The Relationship Between Student Engagement and Academic Performance in Online Education

Ievgeniia Kuzminykh*
Department of Informatics, King's College London, UK

Bogdan Ghita
School of Engineering, Computing and Mathematics, University of Plymouth, UK

Hannan Xiao
Department of Informatics, King's College London, UK

ABSTRACT

In recent years, online education has become a mature, recognised, and heavily used alternative for delivering higher education programmes. Beyond its benefits, online education faces a number of challenges, some of which relate to its engagement and impact on student performance. To support the ongoing research into the complex relationships developed, this research investigated the relationship between engagement and academic performance for students that undertake standalone online programmes. The study uses as input the module content engagement data, as collected from an e-learning platform, including the number of content views, forum posts, completed assignments, and watching of videos. The study used Pearson correlation to evaluate the relationship between learner engagement and academic performance. The analysis revealed that the student engagement was positively correlated to the student performance both for individual modules as well as across the cohort. In addition, correlation between initial engagement with individual subjects and the overall engagement was also strong, indicating both variables lead to improved academic results.

CCS CONCEPTS

• **Social and professional topics**; • **Professional topics**; • **Computing education**; • **Student assessment**;

KEYWORDS

student engagement, academic achievement, correlation, e-learning

ACM Reference Format:
Ievgeniia Kuzminykh, Bogdan Ghita, and Hannan Xiao. 2021. The Relationship Between Student Engagement and Academic Performance in Online Education. In *2021 5th International Conference on E-Society, E-Education and E-Technology (ICSET 2021), August 21–23, 2021, Taipei, Taiwan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3485768.3485796

1 INTRODUCTION

Over the past years teaching methods have been subject to a digital transformation and a shift towards distance learning. The interaction limitations and overall context of the COVID-19 global pandemic coerced the education system to adjust the traditional approaches for knowledge acquisition and the associated methods of assessment. The various regional and national lockdowns required universities to adapt their on-campus face-to-face programmes to online delivery by implementing emergency remote teaching specifically focused on lecture-based classes, with some considerations for more active learning environment [16, 17]. Such changes, albeit ad-hoc in nature, supported students in their knowledge journey, making every effort to ensure that quality does not lose its value.

In the wider educational context, many universities have introduced standalone online learning along with the traditional face-to-face (on-campus mode) programmes, as an alternative provision that adapted the didactic material for a flexible active learning format with the availability of resources.

One of the underlying assumptions in education is that a high level of engagement can affect the quality characteristics of education, which leads to higher academic performance, and increased retention rates [10], as proven in a number of studies [1–3, 5, 9, 14, 19, 20, 26] . Historically, such studies focused on the analysis of on-campus student engagement. While their results are valid, the challenges associated with maintaining student motivation and engagement are significantly higher in a full online education environment. The basic requirement of keeping the student's attention is naturally enforced in a face-to-face environment, which assumes direct interaction between the lecturer and the audience. With the emergence of standalone online programmes where the flip class approach prevails, students are likely to commit their diminishing attention, be subject to a number of distracting factors, in an environment of greater emotional stress due to the lack of direct connection and support of academic and technical on-campus support [7, 21]. Subsequently, this raises additional concerns related to the relationship between student engagement and academic achievement. In addition, metrics for determining student involvement, such as attendance and physical abilities, which were the de-facto indicators in face-to-face or passive methods of learning, became ineffective [20], requiring a review of the ecosystem and definition of new metrics.

A feature of the standalone online programmes is that they are specially developed, taking into account minimal interaction with the educator/instructor, full reliance on a self-explanatory training content, and provision of unequivocal and precise instructions for completing assignments, as well as ensuring remote availability of practical environments [4]. To complete the learning experience, teacher-led seminars and sessions are used for discussion purposes only, aiming to resolve any questions relating to the learning journey, and sharing the difficulties faced by learners. In this context, the teacher plays a mentoring role rather than a knowledge transmitter role [8], providing for the functions of advising and consulting

*Corresponding author, email: ievgeniia.kuzminykh@kcl.ac.uk.

