

MOOC DESIGN PRINCIPLES: TOPIC MODELLING-PyLDAvis VISUALIZATION & SUMMARIZATION OF LEARNERS' ENGAGEMENT

Daniel F.O. Onah¹, Elaine L. L. Pang²

¹*Department of Information Studies, University College London (UK)*

²*Educational Development, Lancaster University (UK)*

Abstract

This research is based on the thought process that informed the design of a virtual reality massive open online course (MOOC) in FutureLearn. One of the main significant aspects of developing an online course for a large cohort of learners is to first consider the structure and pedagogic instructions and theories underpinning the design. This research covers aspects of good pedagogical design practice and perspectives used for structuring the content of the course. Designing an online course such as a MOOC requires careful consideration of the structure of delivery and the pedagogic flow of the instructional resources and process. The proposed virtual reality course follows good practice pedagogy from a design science perspective. The course content is structured in a logical manner that guides the learners in directing their studies in a sequential order of engagement that was designed by the course coordinator. This research reveals how good instructional pedagogy theories and principles help to facilitate teaching and learning easily. The development of concepts that lead to self-directed learning makes facilitating online learning very easy for the educators. The learners and course evaluations were done using incremental weekly quizzes and exercises. This course is developed for learners with little or no programming skills or no experience in virtual worlds (Augmented, Virtual and Mixed realities). The concepts were delivered with instructional structure to guide and direct the learners for independent learning no matter their location. In this course, learners' engagement with the resources is taken into consideration and reflective observations are done based on the demography of the registered participants. Action research and qualitative content analysis methods were applied to evaluate the text data from the research. The main purpose of the research is to investigate learners' reasons for engaging with the course structural resources provided and to explore aspects of mitigation for future MOOC designs. The result reveals that majority of the learners registered for the course to learn more about the virtual reality (VR) course and to learn from their interaction with other learners. This research applied some natural language processing (NLP) techniques to extract and summarize the prevalence topics or reasons for engagement from the icebreaker discussion. NLP techniques such as summarization, topic modelling and latent dirichlet allocation (LDA) were used to create the topic models for visualization.

Keywords: MOOC, pedagogy theories, self-directed learning, virtual reality, engagement, latent dirichlet allocation, topic modelling, pyLDAvis, summarization

1 INTRODUCTION

Developing massive open online courses (MOOCs) for large audience required high level of preparation and content structuring. There are several thought processes involved in developing MOOC delivery for learners from different background and mode of learning. MOOC has been predominantly used for delivering online teaching reaching out to large variety of people worldwide. Theories of engagement had laid the background work underpinning the learning design of massive open online courses [10]. Learners learn in relatively new ways progressively in both informal and formal environment. To apply our thought process in creating a massive open online course (MOOC), we investigate the meta-cognitive intersection in designing the learning processes of engagement using the pedagogic theory underpinning the course design. The course learning design follow the path of the taxonomy of course creation and development [31]. This study provides the pedagogic processes applied in development a Virtual World course in FutureLearn. The thought processes that was applied in structuring the course content, resources and the structure of the course. The platform has a well-structured approach to creating course content and delivery. There is existing *markdown* for code placeholders and syntax convention. The guidelines within the format of course presentation is promising because of the way this has been structured for course production. The pedagogic insights

underpinning this course design was based on learners' engagement and interventions to support independent learning goals.

1.1 Planning the course structure

Massive Open Online Courses (MOOCs) attract several learners in an open-access learning environment structure [31], [19]. This attract large numbers of people with a wide range of learning objectives and interests. Some of the participants might engage with the course as lurkers and once they achieve their goals then tend to leave. Hence, dropout becomes a common case as noticeable in most MOOCs [23]. MOOCs focused mostly on understanding and mitigating high drop-out rates and grades of learners without considering the fact that some learners might just be interested in a particular concept within the course provision and once they achieve this they tend to move on to other studies. We cannot conclude or add to the completion rate issues in most MOOC that these set of learners dropout from the course, however we can say their expectations at that particular course engagement was met.

