Proxy Variables of Online Engagement on the Learning Management System

Sylvia Chong and Adam Wong
Singapore University of Social Sciences, Singapore
Email: sylviachong@suss.edu.sg

***Abstract*—The wide-spread adoption of Learning Management System (LMS) technology has fundamentally changed the environment for today's teaching and learning. The LMS has generated a new set of data that could possibly serve as proxy measures of learners' engagement online. LMS data sets can be mined and analyzed to provide meaningful measures of learners' online engagement to support and enhance teaching and learning practices. In particular, this research focuses on the digital traces of LMS data to establish and validate meaningful proxies of online engagement. With the vast data sets available over a LMS, it is timely to identify a methodical approach to extract and transform data features in a modus that is useful for analysis. Hence, the purpose of this study is twofold: (1) to establish a common set of metrics (immediacy, frequency, interval and recency) to codify and quantify learners' online engagement in the LMS; and (2) to validate, using data mining techniques, these metrics of online engagement patterns in relation to the learners' academic performance.**

***Index Terms*—Learning management system, online engagement**

I. Introduction

With the advancement of learning technologies and as teaching and learning become increasingly omni-channel and digitalized, the online learning platform, usually in a form of a learning management system (LMS), becomes an important touchpoint where students and educators engage to exchange information and knowledge. The wide-spread adoption of LMS technology has fundamentally changed the learning environment. Data fed through this LMS becomes a record of any online activity associated with student-to-student, student-to-content, and student-to-instructor interactions. In this teaching and learning environment, a key part of the learning experience of a student transpires into a series of mouse clicks, navigations, and interactions within the LMS.

The study of student behaviour through the analysis of the digital trails that students leave as they navigate through the digital space – how frequently and when they log on, where and what digital resources they use, will become increasingly central to learning support, curriculum design, assessment and quality assurance (Hall, 2015). While some universities have developed dashboards to present some degree of information from the raw data, most have suggested that there is an interest in including a metric around online engagement but it has proved to be challenging so far to develop anything meaningful from the click-data (Sclater, 2014). More importantly, it is in the interest of educational institutions to make better sense of this data source; and with it, the possibilities of aggregating students' online activities and analyzing them in a more comprehensible form to better inform teaching and learning.

There is a need to develop a more robust concept of student engagement in a virtual learning environment. This study aims to establish a set of features engineered from the LMS click-data to represent engagement characteristics observed from student online behaviour. This set of features build on the metrics profile established in an earlier study [1].

II. Literature Review

A. Student Engagement

The concept of student engagement was developed from Kuh's 2009 study on the National Survey of Student Engagement (NSSE) [2]. NSSE was based on the concepts of Chickering and Gamson's [3] "Seven Principles of Good Practice in Higher Education", Astin's [4] concept of student's involvement as well as Pace's [5] notion of quality of effort. Kuh's study combined the three concepts and defined student engagement as a students' involvement as well the amount of invested effort by the student [6]. Student learning engagement has become an important area of educational research in the past decade because of its correlation to student outcomes. Measures of student engagement include behavioural, psychological and cognitive variables such as class attendance, class participation, retention rates, students' perceptions of their learning in relation to their peers, faculty members, and institutions; and course or assignment grades and term grade point averages [7].

In a 2014 study, 85% of students reported the use of LMS in at least one course and 56% of students reported using an LMS in most or all of their courses [8]. This adoption of LMS has generated a new set of data that could possibly serve as proxy measures of student engagement online. On the LMS, student engagement

Manuscript received May 1, 2019; revised September 1, 2019.

doi: 10.18178/ijlt.5.4.296-300

such as logging into the LMS, viewing lesson materials, participating online activities, contributing to discussion blogs, and submitting assignments online can be recorded. These data records are potential proxies of measures of behavioural engagement (time spent logged on, materials viewed), psychological engagement (frequency of participation online, network analysis of discussion posts, number of emails sent and received), and cognitive engagement (number of attempts, content analysis of online discussions). Trowler [9] described student engagement as “the investment of time, effort and other relevant resources by both students and their institutions intended to optimise the student experience and enhance the learning outcomes and development of students, and the performance and reputation of the institution” (p. 3).