ICSET 2021, August 21–23, 2021, Taipei, Taiwan

ACM ISBN 978-1-4503-9015-6/21/08…$15.00
https://doi.org/10.1145/3485768.3485796

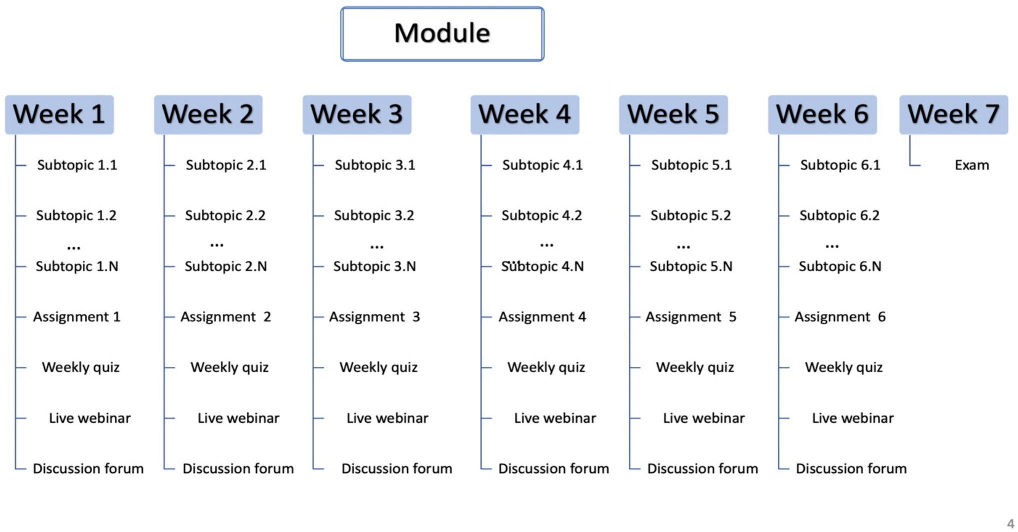

Figure 1: Module structure

information that guides learners throughout the knowledge field [2, 13]. Another specific feature of this form of training is that students consciously choose such programmes, knowing that they are expected to process the material most of the time and placing them in full control for acquiring the knowledge.

The purpose of this research is to evaluate the performance of students enrolled in an on-line programme as well as analyse their engagement and attainment data in order to determine whether there is any correlation between the pattern of engagement with the module materials and the marks awarded for the module. To narrow down the scope of the analysis and ensure relative consistency, the study will focus on a specific group, a set of cybersecurity modules.

The remainder of the paper is structured as follows: section 2 provides a brief description of the educational context for the study, then section 3 follows with a concise overview of the existing research. Section 4 presents the methodology followed by the research and section 5 provides the results and analysis, then section 6 summarises the findings.

2 THE EDUCATIONAL CONTEXT

A core characteristic of typical online programmes is that each subject is delivered in a block-based manner rather than running along with other modules, as it is the case with face-to-face programmes. As a typical example, a face-to-face programme may be delivering N modules (typically 2-3) in parallel during a period of 12 weeks, with 10 weeks allocated for teaching, one reading week, and one revision week, with the students undertaking additional exercising time in the small tutorial groups with teacher or teacher assistant. In contrast, an online programme will be delivering each of the N modules in a 12/N timeframe, with no reading week, no exercising in small groups, and a very small time window (typically 3-4 days) between last lecture and the exam taken. Such a limited timeline is stressful both for educators and students and, should the students not maintain their attention on the respective subject, may affect the quality of education, both teaching and learning. As a result of such design, the students who have a limited interaction or have a slower pace of progress, may not have the necessary time to progress and, at least in the early stages, might show less academic performance.

This study is based on a group of students undertaking a computing MSc, delivered online, which includes a set of specialist modules. Each module consists of six weeks of active learning, followed by an examination. The structure of the module is showed below in Figure 1. Each module includes a series of activities, such as reading the provided web content, undertaking small practical tasks, answering weekly quizzes, participating in the discussion on forum and attending the weekly webinar with the instructor. All these factors might be significant in tracking student engagement for further investigation.

3 RELATED WORKS

For the purpose of our study, we will focus exclusively on the cognitive and behavioural engagement aspects of university students. Analysis of the emotional dimension of the engagement is beyond the purpose of the study. The cognitive dimension refers to the students' personal investment [1], as well as existing learning approaches and self-regulatory strategies [3]. Behavioural engagement refers to behavioural norms, such as attendance and involvement, and would demonstrate the absence of disruptive or negative behaviour [26].