In designing the course, we created the initial plans for the course structure and delivery. We had the initial course development plan, and organised several sectional team meetings with both the Lead Educators and FutureLearn representatives. This meeting was to discuss and agree on the course structure and the flow of the content in a structural manner. Our early decision-making in terms of the course development was very vital for structuring the entire course and weekly interaction with the resources provided. This early plan enabled the development team to apply any intervention in the structure of the weekly content and mode of interaction. Interaction among learners in MOOCs has been hailed to benefit course engagement for constructive learning [5], [24], [25].

1.1.1 What is the format?

In structuring the course content for FutureLearn, we have to create the steps for the various sections. There were steps for articles, videos, discussion, quiz, exercise and so on. Each of the steps are formatted properly in the structure of the course delivery. You can re-order the steps and move them around within the course structure.

This course is designed in such a way to allow learners to self-direct their learning [21]. Learners were given the opportunity to study at a reasonable self-pace and independently. They are encouraged to set their goals and engaged with the resources provided for their own practice and design. They are encouraged to share good practice design with their peers in the course discussion forum. The format of the course provision was to allow learners to take ownership of their individual learning and strategize their learning paths. MOOCs largely allow learners to study at their own pace and take control on how they wish to engage with the learning resources provided [8].

1.1.2 Markdown format

FutureLearn uses markdown for creating code snippets to illustrate specific concepts. This markdown enables the programming code to be well structured and makes this easily identifiable within the clear syntax convention of the language. Images are included in the article steps either as assets or uploaded into the course by using the upload button provided. These images are assigned links that could be copied and then embedded into the course with the markdown format for images.

For the video steps, these are created separately, the videos are added to the step that was provided by FutureLearn before the video description, subtitles, and transcripts can be done. FutureLearn has provided an extensive documentation to guide educators creating courses with them. The documentation contains all you need to know about formatting your course content properly and using the markdown format for programming codes and content creation.

1.2 Overarching plans

The overarching plan for developing the course was to create a simple guided or instructional virtual world course for learners with little or no programming skills. The learners do not need to have any idea about virtual worlds. Our main goal is to create a course that will improve the learners' digital

skills in virtual worlds. The course contains several interactive activities such as exercises, discussion forums and external resources that they could engage with while participating with the resources. In designing a course for MOOC learners' engagement and participation, careful planning should be considered before the creation of the course content and other resources. The course resources cover articles, videos, practical exercises, quizzes and external resources necessary for further reading. The course design could then be structured in a sequential manner in order for it to flow in the direction that would encourage effective learners' engagement.

1.2.1 Who is this course for?

The virtual world course is developed for participation from anyone. The course was created for worldwide audience including students, academics, and professionals and so on. The content is suitable to all participants no matter their level of understanding, background and culture. We aimed to reach out to learners participating in the form of distance learning and online delivery. In order to engage more participation, we will be facilitating the learners' support and providing instructional guidance in the course [15]. The course is open to everyone no matter their location and provided learners have access to good internet services and facilities that they could use to access the course resources. As this is a Massive Open Online Course, learners can engage with the course using any devices such as desktop PC, laptops and mobile devices such smartphones, tablets and so on.

2 LITERATURE REVIEW

Learners' engagement in e-learning is a topic that has been getting increasing attention lately especially due to the Covid19 pandemic [9]. Studies have indicated that online learning has a positive impact on student engagement [14]. Learners' engagement is believed to be an important benchmark and indicator of the quality of the learners' experiences and investment in learning activities on any course or programme [1]. Their engagement shows us the level of active learning that is taking place [16]. Researchers have attempted to study learners' engagement using many different definitions and conceptual frameworks [33]. According to Beder et al [1], engagement is mental effort focused on learning and there is a need to understand how learners engage because engagement is a precondition to one's learning progress. Engagement is also defined as the amount of effort dedicated to educational activities that help to bring out ideal performances in learners [13]. According to Reschly and Christenson [30], stated that the term engagement is seen as a multi-dimensional construct.

