Personal and engagement factors in remote learning contributing to high student achievement in a transnational education context: A quantitative study

Hatahet, T., Mohamed, A. A. R., Malekigorji, M., Clerkin, C., & Sandrey, L. (2024). Personal and engagement factors in remote learning contributing to high student achievement in a transnational education context: A quantitative study. *Advances in Online Education*, *2*(4), 330-350. https://hstalks.com/article/8492/personal-and-engagement-factors-in-remote-learning/

Published in:
Advances in Online Education

Document Version:
Peer reviewed version

Personal and engagement factors in remote learning contributing to high student achievement in a transnational education context: a quantitative study.

Taher Hatahet[ab*], Ahmed A. Raouf Mohamed[c], Maryam Malekigorji[d], Caoimhe Clerkin[ab], Liane Sandrey[ab]

[a] School of Pharmacy, Queen's University Belfast, Belfast, BT9 7BL, UK

[b] China Medical University and Queen's University Joint College, Shenyang, Liaoning, China.

[c] School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, BT9 5AG, UK

[d] Institute of Health Sciences Education - Faculty of Medicine and Dentistry, Queen Mary University of London, London, E1 4NS, UK

Correspondence Author:

Dr. Taher Hatahet,

Postal address School of Pharmacy, Queen's University Belfast, 97 Lisburn Road BT9 7BL, Belfast, UK,

Email: t.hatahet@qub.ac.uk

Phone: +44 (0)28 9097 5732.

ORCID: 0000-0003-0171-3304

Abstract

Remote learning, adopted by many institutions as a temporary measure during the COVID-19 pandemic, has been found to be a viable option in many educational settings.

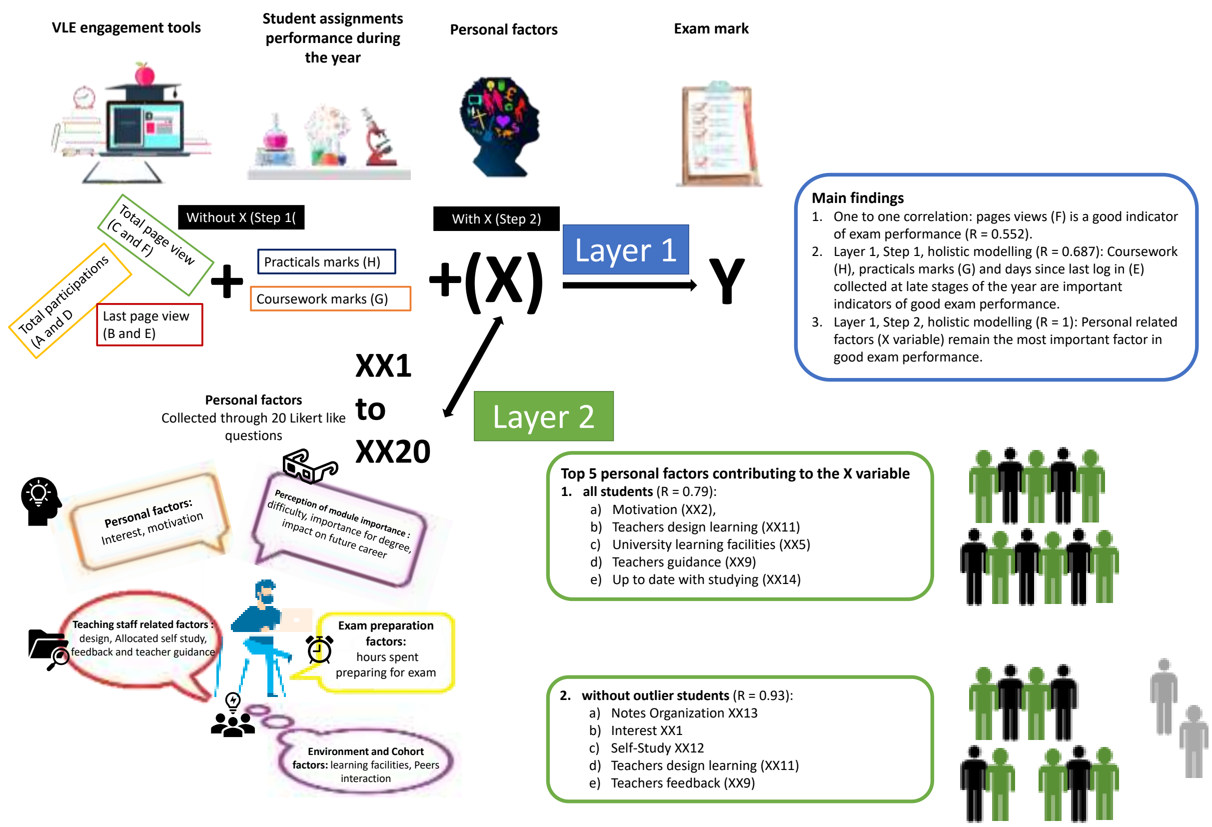
Including transnational higher education, remote learning makes extensive use of technology such as virtual learning environments (VLEs) and video conferencing applications. Understanding how these tools can be used to predict students' academic performance, has brought huge benefits to both students and educators. This revealed the main determinant of student success to be personal factors/individual differences that are yet to be investigated.

Taking exam performance as an outcome variable, this research utilises a BSc pharmaceutical course student perception on a survey to quantitatively identify the personal factors which contribute to academic achievement through the construction of a new X variable. The findings confirm that page views values operate as a simple one-to-one predictor of exam performance regardless of student cohort. In a holistic model, days since last log in, and practical and coursework marks had a high influence on the generated equations. Additionally, self-study skills, note taking, and interest in the subject matter as personal factors were found to be key determinants of student engagement in remote settings. Conversely, less engaged remote learners appeared to struggle with motivation and keeping up to date with their studies, relying to a high extent on teacher support and university facilities during their learning journey. For the first time, this paper addresses personal factors contributing to success in remote learning settings using mathematical modelling and machine learning, thereby avoiding subjectivity and bias in assessment. The findings will be of use to designers of remote learning in general, and more specifically in transnational education.

Keywords:

Academic performance, linear regression, modelling, student engagement, transnational education, virtual learning environment, Personal factors

Graphical abstract



1. Introduction

The role of technology in higher education has witnessed substantial growth in recent years, then further boosted by the emergence of the COVID-19 pandemic (Alzubaidi et al., 2021; Jeffries et al., 2022; Weller, 2007). The ability to store and easily disseminate lectures, online activities such as quizzes or to create asynchronous discussion boards on virtual learning environments (VLEs) catapulted the effectiveness of online learning. As a reaction to the COVID-19 pandemic, video conferencing applications supported live teaching to take place remotely. Further, blending online asynchronous teaching with face-to-face teaching has been proven to be preferred by students, in a Chinese context, and could be implemented to support future transnational educational (TNE) program (Clerkin et al., 2022). However, for better design of learning, the way students interact with blended learning on the personal side has yet to be investigated.