B. *Learning Management Systems (LMS)*

The campus learning environment is profoundly impacted by the pervasive dependence on the internet. Students are continually connected via smart phone, tablet, or laptop. Ownership of internet-enabled devices is high, with over three quarters of students owning a smartphone and six out of ten students owning three or more Internet capable devices [8]. Being connected to the Internet at all times, students have immediate and continuous access to social media, library databases, their institution’s Learning Management System (LMS). This infusion of technology into the campus-based student environment has several implications, including an increased number of student-student and student-faculty communication channels, opportunities to continue discussions from class into an online environment, and the chance to review course materials from any location at any time [10]

Research studies have established a positive relationship between the use LMS use and students’ grades. Dawson and McWilliam’s [11] study concluded that “there is a greater likelihood of achieving better grades if one actively and productively engages with forum activities” (p. 29). Pulford [12] reported that students using the Blackboard to read discussion posts and ask questions of tutors performed significantly better than those who did not participate in the discussion forums available. Gašević, Dawson, and Siemens [13] studied nine Australian undergraduate courses in an Australian college and found the number of logins and views in discussion forums to be significant predictors of academic performance.

Dixson [14] conducted an Online Student Engagement Survey (OSES) comparing groups of students based on self-reported LMS activity, which she categorized as active or passive. While there were no difference in LMS activity between those with high engagement scores and those who had low scores, those who reported have more discussions on the LMS had higher engagement scores. Hamane [15] adapted Dixson’s OSES and found a weak but positive relationship between students’ perceived level of engagement and students’ frequency of logins. Wong and Chong [1] identified patterns of online engagement with positive correlation to adult learners’ academic grades.

III. Purpose of Study

The purpose of the study is two-fold.

1. To establish a common set of metrics (immediacy, frequency, interval and recency) to codify and quantify learners’ online engagement in the LMS; and
2. To validate, using data mining techniques, these metrics of online engagement patterns in relation to the learners’ academic performance.

IV. Methodology

A. *Dataset*

Data for this study came from three undergraduate courses offered during the January 2018 semester at a University in Singapore for adult learners. Learning resources such as course notes, seminar materials, activities, chunked lessons, video recordings, exam revision materials, pre-class quizzes, online discussions, case studies, and assessments are presented on the LMS. Learners are expected to be self-regulated in their own learning and complete majority of the coursework online. There are three face-to-face learning sessions. Students (N = 1187) enrolled in these courses had a total of 253,927 instances of access, averaging about 214 accesses between them.

B. *Feature Extraction*

To illustrate the derivation of the data features, the metrics can be simplified to focus on the entire course; that is, there is no separation of learning resources although differentiation can be made to represent student behaviour on multiple learning resources. This is summarised in Fig. 1.