In a study by Walker and Koralesky [33], they defined learners' engagement as a multi-dimensional construct consisting of three interrelated dimensions including affective, behavioural and cognitive engagement. Connell et al, [7] categorised learners' engagement into three categories; namely the behavioural, emotional and psychological. Redmond et al [28] stated that there are multiple conceptualisations of engagement and these have resulted in diverse views about what engagement is among researchers. This has led to researchers also having different views about the place and interconnectedness of behavioural, cognitive, affective, socio-cultural, ecological and organisational factors [29].

Childs [4] stated that virtual student engagement is possible. In addition, Child [4] supports the work by Fredricks et al [11] and stated that there are three components of learner engagement and these are; emotional, behavioural and cognitive. Pickford [26], highlighted that when the dimensions of learner engagement is grouped in these three modes; emotional (social), behavioural (transactional) and cognitive (academic), the lecturer or instructor must ensure to take into account the individualistic nature of every individual learner's engagement as each learner can choose to engage in a different way. A learner may engage behaviourally first in the course to gain a professional qualification while another learner engages cognitively with the content, concepts and research to challenge themselves in their subject or discipline [26]. With reference to all the various definitions and dimensions of learners' engagement, one important point to note is social relationships are important alongside academic engagement [6].

For this study, we used the definition of learners' engagement discussed in research papers by Childs [4], Fredricks et al [11] and Pickford [26], which explained that learners' engagement is a multi-dimensional construct and is made up of three interrelated dimensions. They are emotional, behavioural and cognitive engagement. This study employs the Virtual Engagement Framework developed by Childs [4] as seen in Figure 1.



Figure 1. Virtual Engagement Framework. Adapted from [4]

2.1 Emotional Engagement

Newmann [20] defined emotional engagement as learners' emotions about learning and this includes their interests, boredom and happiness. In an online course, learners are emotionally engaged when they are interacting with one another in a learning community. By taking part in activities online such as introducing one's self in a course 'Icebreaker' session, it strengthens the learners' sense of value and acceptance remotely. Additionally, the various activities provided by the course developer offer opportunities for the learners to share their interests with the other learners. Thus, this helps to develop strong relationships, a sense of community and a growth mindset by learning from each other.

2.2 Behavioural Engagement

Some observable behavioural characteristics of behavioural engagement include the level of effort that learners contribute towards their learning and the level of learning achievement [17]. Restorative practices play an integral role in promoting positive behaviour in our virtual classrooms [4]. When learners are participating in an online course and put in effort to complete a variety of tasks under the different topics, they develop behavioural engagement as they become more attentive in their involvement with the materials and peers in the course. The resources available online will support the learners to develop a sense of mindfulness, self-awareness and self-regulated learning (SRL) skills [22]. By collaborating with other learners in the online activities, learners also have the opportunities to form virtual circles where they develop group norms. They also get to share viewpoints and develop critical thinking skills.

2.3 Cognitive Engagement

Meece et al [18] stated that learners' specific goals more than intrinsic motivation foster engagement. As learners work through activities and course materials in an online course, they are actively engaged with the course resources. Their learning process would involved steps such as having authentic learning experiences, asking questions, using various learning strategies [4]. They are also making connections to previous knowledge as well as to the real world. Thus, the course developer uses virtual teaching and learning resources to provide relevant learning experiences for the learners in an online mode. Printrich [27] emphasizes that learners' do monitor their academic tasks and environment while being engaged.

The pyLDAvis has two fundamental displays; the left panel as observed in all figures (Figures 4 and 5) visualize all topics as in the form of a circle in a two-dimensional plane determined by *Jensen-Shannon divergence* between topics. Multidimensional scaling was used to project all the inter-related topic distance to two dimensions. Each topic is encoded in the circles and the larger the circle the more prevalence is the topic.

The second part is the right panel which resonate with the horizontal bar chart which represent each individual term of a topic that are very useful and meaningful from the selected circle. The chart is made up of automated overlaid bars which are associated to the corpus-wide-frequency of the term and as well as the topic-specific frequency of the term.

The slider (λ) in the visualization depict the significance or relevance of the rank terms. It is worth knowing that the terms of the topic are ranked in descending order by default in accordance to the topic-specific probability ($\lambda = 1$).