1.1. Transnational education (TNE)

Transnational education (TNE) or cross border education can be defined as the delivery of Higher Education (HE) to undergraduate and postgraduate students in a country other than the country in which the awarding institution is based (Tsiligiris et al., 2018), with the earliest definition of TNE being less than three decades old (Sharma & Johanson, 1987). A high increase in competition for university places in many countries has created an opportunity for high-status universities to offer offshore education to students in developing nations. For example, TNE can been delivered by distance or remote learning or through the establishment of a licensed international branch campuses (IBC) (Healey, 2020). This opportunity has seen high interest from students who were denied a place locally, simply failed to secure a place in a top local university or who come from high income families aiming for high quality education for their children and who are capable of supporting high tuition fees (Healey, 2020).

From the perspective of local universities, TNE programs in the form of partnering arrangements can offer competitive advantage in local recruitment through relationships with highly ranked (overseas) universities. As a result of these benefits, the TNE market has become a very competitive place among top universities from countries like the UK, Canada and Australia (Healey, 2020). This has in turn created pressure for institutions involved in TNE to not simply offer education but to provide high quality education that attracts the attention of host governments and the recruitment of good local students.

1.2.Teaching and learning in transnational programs.

TNE programs can be delivered by face-to-face teaching, through remote learning, or through a blended approach which employs elements of both. In the case of face-to-face teaching, ‘traveling’ faculty staff can travel from the main campus to the remote location to deliver in-person teaching, or locally trained staff can be recruited for that purpose; a mix of both these approaches can also be used. On the other hand, remote or distance learning involves the delivery of educational content completely or partially from the main campus through online learning, video conferencing applications, and the use of a ‘virtual learning environment’ (VLE) (Heitz et al., 2020). This can be asynchronous (e.g., pre-recorded materials stored on VLEs) or synchronous (delivered live through online teaching platforms such as Microsoft (MS) Teams or Zoom). Finally, both face to face and remote learning can be delivered in a blended approach where pre-recorded materials are made available to students followed by live - either online or in class - teaching on several occasions within the module. The emergence of COVID-19 pandemic pushed everything to remote learning (Alzubaidi et al., 2021).

An understanding about how students can succeed in remote learning settings is crucial for the design of the optimised learning environment. One of TNE’s main challenges is developing learning strategies which keep students engaged with content remotely, and therefore understanding personal factors associated with learners’ engagement and how they can be monitored are the focus of this project.

1.3.Student learning in remote learning

1.3.1. Learning approach is affected by context.

Students' approach to studying can be a reflection of their perception of the learning environment and hence can be affected by remote learning settings (Richardson, 2005). Furthermore, how interested students are in a subject area and/or how intrinsically or extrinsically motivated they are may be linked to which learning approach they prefer form superficial to deep and through use of strategic learning (Richardson, 2005). Understanding motivation dynamics in remote learning can be crucial to understand academic achievement.

1.3.2. Motivation and interest in remote learning

Motivation to learn is associated with student success in learning and has both intrinsic and extrinsic origins. Intrinsically motivated behaviour can be linked to the term 'self-determined' according to self-determination theory (SDT) proposed by Deci and Ryan, where clear internal intentions are the driving force for the action of learning (Deci & Ryan, 1994). Students who are self-determined are prepared to behave autonomously as the result of their own desire to learn the subject, which is a key factor in being able to cope with changes to the education system e.g. the switch to remote learning from face to face teaching (Chiu, 2022; Müller & Louw, 2004). Intrinsically motivated students have individual interests which coincide with the subject of study (Hoffmann et al., 1998). These interested students will engage in classes and other learning activities pushed by an inner appreciation of the degree or the module they are studying.

On the other hand, extrinsically motivated learner behaviour is linked to the learning context, and the learning environment and design of learning materials can impact such motivation (Chiu, 2022; Deci & Ryan, 1994; Hoffmann et al., 1998; Shah et al., 2021). Extrinsic motivation derives from the complex interaction between external, introjected, identified and integrated regulations (Müller & Louw, 2004). Students driven by extrinsic motivations are more sensitive to changes in the learning context, such as the switch from social, face to face learning to the more individualistic remote learning. Instead of continuing to pursue the deep learning approach which they may have followed during face-to-face teaching, top students may adopt a more strategic approach in response to a switch to remote learning.

SDT can be used to assess motivation in online (and remote) learning settings. SDT tackles autonomy, relatedness, and competency as determinants of motivation (Chen & Jang, 2010). It follows Chen and Jang's path of causation. This research "contextual support (matched with environment and cohort and teaching staff factors) can help students to satisfy their need (perceptual view of module importance to degree or career), building students self-determination (during year and close to exam preparation) to meet final learning outcome," (Supplementary data Table 1) (Chen & Jang, 2010). In the case of motivated students, which skillset is needed to cope up with remote learning settings is worth investigating.

1.3.3. Self-study skills in remote learning

The notion of being autonomous is central in human activity to improve self-esteem, a construct that shadows motivation and well-being (Deci & Ryan, 1995). Equally important as autonomy in personal behaviour is the acquisition of knowledge. Higher education graduates are people who have built knowledge and are capable of finding ways to expand it throughout their lives (Krzyżak & Walas-Trebacz, 2021). With conventional teaching such as face to face lectures, tutorials and practicals, student autonomy manifests in exam preparation, completing assignments and even taking notes during lectures (Chiu, 2022). In comparison, remote learning requires a higher level of autonomy which goes beyond simply responding to learning opportunities into the realms of creating such opportunities through engagement with pre-recorded materials. This autonomy can be linked to self-study skills, which involve not only completing the learning tasks independently but also self-assessing own's understanding and adjusting learning strategies accordingly (Ayres & Paas, 2009).

In a previous study, VLE factors related to success were studied in remote learning scenarios across a transnational education programme between the United Kingdom and China (Hatahet et al., 2022). Student exam performance, engagement factors on the VLE, coursework and practical marks were utilised to generate a mathematical model as a predictive tool for students' academic achievement. The generated models showed high correlation coefficients (> 0.65) between VLE engagement indicators, assignments marks and final exam marks (Hatahet et al., 2022). However, the highest contributor to students' academic performers that can create a model with perfect fit was an unknown (X) variable. The assumption is that this variable is more related to personal differences among students (Hatahet et al., 2022).

1.4.Educational context

This project aims to investigate and identify the factors which contribute to this unknown X variable from the student perspective by modelling student responses to a questionnaire (Figure 1). This will be achieved through:

a) Aiming to generate a model with perfect correlation. VLE engagement indicator, coursework, practical marks, and final exam marks will be modelled with an aim of developing an R close to 1, thus calculating the X variable (personal factor) for each student in this cohort (Layer 1 modelling).

b) Using the calculated X variable as an output data and modelling students answers to a perception questionnaire (Layer 2 modelling) to identify key personal factors associated with good exam performance in remote learning settings.

2. Methods

2.1. Study participants

VLE engagement reports were collected alongside coursework and final exam marks for 56 students studying Level 2 BSc Pharmaceutical Sciences during the academic year 2020-2021. The students are enrolled off campus at China Medical University - the Queen's University Joint College (CQC) in Shenyang, Liaoning province, China. The transnational education programme is jointly run with China Medical University as a TNE partner institution. The individual scores used in this paper are presented in Supplementary data Figures 1 and 2.

