INTED 2019

13th International Technology, Education and Development Conference

11-13 March, 2019
Valencia (Spain)

CONFERENCE PROCEEDINGS

Exploring New Frontiers in Education

INTED 2019

13th **International Technology, Education** and **Development Conference**

11-13 March, 2019
Valencia (Spain)

CONFERENCE PROCEEDINGS

Published by
IATED Academy
iated.org

INTED2019 Proceedings
13th International Technology, Education and Development Conference
March 11th-13th, 2019 — Valencia, Spain

Edited by
L. Gómez Chova, A. López Martínez, I. Candel Torres
IATED Academy

ISBN: 978-84-09-08619-1
ISSN: 2340-1079
Depósito Legal: V-247-2019

Book cover designed by
J.L. Bernat

MOOCS LEARNING ASSESSMENT: CONCEPTUALISATION, OPERATIVISATION AND MEASUREMENT

Maria Carannante[1], Cristina Davino[2], Domenico Vistocco[3]

[1]Federica Weblearning Center for Innovation and Dissemination of Distance Education, University of Naples Federico II (ITALY)
[2]Department of Economics and Statistics, University of Naples Federico II (ITALY)
[3]Department of Political Science, University of Naples Federico II (ITALY)

Abstract

Massive Open Online Courses differ from traditional forms of learning, because of their free access, abundance of resources and of user interaction tools available. Notwithstanding these features, some learners reach their achievement while others do not. In learning analytics literature there are many contributing relating to how Massive Open Online Courses learners' behaviours determine performance. The paper proposes a study to face the conceptualisation, operativisation and measurement of learning and engagement, two determinists of learners' performance. Moreover, considering the multidimensional and complex nature of Massive Open Online Courses learning assessment, a Partial Least Squares model is also proposed.

Keywords: learning analytics, multivariate modelling, engagement, partial least squares path modelling.

1 INTRODUCTION

The Massive Open Online Courses (MOOCs) phenomenon emerged with much popularity in recent years so that MOOCs are becoming highly relevant and widespread learning tools in higher education institutes both in traditional and distance universities. MOOCs allow to offer free access to university courses to many users localized geographically all over the world and they provide a self-regulated learning (SRL). In online learning, SRL represents an opportunity for learners to have the control on various aspects of their own learning process [1]. The theoretical framework of SRL can be synthesised in the following four aspects: students have an active role to obtain their own knowledge; students can monitor, control and evaluate their cognition and motivation, they have the opportunity to adequate cognition to their goals and perspective and skills and they can interact with the context and other learners with the same characteristics. The interaction features of SRL systems are enriched by the wealth of learning tools offered by MOOCs which are "based on video lectures, multiple-choice quizzes or peer-review assessments, discussion forums and documents" [2]. It results that all the actions/reactions of each student will be recorded, from right/wrong answers, to time spent on each resource, to the number of visualized resources.

In this framework a research challenge is how to develop proper learning analytics (LA). One of the most cited definition of LA is 'the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs' [3].The aim of this paper is to propose food for thought in the direction of learning assessment of MOOCs. In particular the paper starts from the complex nature of the evaluation of the effectiveness and efficacy of MOOCs and focuses on exploring the main factors affecting students' performance: learning and engagement. *Engagement* is a multidimensional concept, involving aspects of learners' emotion, socialisation, behaviour and cognition [4]. It can be defined as the investment of time, energy, and effort in learning, that is a degree of participation, individual or collective, to the learning process. In educational literature, there are three distinct domains of engagement: behavioural, cognitive, and emotional and it is not limited to all the actions taken by learners to achieve their own goals. *Learning* can be defined as the process of acquiring new, or modifying existing, knowledge, behaviours, skills, values, or preferences [5]. As both concepts are complex and cannot directly measured, conceptualisation, operativisation and measurement of learning and engagement are very interesting challenges in the educational science debate. On one hand, the paper provides a possible point of view to such a debate suggesting a possible way forward identifying the different determinants of each concept and, for each of them, a proper set of indicators. On the other hand, it is provided a possible model to measure the impact of engagement and learning on the performance. Actually, the network of relationships can be structured as complex as desired because the statistical literature offers many goodness of fit indexes and information criteria to choose the best theoretical model. The impact of engagement and learning on the students' performance is explored through Partial Least Squares path modelling (PLS-PM) [6] [7] one of the most widespread approaches to study a network of relationships between concepts that cannot be directly measured.