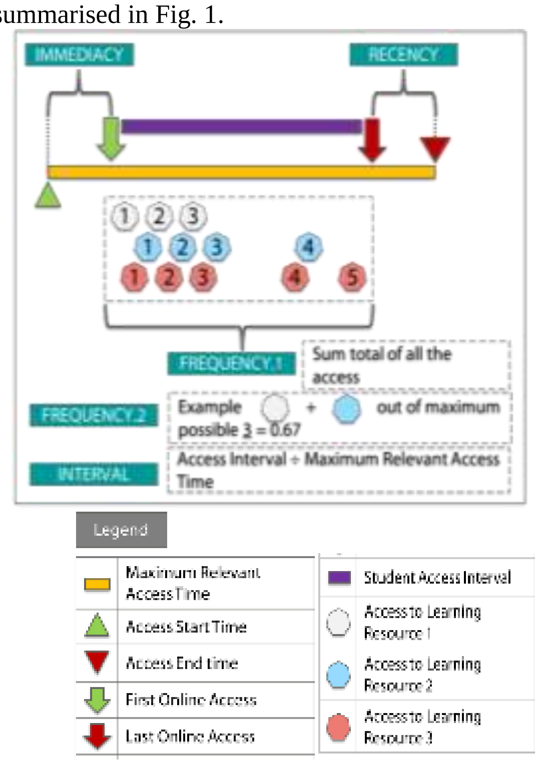

Figure 1. Extraction of the engagement metrics (immediacy, frequency.1, frequency.2, interval, recency)

The five indicators of online engagement proposed in this study is presented in Table I. Data preparation procedures will be discussed in the later sections.

TABLE I. ONLINE ENGAGEMENT METRICS — IMMEDIACY, FREQUENCY.1, FREQUENCY.2, INTERVAL AND RECENCY

Online engagement metric	Description
Immediacy	Measures the time lapse between the access start time and the start of the first online access
Frequency.1	Measures the number of episodes of online access over the access interval within a relevant given period
Frequency.2	Measures the proportion of distinct online access to a relevant collection of online learning resource
Interval	Measures the proportion of access interval relative to the maximum relevant access time
Recency	Measures the time lapse between the start of the last online access and the access end time

C. *Data Transformation*

To facilitate downstream analysis and in the leading steps to computing an overall index metric, two additional data preparation procedures are performed on the original metric values. The following formula is applied to normalize each metric:

Normalized Metric Value = (Original Metric Value – Min) / (Max – Min)

To set the metrics in the same and right direction, the following formula is applied to immediacy and recency:

Reverse-Scored Normalized Metric Value = 1 – Normalized Metric Value

In place of the metric in its original or normalized and reverse-scored metric value, additional procedures are applied to summarise a student's relative engagement behaviour in banded groupings based on the values of the mean and standard deviation of the distribution of the metric values. For example, selecting a +/-2 mean/standard deviation as the binning method results in five brackets ("buckets"). With this, a student's engagement behaviour can be explained in a relative term to his course cohort (i.e. low LMS engagement (↓), moderate LMS engagement (—), or high LMS engagement (↑)).

The overall index metric (or LMS engagement score) is a linear combination of the normalized and reverse-scored normalized metric values by multiplying each component metrics with a constant. A constant is set to 1 for all component metrics to differentiated weights to reflect the importance of the component metrics.

D. *Exploratory Data Analysis*

The analysis of these metrics represent proxies of engagement online: *Immediacy* - their sense of urgency or excitement; *Frequency 1* - how frequently do they access the digital resources; *Frequency 2* - what digital resources they use and how much, *Interval* - how much time they put in relative to the maximum possible access time given, and *Recency* – their sustained efforts for the most recent access prior to access end time (or an end-of-course assessment). The association between learners' LMS engagement level and the consequences of learning (their course grades), is analysed using curve estimation and analysis of variance (ANOVA) (see Table II).

TABLE II: ONE-WAY ANALYSIS OF VARIANCE OF GRADES BY LEVEL OF LMS ENGAGEMENT

		df	Mean Square	F	Sig.
COURSE A	Between Groups	4	3801.460	22.541	0.000
	Within Groups	382	168.650		
	Total	386			
COURSE B	Between Groups	4	3630.436	19.894	0.000
	Within Groups	574	182.489		
	Total	578			
COURSE C	Between Groups	4	1154.854	7.948	0.000
	Within Groups	197	145.304		
	Total	201			

V. RESULTS AND DISCUSSIONS

An understanding of the LMS engagement behaviour which proceeds from observations is predicated by interpreting the LMS engagement metric value. It is possible to differentiate LMS engagement behaviour by separating the interpretation of each metric. However, since the LMS engagement score is an aggregation of each metric (such that a higher index value reflects a higher level of LMS engagement), we shall only infer LMS engagement behaviour from the LMS engagement score to make conclusions about the study population in simplest terms.