2.2. Study design

Two layers of modelling were investigated in this paper (Figure 1). The first set of modelling was based on a published work on modelling Canvas VLE engagement indicators with coursework and final exam marks (Hatahet et al., 2022). Following the same principle using the Industrial Pharmaceutics module in BSc Pharmaceutical Sciences, student's engagement reports were collected and modelled with students' coursework and final exam marks. Continuous engagement data, and practical and coursework marks were set as the input data while the final exam was set as the output data. Data was collected twice in the academic year. Total number of participation occasions (A and C), days since last log in (B and E) and page views (C and F) were used as indicators for students' level of engagement with the VLE, with first and second data collection points in November and March in the academic year 21-22, respectively. In parallel, all practicals

(G) and coursework marks (H) were averaged. The final examination mark was abbreviated as (Y). The used coding was similar to the previously published work (Hatahet et al., 2022). The obtained layer 1 equations were compared to a previous work in terms of correlation coefficients and statistical significance values. The most important variables were also compared between generated models. The models were created with and without an extra X variable that represent the personal factors associated with near perfect prediction of final exam mark ($R^2 \approx 1$).

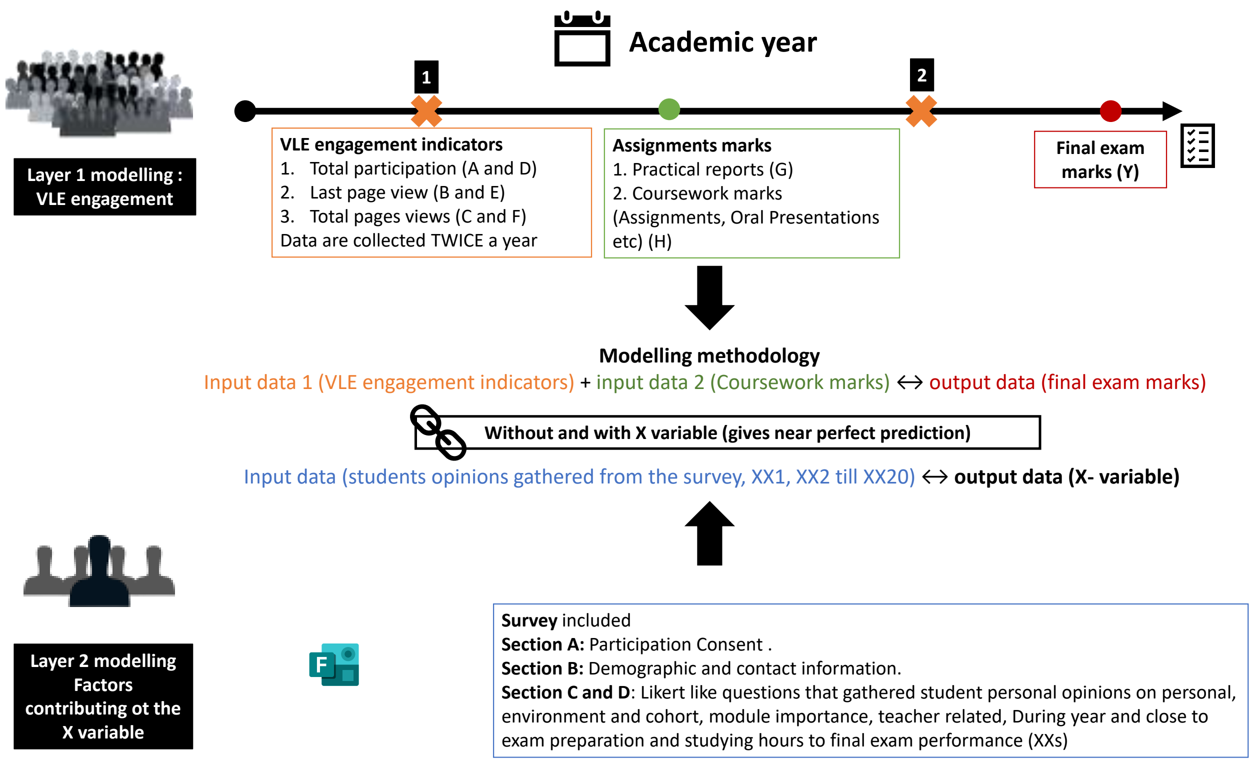


Figure 1: A schematic representation of the study design.

The second layer of modelling was based on assessment of the main influences on the X variable according to students' self-reported personal assessment. A survey was developed using Likert like questions to investigate students' opinions in 7 areas related to module exam performance (Supplementary Table 1). The survey instrument was peer reviewed and upon ethical approval by Faculty of Medicine and Life Sciences ethical committee (Reference Number: MHLS 21_93) was distributed to students after finishing their final year exam (n = 56). The survey remained open for one month. Cronbach's alpha was used to assess the instrument's validity. The final survey was composed of 4 sections.

- Section A: contained 7 questions to confirm participant consent and the voluntary nature of participation in the study and consent to the potential publication of anonymised findings in accordance with the obtained ethical approval.
- Section B contained 3 general information questions in order to be able to connect individual participants to their module performance data.
- Sections C and D contained 20 Likert style questions that gathered student personal opinions on factors related to their performance in the final module exam (Supplementary Table 1).

Forty five percent of students completed the survey (1 mathematical outlier, section 2.3) and their answers were used to create a mathematical model that links their obtained X variable score from Layer 1 step 2 and the individual answers (Layer 2). Correlation coefficients and statistical significance data were determined for the obtained equations with and without outliers. Afterwards, individual contributors (XXs) in the equations were ordered in terms of importance for the model (Layer 2) and were compared between equations generated from modelling students (with or without outliers).

2.3. Data modelling and analysis

The modelling process is divided into two layers as outlined in Figure 2. MATLAB software was used for the modelling in this paper (*MathWorks - Makers of MATLAB and Simulink - MATLAB & Simulink*, n.d.). For the first layer, the prediction variables are denoted by letters from A to H.

The output variable (exam mark) is symbolised by Y. The first layer consists of two separate steps. The first step is to produce a simple linear regression model between the predictors and the output. The second step aims to find the optimal value of an additional variable (symbolised by X) that can be added to improve the regression model accuracy. This was formulated as black-box optimisation (Audet & Hare, 2017) and solved using the Genetic Algorithm (GA) (Whitley, 1994) due to its ability to provide good solutions in a reasonable time.

GA is an optimisation algorithm based on a search process used to find optimal or near-optimal solutions for complex problems based on the genetics selection principle. The objective function of this optimisation process is to maximise the Pearson correlation coefficient (R-Value) using a decision variable 'X' that has a lower bound of zero and an upper bound of ten. The GA optimiser iterates by updating the decision variables until convergence which represents an optimal value of the Pearson correlation coefficient (R=1) for each student. The output of this step (Step 2 of Layer 1) is another regression model with a new predictor (X-variable). The outliers were detected and discarded using the generalised extreme studentised deviate test for outliers (Walfish, 2006). This iterative method removes one outlier per iteration based on hypothesis testing. It assumes that the data is normally distributed and has the capability to perform well when there are multiple outliers masking each other. The previous steps of the first layer (Step 1 and Step 2) are then repeated without outliers (N = 42). Therefore, a total number of four prediction models have been developed (Equations 1 to 4).