Figure 4 reveals similar result from the word-cloud full content analysis that reveals that learners engage in the course because they have little or no knowledge of virtual reality and they wanted to learn about the application.

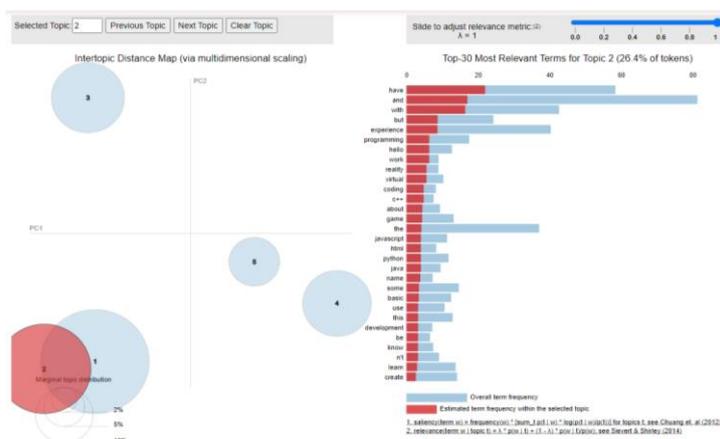


Figure 4. Interactive topic model visualization of the ‘icebreaker’ document with pyLDAvis

In Figure 5, we observed similar results from the word-cloud content analysis for the summary document which shows that most learners reasons for engaging with the course was to know more about virtual worlds. This further proves other studies that mentioned learners’ curiosity to learn new concepts in MOOCs.

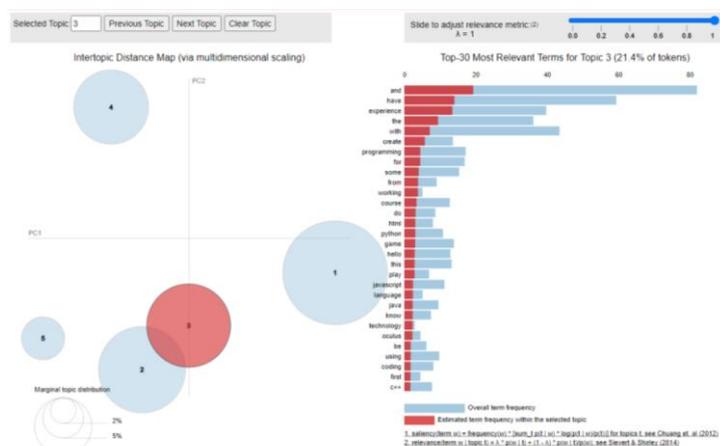


Figure 5. Interactive topic model visualization of the ‘icebreaker’ summary documents with pyLDAvis

5 CONCLUSIONS & RECOMMENDATIONS

This study introduced best practices in course development and creation using the FutureLearn platform. The course structure is designed in such a way that allows learners' participation in a guided and instructional manner. There are no initial prerequisites to participate in this virtual world course. The overarching goal is to create a course that covers virtual reality, briefly describe augmented reality and mixed or extended reality. We aimed to reach out to people at all levels and profession. We have developed the course to cover resources suitable for undergraduate students, postgraduates, academics and professionals who are looking at ways of improving their digital skill set. This study covers reflections on the thought processes of developing a Massive Open Online Course (MOOC) from the perspective of the Lead Educator. There are no data to present now, as this is an ongoing course. The course started in March 2021 and learners are currently engaging with the resources.

The study applied natural language processing (NLP) techniques such as Latent Dirichlet Allocation (LDA), summarization and topic modelling to extract the text data, pre-processed and pyLDAvis was used to visualise the emerging terms from the topics. We designed an interactive web-based pyLDAvis visualization of the text, which demonstrates the significance of a topic based on the proportional size of the circle. Although there are no prerequisites needed for this course, our result shows that most of the learners engaging with the course have little experience in some programming languages as revealed in the result section. Another prevalence term was experience which prominently reveals that majority of the learners in this course have little or no experience in designing virtual world applications. Therefore, they registered to engage with this course and learn how to design their own virtual world application as well as enhance their knowledge and progress with their interaction with peers.