The analysis refers to the courses offered by the Platform FedericaX, the EdX MOOCs platform of the "Federica WebLearning" Center at University of Naples Federico II. In 2015, the University of Naples, Federico II has entered the world of MOOCs launching a new series of online courses with more interactive features and more multimedia contents. To date around 75 courses are offered, and they range from basic introductory courses to topics of particular public interest. FedericaX courses are provided into two versions: an instructor-paced, that is a course following a set schedule, with specific dates for assignments, availability of course materials and exams, and a deadline within which learners have to complete the course and get certification, and a self-paced version, where all of the course materials are available as soon as the course starts and assignments and exams do not have due dates, so a learner can progress through the course at its own speed and can request a certificate as soon as you have a passing grade in the course, even without complete all of the course materials [34]. Data generated from courses and learner activities are available to EdX partner institutions through the EdX data package which consists of a set of compressed and encrypted files that contain event logs and database snapshots. The data package includes event logs (also known as the clickstream), course content exports, courseware database exports, forums database exports and open response database exports.

2 CONCEPTUALISATION AND OPERATIVISATION

In order to define the MOOC assessment dimensions, it is important to distinguish between learning and engagement concepts. While we can define learning easily as the process of acquiring new, or modifying existing, knowledge, behaviours, skills, values, or preferences [5], engagement is characterised by different concepts and dimensions. The most used are behavioural, emotional and cognitive engagement [8]. According to this perspective, behavioural engagement includes all the actions taken in learning activities, emotional engagement is defined as affective feelings for coursework, teachers or institutions and cognitive engagement is the attitude to undertake in even difficult tasks and to create efficient learning strategies. According to [9] [10] researchers should take care to ensure that construct definition drives their choice of measures rather than the selection of measurement determining how engagement is conceptualized in the research. Researchers must decide if their engagement is behavioural, cognitive, or emotional, or a combination of them. These decisions affect the choose of appropriate indicators. If both cognitive and emotional engagement are important for the research, then clearly a method of measurement that captures both dimensions is essential. In particular, if we see emotional engagement as students' emotional reactions to academic subject areas such as science or to school more generally [11], in MOOCs' context it become the learner's feelings for learning process participation and reach its achievement. That is higher a learner with positive emotions with the subjects of course and in general with learning by MOOCs is more likely to complete the course or get certification and more regularly than the others.

External information such as demographic data should also be taken into account.

2.1 Socio-demographic dimension

The first dimension relates to learners' socio-demographic characteristics. These indicators, typically, are provided by the user during the registration phase (for the EdX fields see [12]).

Table 1. Indicators for socio-demographic dimension.

Subdimensions	Indicators
-	Age
	Gender
	highest level of education attained
	Nationality

2.2 Learning dimension

The learning dimension can be structured into three sub-dimensions: frequency-based actions, time-based actions and interactions. Frequency-based data are the simplest and most used data for synthesizing and summarizing data from tracking log (e.g., [13]). Despite its simplicity, frequency can provide many useful information as the distribution of user events, in order to identify different behavioural patterns among learners. Time-based data give information about duration of time spent studying. In literature, the quantity of time spent in learning activities is a predictor of student performance [14]. According to [15],in a SLR context, analysing only the time spent in one or more learning activities is not enough, but it is necessary to analyse indicators about how a learner spent its time. [16] demonstrates the importance of time management in SRL activities. The interaction sub-dimension relates to discussion forums activity and social learning dimension, to which MOOCs platforms pay many attentions. Discussion forums activity not only allow peer-to-peer learning and a direct interaction between teacher and learners, but they help to reduce drop-out rate. Furthermore, [17] argued that in MOOCs contexts, learning does not depend only on regular coursework completion, but also on related activities like interaction. Then, forum data activities can be used not only to predict performance [18] but also to find learners' risk of dropping out factor and predict dropping out rate.[19].

Table 2. Indicators and sub-dimensions for learning dimension.

Subdimensions	Indicators
Frequency-based	Rate of active days
	Rate of different videos watched
	Rate of video re-watching
	Rate of different page visited
	Rate of video backward seeing
	Rate of speed change watching video
Time-based	Rate of time spent on each video
	Rate of videos watching time
	Mean time spent solving quizzes
Interaction	Number of forum pages viewed
	Number of provided response
	Number of provided threads

2.3 Engagement dimension

The engagement dimension can be structured into three sub-dimensions: reason of enrolment, regularity and procrastination. The reason of enrolment is related to the mode of enrolment and to the motivation [20]. As log data about learners' motivations are often poor and present lots of missing rows, they are usually collected through ad hoc surveys ([21] [22]). The regularity sub-dimension is a time-based dimension that analyses how a learner spends his/her time on the platform and how he/her organizes its own learning roadmap. You [15] proposed, among others, regularity as an alternative indicator to analyse time dimension in learning. The procrastination sub-dimension is a key factor of self-regulation and lifelong learning [23] since it can be viewed as the failure of the learner to organise its own learning process. Empirically procrastination has been analysed by Howell and Watson [24] and Tuckman [25]; the authors confirmed not only that procrastination has a negative impact on performance, but also that this impact is greater in a SRL context.