The second layer uses prediction variables representing the personal factors influencing the final exam mark of students. These factors were captured using a 20-factor survey as described in Section 2.2. Such factors were used as prediction variables to generate a prediction model using linear regression, symbolised as XX1 to XX20 (Supplementary Table 1). As illustrated in Figure 2, the output of this layer is the X variable generated from layer 1 step 2 that is linked to all the personal factors affecting students' performance (XX2 to XX20). The objective function was again to maximise Pearson correlation coefficient (R-Value).

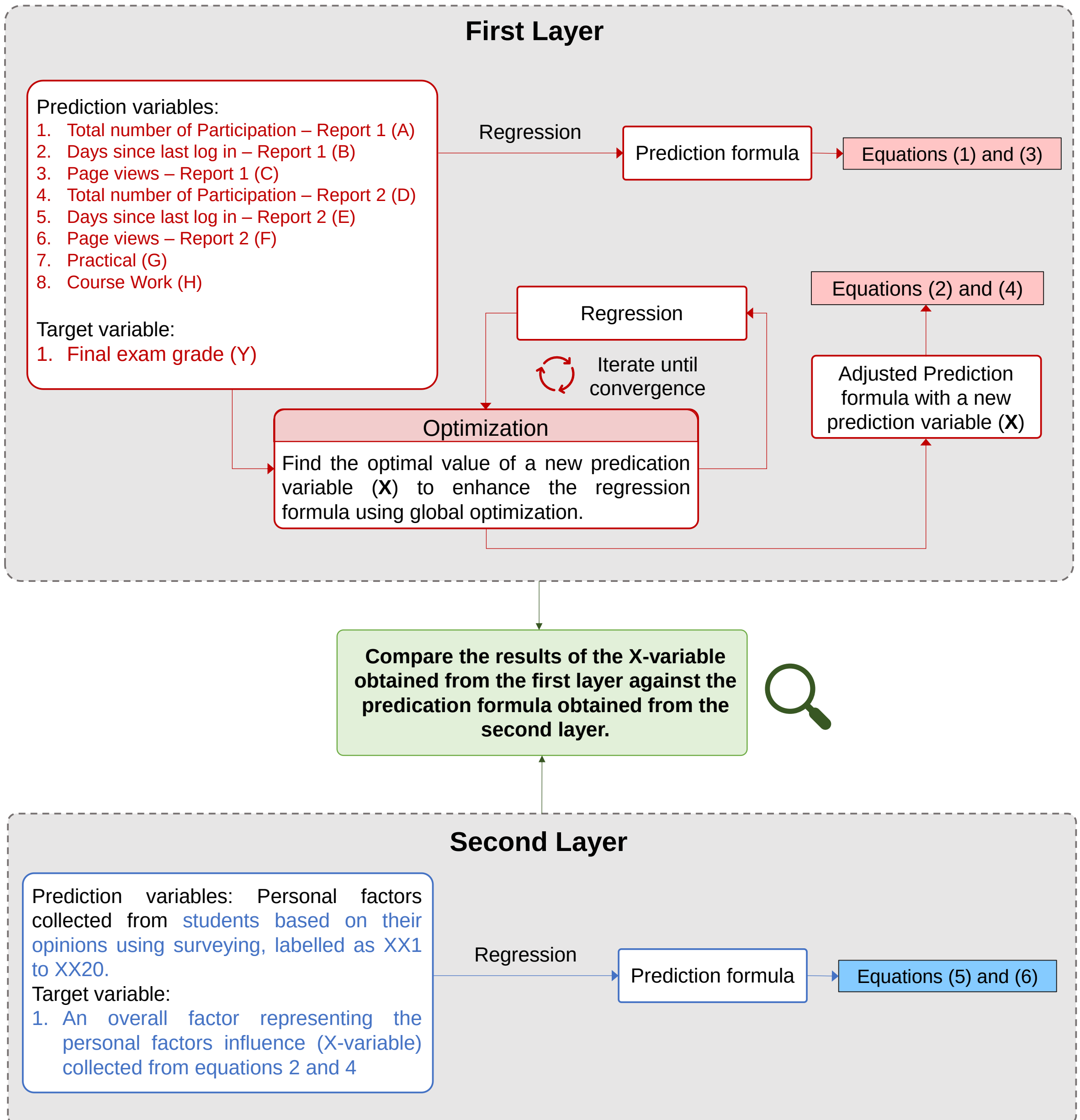


Figure 2: A schematic representation of the stages of the modelling process

2.4. Data storage and accessibility

All collected data, whether collected for VLE engagement, module marks or survey answers, was coded and anonymised in a one step process; any identifiable data was subsequently permanently deleted. All data was encrypted and stored on a password protected disk drive only accessible to the authors.

3. Results

3.1. Layer 1 modelling: virtual learning environment engagement, coursework, and final exam marks prediction

Table 1 presents correlation coefficients between individual variables and the final examination mark. Page views from VLE engagement reports along with coursework marks appeared as being important predictors (both with R-value>0.5).

Table 1: Pearson Correlation Coefficients between virtual learning environment engagement and coursework variables A-H and final exam mark Y with (n= 55) or without outliers (n = 42).

Data source	Variable	Pearson Correlation Coefficients (R-value)	
		with outliers (N = 56)	without outliers (N = 42)
Virtual learning environment engagement repot 1	Total number of Participation **A**	0.240	0.184
	Days since last log in **B**	-0.238	-0.190
	Page views **C**	**0.400**	**0.532**
	Total number of Participation **D**	0.228	0.176

Virtual learning environment engagement repot 2	Days since last log in **E**	0.044	0.145
	Page views **F**	**0.378**	**0.552**
In-year continuous assessments	Practicals **G**	0.219	0.197
	Coursework **H**	**0.492**	**0.503**

Table 2 presents detailed layer 1 modelling equations. Days since last log in (E), Practical (G) and Coursework (H) performance were the main contributors to the equation with and without outliers (equations 1 and 3). When adding the X variable without outliers (Equations 4), X becomes the most important variable to the equation followed by Days since last log in (E), Coursework (H) and practical performance (G). Furthermore, report 2 data (D, E and F) weighted more in the equations than report 1 data (A, B and C) which were collected earlier in the academic year. The removal of outliers enhanced the role of Days since last log in (E, which highlights that outliers are students with either very low or very high engagement associated with their final exam mark).

Generated equations metrics (Table 3) reflect higher correlation coefficients obtained by removing outliers (Pearson correlation coefficient 0.594 v.s 0.687). All models were statistically significant ($p < 0.05$). It is worth noting that generated models showed higher correlation coefficients to page views alone (Table 3 v.s Table 1)

Table 2: List of the generated equations (Layer 1) from simple linear regression of virtual learning environment engagement (A-F), coursework marks (G and H) variables and final exam marks (Y) with (N = 55) or without the X variable (N = 42).