How well do we need to create course resources for MOOCs delivery? MOOCs development without proper thought process of the learners' engagement tend to lead towards high dropout rates and participants disengaging with the course resources [23]. Massive Open Online Courses (MOOCs) should be developed to engage learners effectively considering the different learning styles and mode of independent learners. Course facilitation is very essential for any MOOCs delivery and this help in introducing in-person interaction with the learners and these could motivate them in their participation and effective course engagement.

5.1 What next?

This course is an ongoing course within Lancaster University in collaboration with FutureLearn and the Institute of Coding. In addition to this course, plans for further course design and creation is on the way for participation. This is just the first course out of many that will be developed by the University for Worldwide Participation. These thought processes would enable the creation of other courses using the approaches discussed in this study.

ACKNOWLEDGEMENTS

We acknowledge funding from the Institute of Coding, School of Computing and Communications at Lancaster University. Appreciation to The Joy Welch Educational Charitable Trust for research grant. Finally, we acknowledge FutureLearn for providing the platform for this course design and development.

REFERENCES

- [1] Beder, H., Tomkins, J., Medina, P., Riccioni, R., Deng, W. (2006) Learners' Engagement in Adult Literacy. Education. NCSALL Reports #28. National Center for the Study of Adult Learning and Literacy (NCSALL). Harvard Graduate School of Education, Cambridge, Massachusetts, United States.
- [2] Bengfort, B., Bilbro, R. and Ojeda, T., 2018. Applied text analysis with python: Enabling language-aware data products with machine learning. " O'Reilly Media, Inc."
- [3] Blei, D.M., Ng, A.Y. and Jordan, M.I., 2003. Latent dirichlet allocation. *The Journal of machine Learning research*, 3, pp.993-1022.
- [4] Childs, M. (2021) Virtual Student Engagement Isn't Impossible. In Respond & Reimagine: Putting Students' Needs First. ASCD Express, Vol. 16, No. 13. Alexandria, Virginia, USA.
- [5] Chua, S.M., Tagg, C., Sharples, M. and Rienties, B., 2017. Discussion analytics: Identifying conversations and social learners in FutureLearn MOOCs. *MOOC analytics: live dashboards, post-hoc analytics and the long-term effects*, pp.36-62.
- [6] Coates, H. (2006). Student engagement in campus-based and online education: University connections. New York, NY: Routledge.
- [7] Connell, J.P.; Halpem-Felsher, B.L.; Clifford, E.; Crichlow, W.; Usinger, P. (1995) Hanging in there: Behavioral, psychological, and contextual factors affecting whether African American adolescents stay in high school. *Journal of Adolescent Research*. 10, 41–63.
- [8] Deng, R., Benckendorff, P. and Gannaway, D., 2020. Learner engagement in MOOCs: Scale development and validation. *British Journal of Educational Technology*, 51(1), pp.245-262.
- [9] Dhawan, S. (2020) Online Learning: A Panacea in the Time of COVID-19 Crisis. *Journal of Educational Technology Systems*. 49(1), 5-22.
- [10] Ferguson, R. and Clow, D., 2015, March. Examining engagement: analysing learner subpopulations in massive open online courses (MOOCs). In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 51-58).
- [11] Fredericks, J., Blumenfeld, P., & Paris, A. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, 74(1), 59–109. Retrieved from <http://www.inquirylearningcenter.org/wp-content/uploads/2015/08/Fredricks2004-engagemt.pdf>
- [12] Haque, M.M., Pervin, S. and Begum, Z., 2013. Literature review of automatic single document text summarization using NLP. *International Journal of Innovation and Applied Studies*, 3(3), pp.857-865.
- [13] Hu, S., Kuh, G.D. (2002) Being (Dis)Engaged in Educationally Purposeful Activities: The Influences of Student and Institutional Characteristics. *Research in Higher Education* 43, 555–575 (2002). <https://doi.org/10.1023/A:1020114231387>
- [14] Junco, R., Elavsky, C. M., & Heiberger, G. (2013). Putting twitter to the test: Assessing outcomes for student collaboration, engagement and success. *British Journal of Educational Technology*, 44(2), 273–287. <https://doi.org/10.1111/j.1467-8535.2012.01284.x>
- [15] Kizilcec, R.F. and Schneider, E., 2015. Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(2), pp.1-24.
- [16] Lee, J.; Song, H.-D.; Hong, A.J. Exploring Factors, and Indicators for Measuring Students' Sustainable Engagement in e-Learning. *Sustainability* 2019, 11, 985. <https://doi.org/10.3390/su11040985>
- [17] Marks, H.M. (2000) Student Engagement in Instructional Activity: Patterns in the Elementary, Middle, and High School Years. *American Educational Research Journal*, 37, 153-184. <http://dx.doi.org/10.3102/00028312037001153>
- [18] Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology*, 80, 514–523.