Table 3. Indicators and sub-dimensions for engagement dimension.

Sub-dimensions	Indicators
Reason of enrolment	Mode of enrolment
	Motivation
Regularity	Interval of days between activities
	Mean of time activity in a module
	Rate of return on previous modules
Procrastination	Time between module release and learners' activity
	Followed lesson ordering
	Difference in activity between starting and closing date

2.4 Performance dimension

Performance represents the outcome dimension as it estimates the degree of effective learning of users through certifications and quiz scores.

Table 4. Indicators and sub-dimensions for performance dimension.

Sub-dimensions	Indicators
-	Rate of correctly solved quizzes
	Rate of correctly solved quizzes for each module of the course:
	Rate of correct answers at first attempt
	Difference in the rate of correct answers among modules (improvement during the course)

3 METHODOLOGY AND MODEL DEFINITION

Partial Least Squares Path Modeling (PLS-PM) is a structural equation model with a component-based estimation approach; it has been proposed as an alternative to the classical covariance-based approach models (CBSEM). Structural equation models are largely used in educational science. For instance, Marks et al. [26] proposed an on-line learning model using LISREL methodology for the analysis of perceived learning and perceived satisfaction of students in function of student-student, student-instructor and student-context interactions. Marsh et al. [27] compared exploratory structural equation model and confirmatory factor analysis to propose a motivation and engagement model based on of the Motivation and Engagement Scale across gender and time. Park [28] proposed a LISREL model to analyse behavioural intention with respect to online courses in function of individual, social and organizational factors, perceived usefulness of online tasks and online skills. Bakker et al. [29] analysed the leaning dynamics in a work environment. Using moderated structural equation modeling, they defined the impact of work engagement and conscientiousness on active learning and work performance.

PLS-PM establishes its role in the international scientific literature as a soft modeling approach alternative to the CBSEM because it does not require restrictive distribution hypotheses and huge samples. This is a very interesting feature especially in those application fields where such assumptions are not tenable, at least in full. On the other side, the classical parametric inferential framework cannot be applied, and resampling methods are required to obtain empirical confidence intervals and hypothesis testing. PLS-PM is a statistical approach for modeling complex multivariate relationships among blocks of observed indicators (also known as manifest variables, MVs) and unobserved concepts named latent variables (LVs).A PLS path model is made of a measurement (or outer) model relating each block of MVs to its corresponding LV and a structural (or inner) model connecting the LVs in accordance with a network of linear relationships. The PLS-PM algorithm allows to estimate separately the blocks of the measurement model and the structural model. The measurement model estimation depends on nature of blocks, which can be distinguished in reflective and formative. In a reflective block, each latent variable is defined by a unique underlying concept reflected by the MVs. For this reason, MVs needs to covary and LVs needs to be unidimensional. In a formative block, a single variable or a group of variables represent a dimension of latent variable. For this reason, it is not necessary to assume unidimensionality of latent variable, but it is important that manifest variables do not covary, since LVs are a linear combination of manifest variables. For further information about basic concepts about PLS-PM see [6], [7], [30] for a review of method based on PLS-PM see [31] [32] for a comparison between covariance-based model and component-based models see [33].

Our hypothesis is to find the relationship between learning and engagement blocks and analyse their impact on performance. In Figure 1 a possible structural model is proposed considering that the socio-demographic, the learning and the engagement blocks are formative while the performance block can be considered as reflective. Dotted arrows represent optional links.