When referring to the students' engagement, we discuss the active participation in learning through such activities as reading of the materials, completing assignments, questions and answers in the forums, and watching the video materials.

Many research studies have shown a positive relation between student engagement and academic performance with higher engagement level associated with better grades [5, 10, 14]. Based on their approach and analysis, there are several directions for investigating student engagement. First, the studies that are devoted to determining of levels of student engagement [22][11]. Oriogun in [22] used student activities on the forum, expressed in the form of questions and answers, suggestions and explicit complex solutions, to determine their engagement levels. Kamath et al. [11] used picture recognition to measure student engagement. However, these studies only focused on one of the metrics as forum posts or video images did not consider other metrics. In addition, if the student did not have a video picture, their level could not be determined.

Other authors have used complex metrics such as the number of questions and answers on the forum and during live sessions, the number of interactions with instructor, and the time spent on the course to determine the engagement level of the students [24]. Similar metrics were used by the authors of [23] in their work, but also took into account the number of content views and assignment completion. However, the aim of these studies was to categorically define the level of involvement (on a multi-level scale) but not to study the effect of this level on academic performance.

The second direction of study is devoted to the influence of student motivation on the resulting academic performance. The authors in [28] concluded that the personal valuation of future goals promote prioritisation in achieving these goals, which, in turn, affects the engagement. On the same area of investigation, Shell and Husman [25] concluded that students who relate school subjects with the desired profession present superior cognitive skills and greater engagement in learning objectives and tasks. Ghazvini studied in [9] the relationship between academic self-concept and academic performance, and verified that the first positively predicts general performance in literature and mathematics. Similarly, the authors in [15] measured overall student engagement and its influence on academic achievement. The results showed the medium positive correlation between the facets of student engagement (behavioral, emotional, and cognitive) and academic achievement. However, for the analysis, the authors used already existing studies and their meta-data, and did not investigate any additional or new data.

Another dimension of the studies in this field is focused on e-learning. Among the latest research, the authors of [19] conducted an experimental analysis of engagement using rule-based machine learning algorithms and a number of identified metrics such as confidence and lift. The set of engagement metrics used included both frequency-related metrics and time spent on different tasks of the second year undergraduate students. They also collected data on exam scores, quizzes and assignments during the course. Analysis showed that a positive correlation exists between students' engagement level and their academic performance in a blended e-learning environment.

Finally, a significant number of research studies focused on investigating impact of student engagement on the resulting academic performance [5, 10, 14, 18]. The progress of academic performance regarding the engagement across school years was investigated by authors in [12] and they determined that students' engagement decreases as they progress from primary to secondary school and then to university.

The majority of the prior research used methods for collecting student engagement parameters inherent to the on-campus, face-to-face education model, such as class attendance and post-course survey. The authors in [14] investigated the relationship between engagement and reading performance in school and concluded that there was a positive correlation. Both [5] and [18] used a questionnaire to measure student engagement and investigated the relationship with academic achievement, although the latter work was conducted for an e-learning platform. The results also showed a positive correlation between student engagement (interaction with e-learning environment in the second study) and student performance.

As highlighted, the primary focus of the existing research has been on-campus learning or a blended learning model. Moreover, the papers that used the e-learning environment to identify relationships had limitations relating to their approach in measuring the engagement factors.

This paper aims to strengthen the state of the art in the area of student engagement and attainment by measuring students engagement in an e-learning (Moodle) environment and correlating it with their resulting academic performance. This investigation demonstrates that reliable metrics can also be identified for online learning environments for assessing engagement. In addition, this study is exclusively focused on a standalone online program, with its own specific characteristics described above.

4 METHODOLOGY

This study analyses the relationship between engagement and academic performance using as input the student interaction with online materials; the relationship between the two factors is determined using Pearson correlation. As part of the study, the engagement data and marks were anonymously collected for each student, then statistical analysis was applied to determine the correlation between the two. The overall steps of experiments are displayed in Figure 2. Student names and ID will be removed during the data collection phase; records associated with each student will be linked to a random ID, generated at collection time, and therefore no names or personal identification will be present in the resulting dataset.