TABLE III. ONLINE ENGAGEMENT SCORE

Course	N	Range	Min	Max	Mean	Std D
A	388	4.12	0.88	5.00	3.22	0.70
B	579	3.76	1.01	4.77	3.38	0.61
C	202	3.77	0.98	4.75	3.43	0.57

Based on the summary online engagement scores (Table III) from the LMS engagement behaviour of students (N = 1187) of the three courses, the mean of the overall index metric stood between 3.22 and 3.38. LMS engagement score varies between 0.88 and 5.00 for Course A, 1.01 and 4.77 for Course B, 0.98 and 4.75 for Course C. The reported means have a SD of 0.70 (or 3.22 ± 0.70), 0.60 (or 3.38 ± 0.60) and 0.57 (or 3.43 ± 0.57) respectively.

A. *Correlation between Student Performance and LMS Engagement*

More importantly, the study is interested in the relationship between course achievement and LMS engagement. Curve estimation can be used to find the best fit to the data points and to summarize the relationships between two variables of interest [16]. From the resulting curve-fitting model and with the aid of a scatterplot (where each student's LMS engagement score (x-axis) is plotted against their respective course grades (y-axis), the study is able to examine the relationship (the direction) of course achievement and LMS engagement behaviour (Fig. 2). Without any knowledge of *a priori* variables, but hypothetically, a higher course grade is associated with a higher level of LMS engagement. From what is observable, *a posteriori*, conclusions about the two variables can be drawn from the analysis. There is a positive correlation between course achievement and level of LMS engagement. This behaviour was also consistent across three courses.

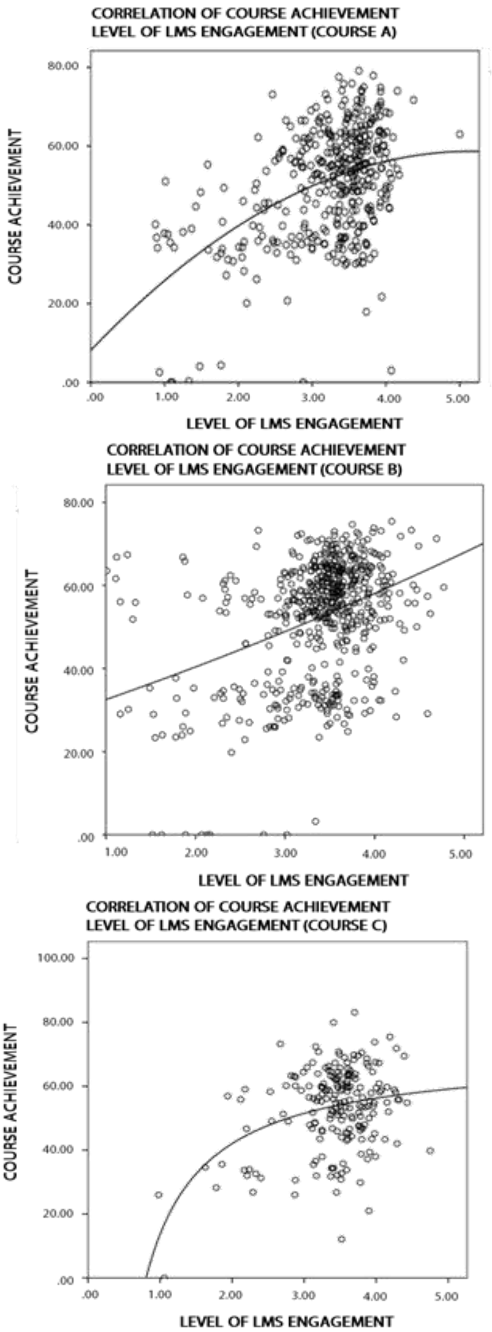

Figure 2. Scatterplots showing a curvilinear relationship between student's course achievement and LMS engagement score

Information gathered from curve estimation allows us to infer that an increase in LMS engagement is likely to have positive effect on course grades.