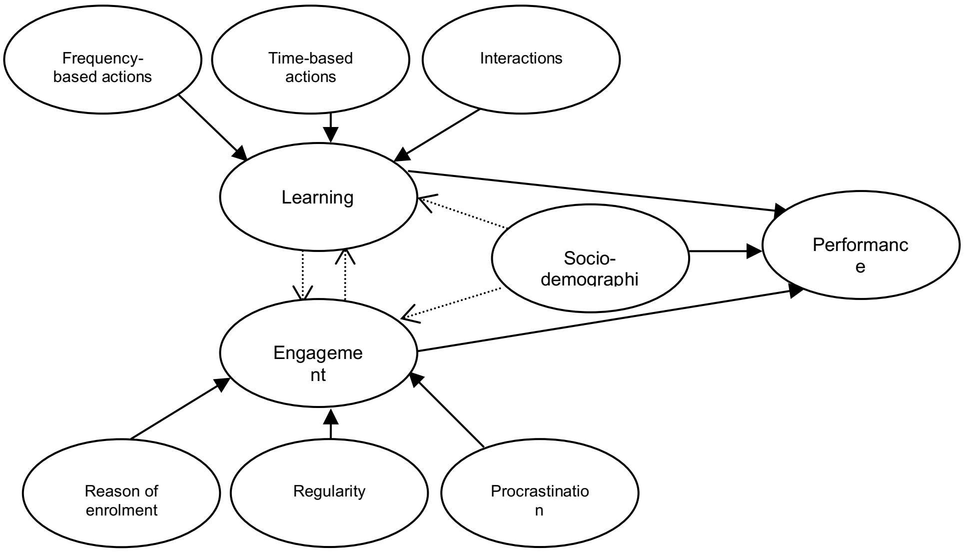

Figure 1. A structural model for MOOCs learning assessment.

ACKNOWLEDGEMENTS

Thanks to Federica Weblearning Center for Innovation and Dissemination of Distance Education,

University of Naples Federico II.

REFERENCES

[1] P. Winne and A. Hadwin "The weave of motivation and self-regulated learning". in *Motivation and self-regulated learning: Theory, research, and applications* (D. Schunk and B. Zimmerman eds.) pp. 297–314, Mahwah, NJ: Erlbaum 2008.

[2] M. Khalil, M. Ebner, "What Massive Open Online Course (MOOC) Stakeholders Can Learn from Learning Analytics?. Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy", in *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*(J.M. Spector, B.B. Lockee and M.D. Childress eds.), Springer International Publishing, 2016.

[3] L. Johnson, S. Adams Becker, V. Estrada, A. Freeman, *Horizon report: 2014 Higher Education*, Austin TX, USA: The New Media Consortium, 2014.

[4] G. Siemens and P Long, "Penetrating the Fog: Analytics in Learning and Education", *Educase Review*, vol. 46, no. 5.Retrieved fromhttp://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learningand-education.

[5] R. Gross, *Psychology: The Science of Mind and Behaviour* 6th Edition, London: Taylor & Francis 2010.

[6] V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang, Handbook of Partial Least Squares: Concepts, Methods and Applications, Springer, 2010.

[7] H. Wold, “Partial least squares,” in Encyclopedia of Statistical Sciences, S. Kotz and N. Johnson, eds., John Wiley &Sons, 1985.

[8] J. Fredricks, P.C: Blumenfeld, and A.H. Paris, "School Engagement: Potential of the Concept, State of the Evidence," *Review of Educational Research*, vol. 74, no. 1, pp. 59-109, 2004.

[9] G.M. Sinatra, B.C. Heddy and D. Lombardi, “The Challenges of Defining and Measuring Student Engagement in Science”, *Educational Psychologist*, vol. 50, no. 1, pp. 1–13, 2015.

[10] R. Azevedo, “Defining and Measuring Engagement andLearning in Science: Conceptual, Theoretical, Methodological, and Analytical Issues”, *Educational Psychologist*, vol.50, no. 1, pp. 84–94, 2015.

[11] R, Pekrun and L. Linnenbrink-Garcia, “Academic emotions and student engagement”, in, *Handbook of research on student engagement*(S.L. Christensen, A.L. Reschley, and C. Wylie eds.), pp. 259–282, New York, NY: Springer 2012.

[12] EdX research guide. Accessed 15 January 2018. Retrieved from https://media.readthedocs.org/pdf/devdata/latest/devdata.pdf

[13] W. M. Reed and J.M.Oughton ”Computer experience and interval-based hypermedia navigation”, *Journal of Research on Computing in Education*, vol. 30 no. 1, pp. 38–52, 1997.

[14] J. Lee, “Racial and Ethnic achievement gap trends: reversing the progress toward equity?”*Educational Researcher*, vol. 31, no. 1, pp. 3–12, 2002.

[15] J.W. You, “Identifying significant indicators using LMS data to predict course achievement in online learning”, *Internet and Higher Education*, vol. 29, pp. 23–30, 2015.