Modelling type	**Equation number**	**Equation**
With outliers		

Simple linear regression model without optimization	1	$Y = 68.62 + \mathbf{1.43E} - \mathbf{0.61G} + \mathbf{0.416H} + 0.281A - 0.17D - 0.161B + 0.0043C + 0.004F$
Simple linear regression with a new predictor (adjusting variable X)	2	$Y = 1.23 + \mathbf{10.92X} + \mathbf{0.951E} + \mathbf{0.457B} + \mathbf{0.374H} + 0.242A - 0.209D - 0.186G + 0.012C + 0.002F$
Without outliers		
Simple linear regression model without optimization	3	$Y = 108.1 + \mathbf{3.15E} - \mathbf{1.2G} + \mathbf{0.299H} - 0.192D + 0.082A - 0.0587B + 0.0118F + 0.0102C$
Simple linear regression with a new predictor (adjusting variable X)	4	$Y = -114.4 + \mathbf{11X} + \mathbf{5.54E} + \mathbf{0.845G} + \mathbf{0.554H} + 0.4A + 0.122B - 0.231D + 0.0001F - 0.0024C$

Table 3: Statistical metrics of the generated models, with or without the X variable (Layer 1).

Modelling type	Equation number	Number of observations Error degrees of freedom	Root Mean Squared Error	Pearson Correlation Coefficient (R-value) R^2 Adjusted R^2	P-value
With outliers					
Simple linear regression model without optimization.	1	56 47	11.1	0.594 0.353 0.243	5.6×10^{-3}**
Simple linear regression with a new predictor (adjusting variable X)	2	56 46	0.224	1 1 1	2.3×10^{-73}***
Without outliers					
Simple linear regression model	3	42 33	10.2	0.687 0.474 0.343	3.6×10^{-3}**

without optimization. Simple linear regression with a new predictor (adjusting variable X)	4	42 33	0.0991	1 1 1	9.66×10^{-69}

3.2. Layer 2 modelling: understating the main influences on the X variable according to students' personal assessment.

Reliability analysis indicates that the questionnaire survey instrument developed for use in the research has acceptable reliability and internal consistency, with a Cronbach's alpha value of 0.924 (Barbera et al., 2020).

Table 4 shows correlation coefficients between the individual perception variables and the X variable value obtained from layer 1 modelling. All the variables (XX) showed weak correlation to X with 14 variables out of 20 showing R-value scores of less than 0.2. The remaining 6 variables with correlations greater than 0.2 revealed 5 negative correlations and 1 positive correlation (the self-study factor). In summary, the highest individual correlations between individual variables and the X variables, whether negative or positive, were centered around teaching staff, preparation throughout the year, and module importance for the degree (40 CATS module in 120 CATS year), which were not greatly affected by outliers.

Table 4: Pearson Correlation Coefficients between student perception individual variables XX1 to XX20 and the X variable value from layer 1 modelling with (N = 25) or without (N = 24) outliers.

Area of variables		Pearson Correlation Coefficients (R-value)

	Description of Variable (Variable symbol)	Without Outliers	With Outliers
Personal factors	Interest XX1	-0.075	-0.047
	Motivation XX2	-0.063	-0.056
	Sleep Before Exam XX3	0.060	0.065
Environment and cohort factors	Peers' Communication XX4	0.040	0.040
	University Leaning Facilities XX5	-0.127	-0.108
Perception of module importance	**For Degree XX6**	**-0.259**	**-0.256**
	For Future XX7	-0.051	-0.039
	Difficulty XX8	-0.132	-0.133
Teaching staff related factors	**Guidance XX9**	**-0.288**	**-0.287**
	Feedback XX10	**-0.226**	**-0.215**
	Design XX11	**-0.253**	**-0.206**
During year preparation	**Self-Study XX12**	**0.202**	**0.203**
	Notes Organization XX13	-0.044	-0.044
	Up to date Studying XX14	0.081	0.083
	Study Plan XX15	**-0.266**	**-0.254**
Close to exam preparation	Enough preparation time Close to Exam XX16	-0.056	-0.059
	Practicing Quiz XX17	0.010	0.020
	Practicing Exam XX18	-0.142	-0.110
	Understand Not Memorise XX19	-0.126	-0.084

	Studying hours XX20	0.171	0.161

Table 5 presents equations generated for layer 2 modelling, where the effect of outliers is remarkable. Removing the outliers results in a change to the most important variables in equations (5 and 6). Both equations (5 and 6) showed high correlation coefficients (R-value>0.7) but were not statistically significant (p>0.05). Of the top 5 variables that contribute to equation 5 (without outliers), three appear to be student centred (coming from personal factors and during year preparation sections), while two relate to teaching staff support. In parallel, including all students in equation 6, the 5 top variables shift to 2 relating to teaching staff support, 2 student centred variables (personal factors and during year preparation sections) and 1 environment and cohort factors (Table 6).

Table 5: The generated equations from simple linear regression of student perception individual variables XX1 to XX20 and the X variable value from layer 1 modelling with (N = 25) or without outliers (N = 24).

Area of variables	**Description of Variable (Variable symbol)**	**Weighted value of the variable in the equation**	
		Without Outliers (equation 5)	**With Outliers (equation 6)**
Intercept		4.629	5.013
Personal factors	Interest XX1	**-3.832**	0.204
	Motivation XX2	0.426	**-1.739**
	Sleep Before Exam XX3	-0.164	0.032

Environment and cohort factors	Peers' Communication XX4	**1.749**	0.228
	University Learning Facilities XX5	0.816	**0.942**
Perception of module importance	Fore Degree XX6	0.648	0.108
	For Future XX7	1.314	-0.437
	Difficulty XX8	0.213	-0.027
Teaching staff related factors	**Guidance XX9**	**-1.503**	**0.739**
	Feedback XX10	**-1.883**	-0.437
	Design XX11	**-2.044**	**-1.200**
During year preparation	**Self-Study XX12**	**-2.269**	**0.656**
	Notes Organization XX13	**3.849**	0.272
	Up to date Studying XX14	-0.162	**0.720**
	Study Plan XX15	-1.397	**-0.632**
Close to exam preparation	Enough preparation time Close to Exam XX16	0.973	-0.441
	Practicing Quiz XX17	1.199	0.232
	Practicing Exam XX18	0.530	-0.08
	Understand Not Memorise XX19	1.411	0.522
	Studying hours XX20	-0.009	0.014
Simple linear regression equations metrics			
Equation number		5	6
Number of observations		24	25

Error degrees of freedom	3	4
Root Mean Squared Error	0.89	1.31
Pearson Correlation Coefficient (R-value) R^2 Adjusted R^2	0.93 0.87 -0.01	0.79 0.62 -1.26
p-value	0.59	0.96

Table 6: Summary of the five top important variables contributing to layer 2 modelling of student perception individual variables XX1 to XX20 and the X variable value from layer 1 modelling with (N = 25) or without outliers (N = 24).