- [19] Milligan, C. and Littlejohn, A., 2017. Why study on a MOOC? The motives of students and professionals. *International Review of Research in Open and Distributed Learning*, 18(2), pp.92-102.
- [20] Newmann, F.M. (1992) Student Engagement and Achievement in American Secondary Schools; Teachers College Press: New York, NY, USA, pp. 51–53.
- [21] Onah, D.F., Pang, E.L. and Sinclair, J.E., 2020. Cognitive optimism of distinctive initiatives to foster self-directed and self-regulated learning skills: A comparative analysis of conventional and blended-learning in undergraduate studies. *Education and Information Technologies*, 25(5), pp.4365-4380.
- [22] Onah, D.F., Pang, E.L., Sinclair, J.E. and Uhomoibhi, J., 2021. An innovative MOOC platform: the implications of self-directed learning abilities to improve motivation in learning and to support self-regulation. *The International Journal of Information and Learning Technology*.
- [23] Onah, D.F., Sinclair, J. and Boyatt, R., 2014. Dropout rates of massive open online courses: behavioural patterns. *EDULEARN14 proceedings*, 1, pp.5825-5834.
- [24] Onah, D.F., Sinclair, J.E. and Boyatt, R., 2014, November. Exploring the use of MOOC discussion forums. In *Proceedings of London International Conference on Education* (pp. 1-4).
- [25] Onah, D.F.O., Sinclair, J.E. and Boyatt, R., 2015. Forum Posting Habits and Attainment in a Dual-Mode MOOC. *International Journal for Cross-Disciplinary Subjects in Education (IJCDSE), Special Issue*, 5(2).
- [26] Pickford, R (2016) Student Engagement: Body, Mind and Heart – A Proposal for an Embedded Multi-Dimensional Student Engagement Framework. *Journal of Perspectives in Applied Academic Practice*, 4 (2). ISSN 2051-9788 DOI: <https://doi.org/10.14297/jpaap.v4i2.198>
- [27] Pintrich, P.R. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33–40.
- [28] Redmond, P., Heffernan, A., Abawi, L., Brown, A., & Henderson, R. (2018). An online engagement framework for higher education. *Online Learning*, 22(1), 183-204. doi:10.24059/olj.v22i1.1175
- [29] Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, 36(4), 257–267. doi:10.1016/j.cedpsych.2011.05.002
- [30] Reschly, A. L., and Christenson, S. L. (2012). Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 3–20). New York: Springer.
- [31] Rizvi, S., Rienties, B., Rogaten, J. and Kizilcec, R.F., 2020. Investigating variation in learning processes in a FutureLearn MOOC. *Journal of computing in higher education*, 32(1), pp.162-181.
- [32] Sievert, C. and Shirley, K., 2014, June. LDavis: A method for visualizing and interpreting topics. In *Proceedings of the workshop on interactive language learning, visualization, and interfaces* (pp. 63-70).
- [33] Walker KA. and Koralesky KE. (2021) Student and Instructor Perceptions of Engagement after the Rapid Online Transition of Teaching due to COVID-19. *Nat Sci Educ*. 2021;50:e20038. <https://doi.org/10.1002/nse2.20038>