4.1 Data Collection

The data was collected using embedded functionality of the Moodle e-learning platform. We collected data from 139 students on postgraduate level of an online programme on cybersecurity. The data was collected on a daily basis. The following parameters have been collected for measuring the engagement.

- Unique content web-page visits
- Repeated content web-page visits
- Posts on the Discussion forums
- Visits of web page with announcements
- Attempts in doing weekly quizzes
- Post-views of recorded webinars

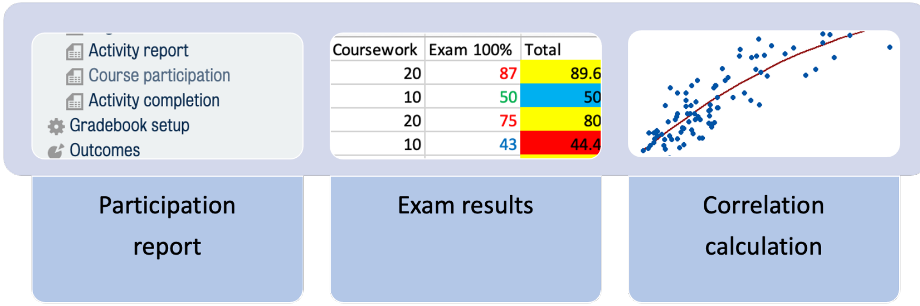

Figure 2: Data analysis methodology

Table 1: Results of correlation analysis between engagement and academic performance

Engagement	Correlation	p-value	Confidence interval	Samples
On network security module	0.57	0.001	[0.258;0.775]	29
On cryptography module	0.383	0.009	[0.101;0.608]	44
On security engineering module	0.62	3.56e-08	[0.443;0.75]	66
Initial engagement across all modules	0.343	3.617e-05	[0.187;0.482]	139
Overall engagement across all modules	0.419	2.86e-07	[0.271;0.547]	139

4.2 Data Extraction and Static Analysis

After raw data has been collected, we made the calculation of mean values of engagement parameters per topic, week, and total after the module completion and put it to the first dataset. The weekly patterns were also investigated but this is out of the scope for this study.

Similarly, the data related to the students final grades were sorted to accomplish the analysis and formed a second dataset.

4.3 Correlation Function and P-value

To measure the relationship between two datasets the Pearson's linear correlation [27] was employed. The correlation coefficient, r, determines both the strength and direction of the relationship between the dependent and independent variables. The r values range from -1.0 (strong negative correlation) to +1.0 (strong positive correlation). For two datasets X and Y, and two variables $x \in X$ and $y \in Y$, when $r = 0$, there is no relationship between the variables, where are the datasets. In addition, the p-value, $p \in [0; 1]$ determines the significance of the results in relation to the null hypothesis, as it describes how likely it is that the data occurred by random chance. The smaller the p-value, the stronger the evidence that you should reject the null hypothesis; the typical threshold for the p-value is 0.05.

5 RESULTS AND ANALYSIS

The analysis included a combined cohort of 139 postgraduate taught students, undertaking three modules over 126 days. The relationship between engagement and performance focused on the correlation across the three modules. The null hypothesis was that overall engagement, calculated as the sum of engagement across all individual elements for each student, has no relationship with the exam mark for the respective student. Applying the Pearson correlation across the entire dataset of 139 results, the resulting correlation factor is 0.419, with a p value of 2.86e-07 and a confidence interval of [0.271;0.547]. Table 1 below presents a summary of the results. In can be concluded that, both for each module as well as across the cohort, the overall engagement correlates with the exam mark.

