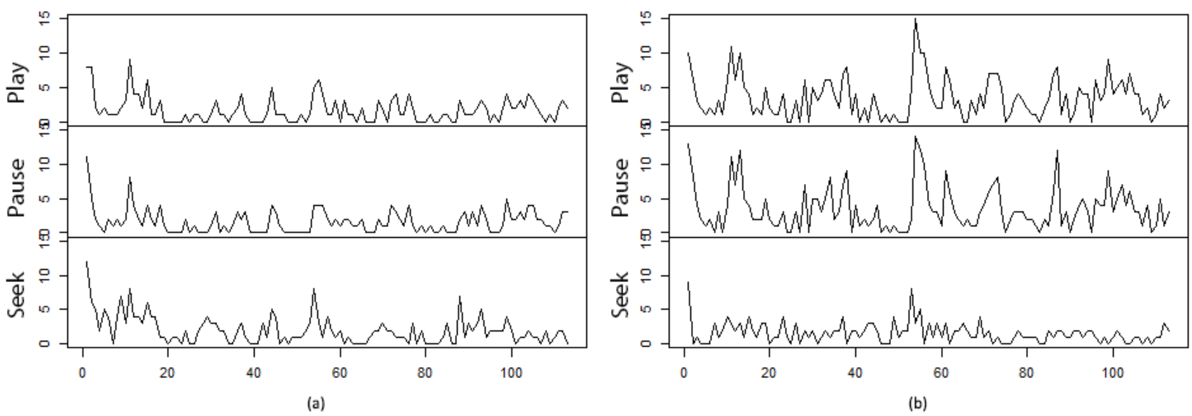

Figure 2. Time Series Graphs (A) Surface Learners (B) Deep Learners

DISCUSSION, CONCLUSION AND SUGGESTIONS

This study aimed to compare the video viewing behaviors of students with deep and surface learning approaches. In order to identify the video viewing behaviors of the students, nine features were specified, and a video player developed by the authors was used in order to collect data related to these features. With this research, first, whether there are any statistically significant differences between deep and surface learners in terms of video metrics was investigated. Then, the correlations between students' scale scores and the scores obtained from video metrics were examined. Finally, the video viewing data of deep and surface learners were visualized via time series graphs and were visually analyzed.

Results of the study showed that when deep and surface learners were compared in terms of video viewing behaviors, statistically significant differences were found only in terms of the number of forward seek between the two groups. When mean ranks of two groups compared it was understood that while viewing a video, surface learners seek forward more than the deep learner. The correlation analysis conducted within the scope of the second research problem revealing a positive and significant correlation between the numbers of forward seeks and students' surface approaches scale scores also supports this finding. When the deep and surface learners' time series graphs drawn in the third research problem are examined. It is observed that the students with the deep approach clicked on pause and play buttons more while viewing the video, and these actions peak in certain parts of the video. On the other hand, surface learners viewed the video more linearly with the random pause, play and seek actions. These viewing patterns can be used to identify problems in digital materials (video here) and to improve their quality (Ogata et al., 2018).

Learning approach is an indicator of what do students do and which methods they follow while carrying out a learning task (J. Biggs et al., 2001). And interaction data is an indicator of to what extent and in what manner students engage with the learning content. Therefore, students with surface approach are expected to also have surface interactions. Akcapinar (2015) investigated the learning approaches of the student groups who are active on different levels on the Moodle platform, and found that the students with low activity on the platform have high surface learning scores and low deep learning scores, while the students with high activity on the platform have high deep learning scores and low surface learning scores. This present study revealed that, compared to deep learners, surface learners engage in fewer interactions, and that these interactions are displayed randomly, rather than to learn. However, the conducted analyses showed a significant difference between the two groups, but solely in terms of forward seeking numbers. Since the session was conducted in a laboratory with a limited number of students, this may have led to a lack of significant differences between the students studied with the surface and deep approaches in terms of other features. For this reason, it would be useful to conduct a similar study on different videos outside a laboratory engaging a large group of students.

As also addressed by J. Biggs et al. (2001), learning approaches are not invariable characteristics of individuals and may vary depending on the factors such as courses, learning tasks, teachers etc. Dynamic prediction of this structure, which is also correlated with academic performance, based on the interaction data without directly asking students is important in terms of timely intervention on students using surface approach and providing them with feedback for encouraging the use of deep learning approaches (Akcapinar, 2016). The findings obtained here can be used to identify students with a surface approach and to develop intervention mechanisms for them. Instructors can use visual analysis results to re-design their learning materials.

BIODATA and CONTACT ADDRESSES of AUTHORS

Gokhan AKCAPINAR is a research assistant at the Department of Computer Education and Instructional Technology at Faculty of Education, Hacettepe University. Dr. Akcapinar received his Ph.D. in Educational Technology in June 2014. Currently he is conducting his Postdoctoral research at the Academic Center for Computing and Media Studies, Kyoto University, Japan. His academic interest areas are Educational Data Mining (EDM), Learning Analytics (LA), Educational Big Data Analysis, Sequential Behavior Analysis and Student Modeling etc. He has published a number of articles in peer-reviewed academic journals and presented papers in local and international conferences.