In order to further assess the effect on performance. A one-way ANOVA was conducted to compare the effect on course achievement for different levels of LMS engagement. Prior to performing the ANOVA, the LMS engagement scores were grouped into bands determined by a +/-2 mean/standard deviation binning method – such that the cut-offs for each band are determined by how much they deviate from the mean. The results suggested that there is a significant effect on course achievement at $p<.05$ level for five levels of LMS engagement for Course A [$F(4, 382) = 22.54$, $p = 0.000$*]. Similar significant effect is being observed for Course B [$F(4, 574) = 19.89$, $p = 0.000$*] and Course C [$F(4, 197) = 7.95$, $p = 0.000$*].

Taken together, these results suggest that high levels of LMS engagement do have an effect on student performance. More importantly, identifying the determinants that influence academic performance is an essential part of educational research. The correlations between performance and all the LMS behavioural features (aggregated to an overall index metric), build a case that factors influencing academic achievement are not limited to prior academic ability, demographics traits, and pre-admission indicators – in particular for adult learners at the tertiary level.

VI. Future Work & Limitations

This paper has proposed possible component metrics for estimating LMS behaviour, and also discussed its effect on performance with three courses from different disciplines. This is especially significant with the number of degree-seeking adult learners on the rise. With wide spread adoption of the LMS, education institutions need to strategize how to best it can support these learners and optimise teaching and learning.

Future studies can use these metrics to measure and attach significance to different online learning activities in relation to the degree of student engagement and its effect on performance. There is currently no quantification of the interplay of these determinants, particularly so for the research of the antecedents, behavioural, and the consequents of learning for adult learners. And this warrants further studies. Multivariate analysis can be performed using machine learning algorithms such as decision trees to create informative models. The insights gathered from such deeper analysis are expected to reinforce decisions around course design, instructional methods, learning support and student advisory. This is a natural progression for learning analytics – that is, from the extraction of data features to the modelling of it.

The key purpose of the study is to provide advisory intervention based on the online learning engagement of adult learners from the use of proxy variables. Future research should focus on the development and deployment of such learning interventions. A monitoring interface which employs the proxy variables can be embedded in the LMS so that instructors and/or learners are able to monitor their learning engagement status with a view for formative assessment view. This will enable reflection on different levels of online learning engagement to improve learning outcomes. This can be extended to providing more focussed support and interventions for the continuous development of learners online.

This study does come with several limitations. Nonetheless, addressing the limitations would also set the directions for future research. The limitations are set in mostly within the difficulties of extracting useful data features from ‘big data’ – that is, limited both by the volume of data and the pre-processing of LMS data. As such, the analysis presented in this paper is performed largely ex-post. That is after a course has ended. The total login time was used as an indicator of extensive online engagement, however the learners may have been distracted while they were online, this would result in lower learning performance compared to that indicated by the recorded time.

However, in terms of education research and practice, there is scope to look at the past to improve future learning. Addressing the limitations of data pre-processing (including data integration, engineering and automated feature extraction) will greatly benefit downstream analysis, and enable near real-time interventions that aim to help students while the course is still in progress. Overcoming these limitations will also lift the gates for progression to predictive and prescriptive analytics with LMS behavioural metrics, and pave the way for adaptive learning and automated advising systems. In line with this, future research can be directed to look to building up the metric profile with the inclusion of other informative metrics to learn other aspects of student online behaviour – such as regularity of effort from time-density based features, and peer effect from social network metrics.

ACKNOWLEDGMENT

This paper is part of an institutional research made possible through a Grant (Code: RF16IRA01) provided by the Singapore University of Social Sciences, Singapore.