While the overall engagement appears to be a good indicator for the resulting mark, it is also interesting to investigate the relationship between initial engagement and overall engagement. For this study, we define initial engagement as the student interaction with the content when first presented with it and during the week when these materials were introduced to the students. As a follow-up analysis, the overall engagement dataset was compared against the immediate engagement, more specifically at the end of the respective week of delivery, having the null hypothesis that the two types of engagement correlate. The resulting correlation factor was 0.831, with a p-value of 2.2e-16 and a confidence interval of [0.771;0.876]. Given the strong correlation, the result suggests that the level of initial engagement of the students was a good predictor for the final level of engagement across all modules. Indeed, as shown in Table 1, initial engagement also correlates with the exam mark

Finally, to complete the analysis, the two sets of results were compared using *cocor* [6] in order to determine whether the two correlations are statistically similar. Cocor uses a slightly different approach for comparing sample sets, whereby the null hypothesis is that the input sample sets are equal. The analysis, using Pearson and Fillon's z returned z=-1.680 and p-value=0.0929, which validates the null hypothesis that the two correlations are equal. This indicates that, for the group of students analysed, monitoring the two types of engagement leads to equivalent statistical results.

6 CONCLUSIONS

This study analysed the relationship between engagement and academic achievement for a group of postgraduate taught students. The results demonstrates that, in an online education context, both overall and initial engagement have an impact on academic performance. In addition, for the analysed data samples, initial and overall engagement are strongly correlated, indicating a consistent pace of study across the duration of the module. This validates and expands the conclusions of prior research, focused on face to face, on-campus education, for online education.

A larger dataset may further strengthen the validity of the results, but a more relevant direction for future research is analysing the pattern of study and investigating any additional factors and metrics that influence academic performance. Amongst such factors, the student attitude, level of interest, and prior knowledge for the specific subject are likely to be of interest and may lead to a more complete picture of the student learning journey.