Top five variables in terms of weight to the layer 2 modelling		
Area of variables	**Without outlier (rank from equation 5)**	**With outlier (rank from equation 6)**
Personal factors	Interest XX1(2)	Motivation XX2 (1)
Environment and cohort factors	--	University Leaning Facilities XX5 (3)
Teaching staff related factors	Design XX11 (4) Feedback XX10 (5)	Design XX11(2) Guidance XX9 (4)
During year preparation	Notes Organization XX13 (1) Self-Study XX12 (3)	Up to date Studying XX14 (5)

4. Discussion

Most VLEs provide engagement indicators that count participation times, days since last log in and page views. However, these indicators are significantly dependent on how the VLE environment is designed. Some modules do not provide many chances for participation apart from simple assignment submission and others do not encompass tasks and activities that require students to frequently visit the VLE. In such scenarios, students log in once a week to collect all the learning materials for the whole week rather than collecting them gradually over the whole period of the course. This reduces applicability, creditability, and versatility of using participation times and days since log in as valid indicators of engagement and diminishes their application to the way the module is designed. On the other hand, page views are not much affected by course design or by a list of activities, as engaged students tend to spend more time on the VLE by browsing more pages than disengaged students. The highest correlation coefficient among VLE indictors was observed with pages views in report 2, (F) equals 0.552 (Table 1). This validates previous studies (Hatahet et al., 2022), where page views scored the highest correlation coefficient compared to all other VLE engagement indicators (R=0.618) (Hatahet et al., 2022). Furthermore, page views were independently studied and showed correlation with final exam marks in others degrees like social sciences (R = 0.35) (Hussain et al., 2018) and education (R = 0.299) (Mogus et al., 2012).

Generated models of VLE engagement indicators with coursework and practicals marks showed the important roles of these indicators in equations 1 and 3 (Table 2 and 3). These also showed that VLE engagement indicators collected at later stages of the year are more relevant to final exam mark predications. This validates previously reported models (Hatahet et al., 2022). The introduction of the X variable (classified as personal-related factors) again seemed to be of the greatest importance in predicting final exam mark, an observation previously found (Table 2 and 3 equations 2 and 4) (Hatahet et al., 2022). In summary, variables that serve as indicators for good exam performance, irrespective of module structure and student cohort, are: (i) page views (F) as a direct one-to-one indicator of good exam performance; and (ii) coursework (H), practical marks (G) and days since last log in (E) collected at late stages of the year. When using holistic modelling,

(iii) personal related factors (the X variable) remain the most significant factor in predicting good exam performance. These findings match previous research into a different cohort of students studying a different module, thereby suggesting the universality of those indicators (Hatahet et al., 2022).

Looking into the nature of the X variable, 20 questions were presented to students and modelled (layer 2) for matching X variable value coming from layer 1 modelling (Figure 1 and Table 4). 6 out of 20 variables (XXs) showed R values of > 0.2 (Cohen et al., 2013). Out of these six, 3 were related to teaching- staff related factors (XX9-XX11), and 2 were about having a study plan (XX15) and being able to self-study (XX12). The last variable was how students perceived the importance of the module for their degree (XX6), which can be linked to students' strategic approach to studying the module (Duff, 2003). It is worth noting that when VLE engagement, and coursework marks are high while the final exam mark is low, a low X variable is needed to compensate this difference in layer 1 modelling. Where negative correlations are observed in Table 4, this is due to students who scored low in the exam, needed a low X variable value but at the same time answered, 'strongly agree' (5) on questions related to the amount of support needed from teaching staff (XX9, 10 and 11). As a result, the lower their exam mark, the lower X and the more support they need from teaching staff.

With regards to variables with > 0.2 correlation individually, the results indicate students to be using the strategic learning approach. This is evident from the high correlation between student's perception of the module's importance to their degree (XX6) and the individual X variable, and hence exam mark (Duff, 2003). Strategic learning is a process in which students may adopt either or both the deep or surface learning approaches in order to achieve high marks (Fry et al., 2008). Furthermore, due to the nature of a remote learning setting, greater importance attaches to teaching staff guidance (XX9), feedback (XX10) and design of learning (XX11) in order to maintain students' engagement with learning materials as a precursor to and high achievement. This accords with the empirical evidence gathered by Chen and Jang highlighting the mediating effect of contextual support and motivation/self-determination in online learning settings (Chen & Jang, 2010). Wang *et al* reported teacher feedback relating to student self-regulation, and emotional feedback that helps students with motivation, to be essential to mitigate the loneliness and isolation associated with online learning (Wang et al., 2021). On the other side, having a study plan (XX15) and being able to self-study (XX12) i.e., having the skill of autonomy, is also crucial for success in remote learning settings. Chen *et al* found perceived autonomy to be the significant factor driving intrinsic and extrinsic motivation in a study of 276 students enrolled in a special education online certificate (Chen et al., 2010). This can be linked to the work of Poropat, which reports a sample weighted correlation corrected for scale reliability, of 0.23 between conscientiousness and academic performance, where conscientious students are defined as being those with higher self-controlled behaviour (Duckworth *et a*l.), an argument mostly valid for children but still applicable in adults in remote learning settings (Duckworth et al., 2012; Poropat, 2009).

Modelling student answers to the questionnaire was carried out to highlight the most important contributors to exam success from a student's perspective (Table 5 equations 5 and 6). Results indicated that outlier students have a significant impact on the order of importance of XXs variables. Outlier students are those with high exam marks, high coursework, and practical marks but at the same time low VLE engagement indicators in layer 1 modelling. Furthermore, outlier students reported lower scores compared to other students in the questionnaire (layer 2 modelling). In terms of personal related factors (Table 5), removing outliers shifted the equation from a higher importance for motivation (XX2) to interest (XX1). This shows that some students (outliers) are not very autonomous in a remote learning setting and as a result, the main driving force for them to keep going with their studies is to build on extrinsic motivation rather than simply being driven by personal or intrinsic interests (Chiu, 2022; Hoffmann et al., 1998; Shah et al., 2021). Again, in the environmental and cohort factors, outliers apparently are the shy students that rely more on university facilities such as libraries, computer rooms and self-study rooms (XX5) rather than peer communication (XX4) to drive their studies. The sense of relatedness and affiliation observed here is similar to Seiver and Troja's experience with psychology students where those with a high need-for-affiliation preferred enrolment in face-to-face rather than online learning (Seiver & Troja, 2014). A key distinction was also observed during the in-year preparation questions, where self-study (XX12) and study plan (XX15) remained important variables with and without outlier inclusion, but where notes organisation (XX13) (which was important in the cohort of students without outliers is replaced by up-to-date studying (XX14) when outliers are included (Table 5). This observation fits with Constructivism learning theory that engaged students will take notes but will face the challenges of how to arrange these meaningfully (Bada & Olusegun, 2015). In parallel, disengaged students' main challenge is how to keep themselves up to date with their timetable and studies. This can be linked to the observed shift from interest (XX1) as a significant factor with engaged students, to motivation (XX2) with disengaged students (Equations 5 v.s 6).

Summarising the top 5 contributing variables for equation 5 and 6 (Table 6) show that when both engaged and disengaged students are included in modelling, the role of the environment and cohort on students' learning experience is emphasised. There is a missing attribute in remote learning as students are studying individually rather than gathering in a classroom to study as a cohort (Seiver & Troja, 2014). Dabbagh and Bannan-Ritland identified the absence of face-to-face interaction with peers and teachers to be a major reason for loss of student motivation (Dabbagh & Bannan-

Ritland, 2005). The current study in remote learning identifies two distinct groups of remote learners. The first group of students were able to adapt (equation 5) to remote learning. They were interested (XX1), able to make a study plan (XX12), and their learning involves organising study notes (XX13). Therefore, these students require educators to design good learning structure (XX11) and to provide constructive feedback (XX10). Good support during remote learning was found to be key to the success of a K-12 class of children in the USA during the COVID-19 pandemic, with the authors reporting that tackling the complex relationship between the self-control to enable freedom in learning and the cognitive need for organisation of leaning resources to be essential in course design to ensure successful online learning outcomes (Carter Jr et al., 2020). The second group of students were those who experience difficulties in learning individually and prefer the social aspect of the learning process. In this group of students, factors shifted in order of importance during remote learning to equation 6. Looking at the whole cohort of students, it is noted that students with remote learning engagement difficulties were working hard to keep themselves motivated (XX2), were struggling to keep themselves up to date with studying (XX14), relied on university learning facilities (XX5) and the way teachers design learning (XX11), and required guidance (XX9) more than feedback to keep on top of their studies.