[16] L. Song, E.S. Singleton, J.R. Hill and M.H.Koh“Improving online learning: Student perceptions of useful and challenging characteristics”, *The Internet and Higher Education*, vol. 7, no. 1, pp- 59–70, 2014.

[17] R. Kizilcec,C.Piech and E. Schneider “Deconstructing disengagement: Analyzing learner subpopulations in massive open online course”, *Proceedings of the third international conference on learning analytics and knowledge*, pp. 170–179, 2013.

[18] N. Gillani and R. Eynon “Communication patterns in massively open online courses”, T*he Internet and Higher Education*, vol. 23, pp. 18–26, 2014.

[19] J. Cheng, C. Kulkarni, and S. Klemmer, “Tools for predicting drop-out in large online classes” *Proceedings of the 2013 conference on Computer supported cooperative work companion*, pp. 121-124, 2013.

[20] A. Littlejohn, N.Hooda, C. Milligan and P.Mustain, “Learning in MOOCs: Motivations and self-regulated learning in MOOCs”, *Internet and Higher Education*, vol. 29, pp. 40–48, 2016.

[21] C. Milligan, A. Littlejohn and A. Margaryan “Patterns of engagement in connectivist MOOCs”,*Journal of Online Learning & Teaching*, vol. 9, no. 2, pp. 149–159, 2013.

[22] B.K. Pursel, L. Zhang, K.W. Jablokow, G.W. Choi and D. Velegol, “Understanding MOOC students: motivations and behaviours indicative of MOOC completion”, *Journal of Computer Assisted Learning*, vol. 32, pp. 202–217, 2016.

[23] A. AlJarrah , M.K. Thomas, and M. Shehab, “Investigating temporal access in a flipped classroom: procrastination persists*”, International Journal of Educational Technology in Higher Education*, vol. 15, no. 1, pp. 1-18, 2018.

[24] A.J. Howell and D.CWatson, “Procrastination: associations with achievement goal orientation and learning strategies”, *Personality and Individual Differences*, vol. 43, no. 1, pp. 167-178, 2007.

[25] B.W. Tuckman, “Relations of academic procrastination, rationalizations, and performance in a web course with deadlines”, *Psychological Reports*, vol. 96, no. 3, pp. 1015-1021, 2005.

[26] R. B. Marks,S.D. Sibley,J. B. Arbaugh, “A structural equation model of predictors for effective online learning”, *Journal of management education*, vol. 29, no. 4, pp. 531-563, 2005.

[27] H.W. Marsh, G.A.D. Liem, A.J.Martin, A.J.S. Morin, and B.Nagengast, “Methodological Measurement Fruitfulness of Exploratory Structural Equation Modeling (ESEM): New Approaches to Key Substantive Issues in Motivation and Engagement”, *Journal of Psychoeducational Assessment*, vol. 29, no. 4, pp. 322–346, 2011.

[28] S.Y. Park, “An Analysis of the Technology Acceptance Model in Understanding University Students’ Behavioral Intention to Use e-Learning”, *Educational Technology & Society*, vol. 12 no. 3, pp. 150–162, 2009.

[29] A.B. Bakker, E.Demerouti, L.L. ten Brummelhuis, “Work engagement, performance, and active learning: The role of conscientiousness”, *Journal of Vocational Behavior*, vol. 80,pp. 555–564, 2012.

[30] M. Tenenhaus, V. Esposito Vinzi, Y.M. Chatelin and C. Lauro, “PLS Path Modelling”, *Computational Statistics and Data Analysis*, vol. 48 no. 1, pp. 159-205, 2005.

[31] J. Henseler and C.M. Ringle, “Using partial least squares path modelling in international advertising research: basic concepts and recent issues”, in *Handbook of research on international advertising* (S. Okazaki ed.), pp. 252-276, Edward Elgar, 2012.

[32] M. Sarstedt, C.M. Ringle and J.F. Hair, “Partial least square structural equation modelling”, in *Handbook of Market Research* (C. Homburg, M. Klarmann and A.Vomberg eds.), pp 1-40, Cham: Springer, 2017.

[33] W. Reinartz, M. Haenlein and J. Henseler, “An empirical comparison of the efficacy of covariance-based and variance-based SEM, *International journal of research in marketing*, vol. 26, pp. 332-344, 2009.

[34] FedericaX, Free online courses from Università degli Studi di Napoli Federico II. Accessed 15 January, 2018. Retrived from https://www.edx.org/school/federicax