REFERENCES

[1] W. Adam and S. Chong, “Modelling adult learners’ online engagement behaviour: proxy measures and its application,” *Journal of Computers in Education*, vol. 5, no. 4, pp. 463-479, 2018.

[2] D. K. George, “What student affairs professionals need to know about student engagement,” *Journal of college student development*, vol. 50, no. 6, pp. 683-706, 2009.

[3] W. C. Arthur and Z. F. Gamson, “Appendix A: Seven principles for good practice in undergraduate education,” *New Directions for Teaching and Learning*, no. 47, pp. 63-69, 1991.

[4] W. A. Alexander, “Student involvement: A developmental theory for higher education,” 1999.

[5] P. C. Robert, Measuring Outcomes of College. Fifty Years of Findings and Recommendations for the Future, 1979.

[6] F. Daniel, “Baccalaureate attainment of college students at 4-year institutions as a function of student engagement behaviors: Social and academic student engagement behaviors matter,” *Research in Higher Education*, vol. 55, no. 5, pp. 467-493, 2014.

[7] Z. Nick, P. Butler, and L. Leach, “Institutional research and improving the quality of student engagement,” *Quality in Higher Education*, vol. 18, no. 3, pp. 329-347, 2012.

[8] D. Eden, D. C. Brooks, and J. Bichsel, “The current ecosystem of learning management systems in higher education: Student, faculty, and IT perspectives,” 2014.

[9] T. Vicki, “Leadership practices for student engagement in challenging conditions,” *Perspectives: Policy and Practice in Higher Education*, vol. 17, no. 3, pp. 91-95, 2013.

[10] S. Rhona, H. Beetham, and S. de Freitas, *Rethinking Learning for a Digital age: How Learners are Shaping Their Own Experiences*, Routledge, 2010.

[11] D. Shane and E. McWilliam, “Investigating the application of IT generated data as an indicator of learning and teaching performance,” Queensland University of Technology and the University of British Columbia, 2008.

[12] D. P. Briony, “The influence of advice in a virtual learning environment,” *British Journal of Educational Technology*, vol. 42, no. 1, pp. 31-39, 2011.

[13] G. Dragan, S. Dawson, and G. Siemens, “Let’s not forget: Learning analytics are about learning,” *Tech. Trends*, vol. 59, no. 1, pp. 64-71, 2015.

[14] D. D. Marcia, “Creating effective student engagement in online courses: What do students find engaging?” *Journal of the Scholarship of Teaching and Learning*, pp. 1-13, 2010.

[15] C. H. Angelique, “Student engagement in an online course and its impact on student success,” Diss. Pepperdine University, 2014.

[16] S. H. Shivalingappa and K. V. Rao. *Advanced Techniques of Population Analysis*, Springer Science & Business Media, 2013.

[17] A. Cichocki and R. Unbehaven, *Neural Networks for Optimization and Signal Processing*, 1st ed. Chichester, U.K.: Wiley, 1993, ch. 2, pp. 45-47.

[18] W. K. Chen, *Linear Networks and Systems*, Belmont, CA: Wadsworth, 1993, pp. 123-135.

[19] H. Poor, *An Introduction to Signal Detection and Estimation*; New York: Springer-Verlag, 1985, ch. 4.

Sylvia Chong is an Associate Professor with the Business Intelligence & Analytics Unit, Singapore University of Social Sciences (SUSS). At SUSS she is the principal investigator for several institutional research projects.

Her research interests are inter‑disciplinary in nature and include both substantive and methodological approaches. The areas of interest include *Quality Management in Higher Education* (Educational Accountability, Teaching & Learning Quality), *Beliefs, identity and epistemology (*Self‑efficacy ‑ educators & learners; Beliefs about teaching and learning), *Evaluation* (Instrument development and validation; issues in tool development & administration; quantitative analysis & reasoning, multilevel modelling; qualitative coding & category systems) and *Learning analytics* (deciphering learning trends and patterns from educational data).