Dr. Gokhan AKCAPINAR
Department of Computer Education and Instructional Technology
Faculty of Education
Hacettepe University, 06800, Ankara, Turkey
Phone: +90 312 2977176,
E-mail: gokhana@hacettepe.edu.tr

Alper BAYAZIT is an Assistant Professor at the Department of Computer Education and Instructional Technology at Faculty of Education, Yeditepe University. Dr. Bayazıt received his Ph.D. degree in Educational Technology at February, 2013. His academic interest areas are educational data mining, learning analytics, artificial intelligence in education, and K-12 computer science education. He has five journal articles published in national and international indexes. Dr. Bayazit currently teaches the following courses: Fundamentals of Object Oriented Programming, Advanced Internet Programming, Database Management Systems, Artificial Intelligence in Education and Semantic Web Applications.

Dr. Alper BAYAZIT
Department of Computer Education and Instructional Technology
Faculty of Education
Yeditepe University, Kayisdagi Cad. 26 Agustos Yerleskesi 34755,
Atasehir, Istanbul, Turkey
Phone: +90 216 5780000 (3694)
E-mail: alper.bayazit@yeditepe.edu.tr

REFERENCES

Ak, S. (2008). Ogrenme yaklasimlarina iliskin kavramsal bir inceleme [A conceptual analysis on the approaches to learning]. *Kuram ve Uygulamada Egitim Bilimleri,* 8(3), 693-720.

Akcapinar, G. (2015). Profiling students’ approaches to learning through moodle logs. Paper presented at the *Multidisciplinary Academic Conference on Education, Teaching and Learning (MAC-ETL 2015),* Prague, Czech Republic.

Akcapinar, G. (2016). Predicting students' approaches to learning based on moodle logs. Paper presented at the *International Conference on Education and New Learning Technologies (EDULEARN16),* Barcelona, Spain.

Bayazıt, A., & Akcapinar, G. (2018). Çevrimiçi Dersler için Video Analitik Aracinin Tasarlanmasi ve Gelistirilmesi [Design and Development of Video Analytics Tool for Online Courses]. 2018, 17(1), 14-25. doi:10.17051/ilkonline.2018.413719

Beheshitha, S. S., Gasevic, D., & Hatala, M. (2015). A process mining approach to linking the study of aptitude and event facets of self-regulated learning. Paper presented at the Proceedings of the *Fifth International Conference on Learning Analytics and Knowledge,* Poughkeepsie, New York.

Biggs, J., Kember, D., & Leung, D. Y. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. *Br J Educ Psychol,* 71(1), 133-149.

Biggs, J. B. (1987a). Student approaches to learning and studying. Research Monograph: ERIC.

Biggs, J. B. (1987b). Study process questionnaire manual. Student Approaches to Learning and Studying: ERIC.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5), 318-331. doi:10.1504/IJTEL.2012.051815.

Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education,* 80, 108-121. doi:http://dx.doi.org/10.1016/j.compedu.2014.08.015.

Giannakos, M. N., Krogstie, J., & Aalberg, T. (2016). Video-based learning ecosystem to support active learning: application to an introductory computer science course. *Smart Learning Environments,* 3(1). doi:10.1186/s40561-016-0036-0.

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: an empirical study of MOOC videos. Paper presented at the Proceedings of the *first ACM conference on Learning @ scale conference*, Atlanta, Georgia, USA.

Hamm, S. (2009). Digital audio video assessment: surface or deep learning - an investigation. RMIT University.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks inonline lecture videos. Paper presented at the Proceedings of *the first ACM conference on Learning @ scale conference*, Atlanta, Georgia, USA.

Kleftodimos, A., & Evangelidis, G. (2016). An interactive video-based learning environment that supports learning analytics for teaching 'Image Editing'.

Marton, F., & Saljo, R. (1976a). On qualitative difference in learning - I: Outcome and process. *British Journal of Educational Psychology,* 46(1), 4-11. doi:10.1111/j.2044-8279.1976.tb02980.x.

Marton, F., & Saljo, R. (1976b). On qualitative differences in learning - II: Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology,* 46(2), 115-127. doi:10.1111/j.2044-8279.1976.tb02304.x.

Ogata, H., Oi, M., Mohri, K., Okubo, F., Shimada, A., Yamada, M., Hirokawa, S. (2018). Learning Analytics for E-Book-Based Educational Big Data in Higher Education. In H. Yasuura, C.-M. Kyung, Y. Liu, & Y.-L. Lin (Eds.), *Smart Sensors at the IoT Frontier* (pp. 327-350). Cham: Springer International Publishing.

Onder, I., & Besoluk, S. (2010). Duzenlenmis iki faktorlu calisma sureci olcegi'nin (R-SPQ-2F) turkceye uyarlanmasi [Adaptation of revised two factor study process questionnaire (R-SPQ-2F) to Turkish]. *Egitim ve Bilim,* 35(157).

Pi, Z., Hong, J., & Yang, J. (2017). Does instructor's image size in video lectures affect learning outcomes? *Journal of Computer Assisted Learning,* 33: 347–354. doi:10.1111/jcal.12183.

Rajaratnam, N., D'cruz, S., & Chandrasekhar, M. (2013). Correlation between the learning approaches of first year medical students and their performance in multiple choice questions in physiology. *NJIRM,* 4(5), 43-48.

Schiltz, G. (2015). Video Analytics: when and how do students use tutorial videos? Paper presented at the Proc. *23rd Int. Conf. on Computers in Education ICCE.*

Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D. E. (2014). Who does what in a massive open online course? *Commun. ACM,* 57(4), 58-65. doi:10.1145/2500876.

Siemens, G., & Baker, R. S. J. d. (2012). Learning analytics and educational data mining: towards communication and collaboration. Paper presented at the Proceedings of the *2nd International Conference on Learning Analytics and Knowledge.*