While both Equations 5 and 6 (Table 5) showed high Pearson corelation coefficients (0.79 and 0.93, respectively), neither of these was statistically significant. Statistically, no significant relationship was found between the variables XX and response variable X. This may have been due to the small sample size (24 and 25 students without and with outliers, respectively). Therefore, some variables may have to be removed from the modelling as they do not have a significant relation to the response variable X (Khot, 2020). Furthermore, adjusted R^2 indicates parsimony. A negative, adjusted R^2 shows the model to be over parametrised, indicating that there are too many variables for the sample size (Tabachnick et al., 2007). For that reason, a simple linear regression was carried out using only the top 5 variables in equations 5 and 6 to generate equations 7 and 8 respectively (supplementary Table 2). The resulting models showed higher Pearson correlation coefficients, and positive adjusted R^2 both of which were statistically significant ($p < 0.05$).

In terms of limitations, the current research has been undertaken using only one cohort of students and the sample size is small to draw generalised conclusions. The work has also been limited to the modelling of student survey responses, with there being no opportunity to conduct focus groups with outlier students. The data gathered in this work will be used in the design of further investigative research into the personal factors identified, involving focus groups and larger sample size by gathering data form multiple cohorts across multiple modules. Based on this research, some recommendations for remote learning designers are outlined in figure 3.

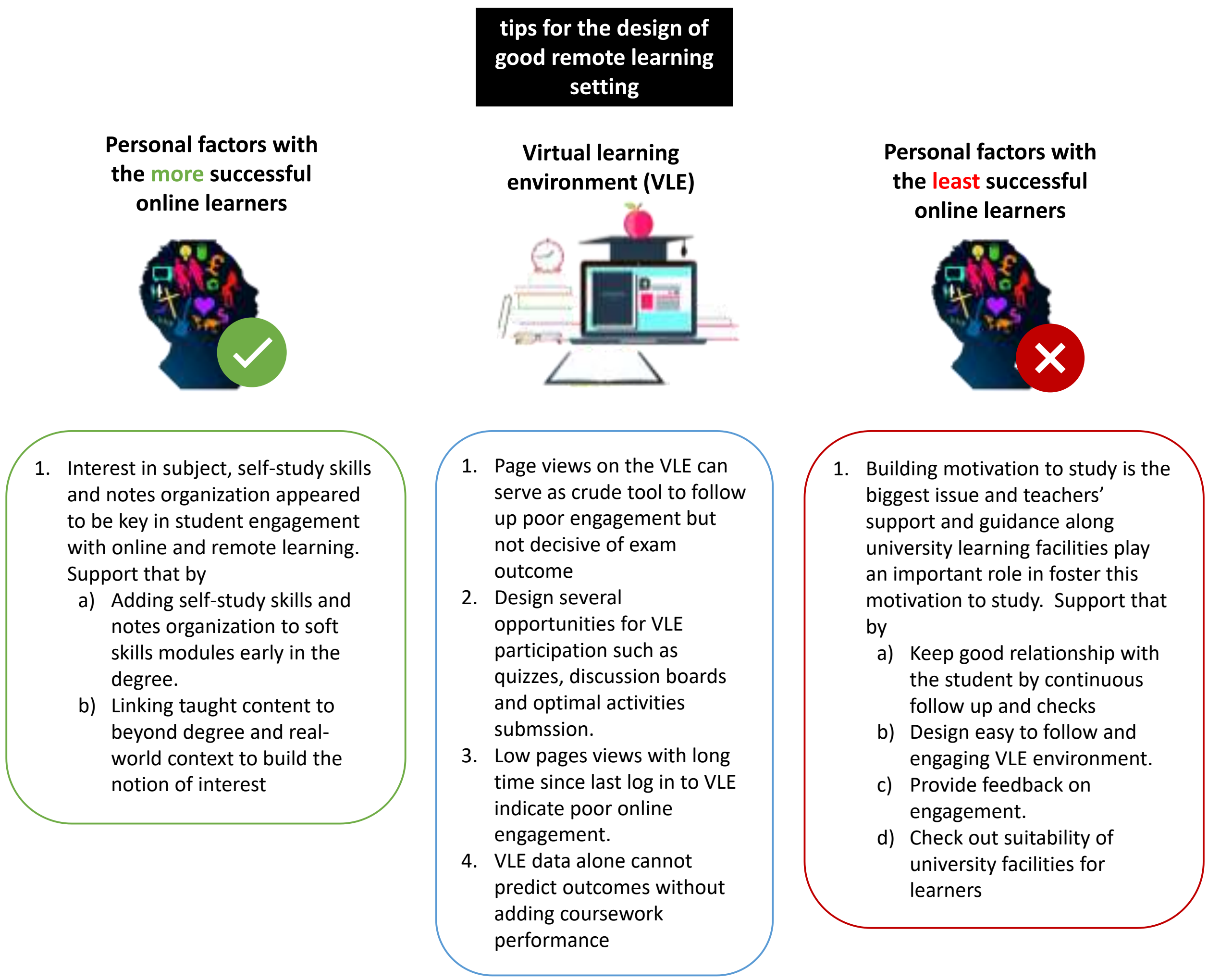


Figure 3: Recommendation for remote learning designers based on modelling results.

5. Conclusions

This study investigated the role of VLE engagement tools, student practical and coursework marks, and student personal factors (X) as predictors of exam performance. In one-to-one correlations, VLE page views – with a correlation coefficient > 0.5 - seemed to be the best VLE related predictor of exam performance, regardless of module design, student cohort or degree programme. Using artificial intelligence to model VLE engagement tools and practical and coursework marks against exam performance resulted in the generation of linear regression models with Pearson correlation coefficients > 0.59. Confirming the findings of previous research (involving a different cohort of students enrolled in a different module and degree programme), the findings showed days since last log in, and practical and coursework marks to be the most important (identified) predictors of good exam performance. An unknown 'X' variable - composed of personal factors and individual differences - was, however, found to outweigh all other variables in importance.

Through modelling student responses to a survey designed to investigate individual differences and personal factors which may be relevant to student engagement, we were able to identify a cohort of engaged students (with remote learning). These individuals were found to be driven by intrinsic subject interest, to have good self-study skills, and be able to organise their notes, to benefit from good teacher design of learning, and to welcome feedback as the key factors for success in their learning journey. On the other hand, including all students - engaged and less engaged - in the modelling revealed that motivation is a key factor in keeping students engaged and up to date with studies. Less engaged students rely more on good teacher design of learning, university facilities and guidance on learning rather than feedback.