REFERENCES

[1] Mary D. Ainley. 1993. Styles of engagement with learning: Multidimensional assessment of their relationship with strategy use and school achievement. J. Educ. Psychol. 85, 3 (1993), 395–405. DOI:https://doi.org/10.1037//0022-0663.85.3.395
[2] John Biggs and Catherine Tang. 2011. Teaching according to how students learn. Teach. Qual. Learn. Univ. (2011), 16.
[3] Phyllis C Blumenfeld, Alison H Paris, and Jennifer A Fredricks. 2004. School Engagement: Potential of the Concept, State of the Evidence. Rev. Educ. Res. 74, 1 (2004), 59–109.
[4] Anders Carlsson, Ievgeniia Kuzminykh, and Rune Gustavsson. 2019. Virtual Security Labs Supporting Distance Education in ReSeLa Framework. Adv. Intell. Syst. Comput. 917, (2019), 577–587. DOI:https://doi.org/10.1007/978-3-030-11935-5_55
[5] Maria J. Casuso-Holgado, Antonio I. Cuesta-Vargas, Noelia Moreno-Morales, Maria T. Labajos-Manzanares, Francisco J. Barón-López, and Manuel Vega-Cuesta. 2013. The association between academic engagement and achievement in health sciences students. BMC Med. Educ. 13, 1 (2013). DOI:https://doi.org/10.1186/1472-6920-13-33
[6] Birk Diedenhofen and Jochen Musch. 2015. Cocor: A comprehensive solution for the statistical comparison of correlations. PLoS One 10, 4 (2015). DOI:https://doi.org/10.1371/journal.pone.0121945
[7] Online Education. 2020. Online: Trending Now. (2020), 1–8.
[8] R Gallegos Nava. 2001. Holistic education: Pedagogy of universal love. (2001).
[9] Sayid Dabbagh Ghazvini. 2011. Relationships between academic self-concept and academic performance in high school students. Procedia - Soc. Behav. Sci. 15, (2011), 1034–1039. DOI:https://doi.org/10.1016/j.sbspro.2011.03.235
[10] Peter Kahn, Lucy Everington, Kathleen Kelm, Iain Reid, and Francine Watkins. 2017. Understanding student engagement in online learning environments: the role of reflexivity. Educ. Technol. Res. Dev. 65, 1 (2017), 203–218. DOI:https://doi.org/10.1007/s11423-016-9484-z
[11] Aditya Kamath, Aradhya Biswas, and Vineeth Balasubramanian. 2016. A crowdsourced approach to student engagement recognition in e-learning environments. 2016 IEEE Winter Conf. Appl. Comput. Vision, WACV 2016 (2016). DOI:https://doi.org/10.1109/WACV.2016.7477618
[12] J. P. Klem, A. M., & Connell. 2004. Relationships matter: Linking teacher support to student engagement and achievement. J. Sch. Health 74, 7 (2004), 262–273.
[13] Ievgeniia Kuzminykh, Maryna Yevdokymenko, Oleksandra Yeremenko, and Oleksandr Lemeshko. 2021. Increasing Teacher Competence in Cybersecurity using the EU Security Frameworks. Manuscr. Submitt. Publ. (2021).
[14] Jung Sook Lee. 2014. The relationship between student engagement and academic performance: Is it a myth or reality? J. Educ. Res. 107, 3 (2014), 177–185. DOI:https://doi.org/10.1080/00220671.2013.807491
[15] Hao Lei, Yunhuo Cui, and Wenye Zhou. 2018. Relationships between student engagement and academic achievement: A meta-analysis. Soc. Behav. Pers. 46, 3 (2018), 517–528. DOI:https://doi.org/10.2224/sbp.7054
[16] Oleksandr Lemeshko, Tetiana Strilkova, Oleksandra Yeremenko, Maryna Yevdokymenko, and Ievgeniia Kuzminykh. 2021. Experience of Adaptation and Organization of Distance Learning in Ukrainian Universities. In Revealing Inequities in Online Education During Global Crises.
[17] Oleksandr Lemeshko, Tetiana Strilkova, Maryna Yevdokymenko, and Ievgeniia Kuzminykh. 2020. Features of creating the virtual cybersecurity lab for distance learning. New Coll. 3, (2020), 41–45.
[18] Mohamed Jama Madar and Othman Bin Ibrahim. 2011. E-learning towards student academic performance. 2011 Int. Conf. Res. Innov. Inf. Syst. ICRIIS'11 (2011). DOI:https://doi.org/10.1109/ICRIIS.2011.6125718
[19] Abdallah Moubayed, Mohammadnoor Injadat, Abdallah Shami, and Hanan Lutfiyya. 2018. Relationship between student engagement and performance in e-learning environment using association rules. EDUNINE 2018 - 2nd IEEE World Eng. Educ. Conf. Role Prof. Assoc. Contemp. Eng. Careers, Proc. (2018). DOI:https://doi.org/10.1109/EDUNINE.2018.8451005
[20] Rabindra Nepal and Ann M. Rogerson. 2020. From theory to practice of promoting student engagement in business and law-related disciplines: The case of undergraduate economics education. Educ. Sci. 10, 8 (2020), 1–13. DOI:https://doi.org/10.3390/educsci10080205
[21] OECD. 2020. The impact of COVID-19 on student equity and inclusion: supporting vulnerable students during school closures and school re-openings. OECD Publ. (2020), 1–37. Retrieved from https://oecdedutoday.com/coronavirus-students-special-education-needs/
[22] Peter K. Oriogun. 2003. Towards understanding online learning levels of engagement using the SQUAD approach to CMC discourse. Australas. J. Educ. Technol. 19, 3 (2003). DOI:https://doi.org/10.14742/ajet.1726
[23] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daum, and Lise Getoor. 2013. Modeling Learner Engagement in MOOCs using Probabilistic Soft Logic. NIPS Work. Data Driven Educ. (2013), 1–7.
[24] Leslie F Reid. 2012. Redesigning a Large Lecture Course for Student Engagement: Process and Outcomes. Can. J. Scholarsh. Teach. Learn. 3, 2 (2012). DOI:https://doi.org/10.5206/cjsotl-rcacea.2012.2.5
[25] Duane F. Shell and Jenefer Husman. 2001. The multivariate dimensionality of personal control and future time perspective beliefs in achievement and self-regulation. Contemp. Educ. Psychol. 26, 4 (2001), 481–506. DOI:https://doi.org/10.1006/ceps.2000.1073
[26] Vicki Trowler. 2010. Student engagement literature review. High. Educ. November (2010), 1–15. Retrieved from http://americandemocracy.illinoisstate.edu/documents/democratic-engagement-white-paper-2_13_09.pdf
[27] Paul Vogt. 2011. Dictionary of Statistics & Methodology. Dictionary of Statistics & Methodology. DOI:https://doi.org/10.4135/9781412983907
[28] Allan Wigfield and Jacqueline Eccles. 2000. Expectancy Value theory of motivation. Contemp. Educ. Psychol. 25, (2000), 68–81.