Overall, the findings highlight the key role that must be played by the teacher in the design of VLEs and flag up areas of personal differences that can be targeted through extracurricular and personal development schemes to support students studying remotely. As such, they may be of great use for universities and academics seeking to maximise student satisfaction in any remote learning context.

Acknowledgment

The authors did not receive any financial support for this study.

Conflict of interest

The authors declare no conflict of interest.

References

Alzubaidi, H., Jirjees, F. J., Franson, K. L., Saidawi, W., Othman, A. M., Rabeeah, Z. H., Al-Obaidi, H. J., Hatahet, T., & Caliph, S. (2021). A global assessment of distance pharmacy education amid COVID-19: teaching, assessment and experiential training. *International Journal of Pharmacy Practice*. https://doi.org/10.1093/ijpp/riab064

Audet, C., & Hare, W. (2017). *Derivative-free and blackbox optimization*.

Ayres, P., & Paas, F. (2009). Interdisciplinary perspectives inspiring a new generation of cognitive load research. *Educational Psychology Review, 21*(1), 1–9.

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, *5*(6), 66–70.

Barbera, J., Naibert, N., Komperda, R., & Pentecost, T. C. (2020). Clarity on Cronbach's Alpha Use. *Journal of Chemical Education*, *98*(2), 257–258. https://doi.org/10.1021/ACS.JCHEMED.0C00183

Carter Jr, R. A., Rice, M., Yang, S., & Jackson, H. A. (2020). Self-regulated learning in online learning environments: strategies for remote learning. *Information and Learning Sciences*.

Chen, K.-C., & Jang, S.-J. (2010). Motivation in online learning: Testing a model of self-determination theory. *Computers in Human Behavior*, *26*(4), 741–752. https://doi.org/https://doi.org/10.1016/j.chb.2010.01.011

Chen, K.-C., Jang, S.-J., & Branch, R. M. (2010). Autonomy, affiliation, and ability: Relative salience of factors that influence online learner motivation and learning outcomes. *Knowledge Management & E-Learning: An International Journal*, *2*(1), 30–50.

Chiu, T. K. F. (2022). Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic. *Journal of Research on Technology in Education*, *54*(sup1), S14–S30.

Clerkin, C., Hatahet, T., Malekigorji, M., & Andrews, G. P. (2022). *Chinese Students' Perception and Expectation of Online and Post-Pandemic Teaching and Learning Approaches in a UK Transnational Program*. https://doi.org/10.3390/educsci12110761

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. routledge.

Dabbagh, N., & Bannan-Ritland, B. (2005). *Online learning: Concepts, strategies, and application*. Prentice Hall.

Deci, E. L., & Ryan, R. M. (1994). Promoting self-determined education. *Scandinavian Journal of Educational Research, 38*(1), 3–14.

Deci, E. L., & Ryan, R. M. (1995). *Human Autonomy BT - Efficacy, Agency, and Self-Esteem* (M. H. Kernis (Ed.); pp. 31–49). Springer US. https://doi.org/10.1007/978-1-4899-1280-0_3

Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology, 104*(2), 439.

Duff, A. (2003). Quality of learning on an MBA programme: The impact of approaches to learning on academic performance. *Educational Psychology, 23*(2), 123–139.

Fry, H., Ketteridge, S., & Marshall, S. (2008). *A handbook for teaching and learning in higher education: Enhancing academic practice*. Routledge.

Hatahet, T., Mohamed, A. A. R., Malekigorji, M., & Kerry, E. K. (2022). Remote Learning in Transnational Education: Relationship between Virtual Learning Engagement and Student Academic Performance in BSc Pharmaceutical Biotechnology. In *Pharmacy* (Vol. 10, Issue 1). https://doi.org/10.3390/pharmacy10010004

Healey, N. M. (2020). The end of transnational education? The view from the UK. *Perspectives: Policy and Practice in Higher Education, 24*(3), 102–112. https://doi.org/10.1080/13603108.2019.1631227

Heitz, C., Laboissiere, M., Sanghvi, S., & Sarakatsannis, J. (2020). Getting the next phase of remote learning right in higher education. *McKinsey Insights*.

Hoffmann, L., Baumert, J., Krapp, A., & Renninger, K. (1998). *Interest and learning: Proceedings of the Seeon Conference on Interest and Gender*.

Hussain, M., Zhu, W., Zhang, W., & Abidi, S. M. R. (2018). Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores. *Computational Intelligence and Neuroscience, 2018*. https://doi.org/10.1155/2018/6347186

Jeffries, P. R., Bushardt, R. L., DuBose-Morris, R., Hood, C., Kardong-Edgren, S., Pintz, C., Posey, L., & Sikka, N. (2022). The role of technology in health professions education during the COVID-19 pandemic. *Academic Medicine, 97*(3), S104.

Khot, S. (2020). *Interaction effect in multiple regression*. https://towardsdatascience.com/interaction-effect-in-multiple-regression-3091a5d0fadd

Krzyżak, J., & Walas-Trebacz, J. (2021). Objective Factors Determining The Development of Online Self-study Skills During The Global COVID-19 Pandemic-from The Perspective of Students' Experiences. *Proceedings of the 38th International Business Information Management Association (IBIMA)*, 1207–1216.

MathWorks - Makers of MATLAB and Simulink - MATLAB & Simulink. (n.d.). Retrieved October 28, 2021, from https://uk.mathworks.com/

Mogus, A. M., Djurdjevic, I., & Suvak, N. (2012). The impact of student activity in a virtual learning environment on their final mark: *Http://Dx.Doi.Org/10.1177/1469787412452985*, *13*(3), 177–189. https://doi.org/10.1177/1469787412452985

Müller, F. H., & Louw, J. (2004). Learning environment, motivation and interest: Perspectives on self-determination theory. *South African Journal of Psychology*, *34*(2), 169–190.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322.

Richardson, J. T. E. (2005). Students' perceptions of academic quality and approaches to studying in distance education. *British Educational Research Journal*, *31*(1), 7–27.

Seiver, J. G., & Troja, A. (2014). Satisfaction and success in online learning as a function of the needs for affiliation, autonomy, and mastery. *Distance Education*, *35*(1), 90–105. https://doi.org/10.1080/01587919.2014.891427

Shah, S. S., Shah, A. A., Memon, F., Kemal, A. A., & Soomro, A. (2021). Online learning during the COVID-19 pandemic: Applying the self-determination theory in the 'new normal.' *Revista de Psicodidáctica (English Ed.)*, *26*(2), 168–177.

Sharma, D. D., & Johanson, J. (1987). Technical consultancy in internationalisation. *International Marketing Review*.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). pearson Boston, MA.

Tsiligiris, V., Ilieva, J., & Hill, C. (2018). *Transnational education: global location, local innovation: a report produced for UUKi*.

Walfish, S. (2006). A review of statistical outlier methods. *Pharmaceutical Technology*, *30*(11), 82.

Wang, H., Tlili, A., Lehman, J. D., Lu, H., & Huang, R. (2021). Investigating feedback implemented by instructors to support online competency-based learning (CBL): a multiple case study. *International Journal of Educational Technology in Higher Education*, *18*(1), 5. https://doi.org/10.1186/s41239-021-00241-6

Weller, M. (2007). *Virtual learning environments: Using, choosing and developing your VLE*. Routledge.

Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, *4*(2), 65–85.