

Education and Information Technologies

Importance of Learning Paths' Design in MOOCs: An analysis under the Socio-cultural Vision of Education Lens --Manuscript Draft--

| | |
|--|--|
| Manuscript Number: | EAIT-D-22-03030 |
| Full Title: | Importance of Learning Paths' Design in MOOCs: An analysis under the Socio-cultural Vision of Education Lens |
| Article Type: | Manuscript |
| Keywords: | Online curricular design; MOOC evaluation; massive online education; active digital learning |
| Corresponding Author: | francisco buitrago Universidad de Los Andes Bogotá, COLOMBIA |
| Corresponding Author Secondary Information: | |
| Corresponding Author's Institution: | Universidad de Los Andes |
| Corresponding Author's Secondary Institution: | |
| First Author: | francisco buitrago |
| First Author Secondary Information: | |
| Order of Authors: | francisco buitrago |
| | Luz Adriana Osorio |
| | Verónica Suárez |
| | Andrés Melani |
| | Gloria Cortés |
| | Jhon A Guerra-Gómez |
| | Marcela Hernández |
| Order of Authors Secondary Information: | |
| Funding Information: | |
| Abstract: | <p>MOOCs have been rapidly permeating online education in the last decade, promoting new ways of learning and teaching based on autonomous work and carefully designed curriculums. A key pedagogical component of such digital courses is learning paths, as a conglomerate of sequential activities that guides the process of learning. Learning paths have been widely described in online education, but little attention to the reasons of its success or failure have been investigated in MOOCs. In this study, we analyzed the relationship between following a proposed learning path in a coding MOOC and the performance in the course. We followed an analytic methodology for retrieving, cleaning, processing, and visualizing data from learning paths in our course, with a subsequent analysis based on concepts related to the theory of socio-cultural vision of education. We found that those students who followed the proposed learning path of the course showed a better performance in comparison to those who decided to navigate through the course in different paths. such effectiveness of the proposed learning path is unraveled considering the socio-cultural concepts of participation and reification, demonstrating why the sequence of activities designed for the course creates a learning environment that facilitates the building of knowledge.</p> |

Importance of Learning Paths' Design in MOOCs: An analysis under the Socio-cultural Vision of Education Lens

Francisco Buitrago-Florez, Universidad de los Andes, Bogotá, Colombia,
sicks@uniandes.edu.co - corresponging author.

Luz Adriana Osorio, Universidad de los Andes, , Bogotá, Colombia,
losorio@uniandes.edu.co

Verónica Suárez, Universidad de los Andes, , Bogotá, Colombia,
verosu@uniandes.edu.co

Andrés Melani, Universidad de los Andes, , Bogotá, Colombia,
af.melani3365@uniandes.edu.co

Gloria Cortés, Universidad de los Andes, , Bogotá, Colombia,
gcortes@uniandes.edu.co

Jhon A Guerra-Gómez, Northeastern University, , Bogotá, Colombia,
ja.guerrag@uniandes.edu.co

Marcela Hernández Hoyos, Universidad de los Andes, Bogotá, Colombia,
marc-her@uniandes.edu.co

Editorial Board
Education and Information Technologies

October 24. 2022

Dr. Editorial board,

Please find enclosed our manuscript entitled “Importance of Learning Paths’ Design in MOOCs: An analysis under the Socio-cultural Vision of Education Lens” by Francisco Buitrago-Flórez, Luz Adriana Osorio, Verónica Suárez, Andrés Melani, Gloria Córtes, Jhon Guerra and Marcela Hernández.

MOOCs have been rapidly permeating online education in the last decade, promoting new ways of learning and teaching based on autonomous work and carefully designed curriculums. A key pedagogical component of such digital courses is learning paths, as a conglomerate of sequential activities that guides the process of learning. Learning paths have been widely described in online education, but little attention to the reasons of its success or failure have been investigated in MOOCs. In this study, we analyzed the relationship between following a proposed learning path in a coding MOOC and the performance in the course. We followed an analytic methodology for retrieving, cleaning, processing, and visualizing data from learning paths in our course, with a subsequent analysis based on concepts related to the theory of socio-cultural vision of education. We found that those students who followed the proposed learning path of the course showed a better performance in comparison to those who decided to navigate through the course in different paths. Such effectiveness of the proposed learning path is unraveled considering the socio-cultural concepts of participation and reification, demonstrating why the sequence of activities designed for the course creates a learning environment that facilitates the building of knowledge.

This manuscript has not been published nor is it under consideration by any other journal. There are no financial obligations with governmental or private organization that can affect the content, results or conclusions of the present manuscript.

Thank you for considering our submission.

Sincerely,

Francisco Buitrago
Universidad de los Andes

Competing interests

The authors declare that they have no competing interests

Author contributions section

FB was a major contributor in writing the manuscript

LO Analyzed and interpreted data

BS was a major contributor in writing the manuscript

AM Analyzed and interpreted data

GC Analyzed and interpreted data

JG was a major contributor in writing the manuscript

MH was a major contributor in writing the manuscript

Importance of Learning Paths' Design in MOOCs: An analysis under the Socio-cultural Vision of Education Lens

Abstract

MOOCs have been rapidly permeating online education in the last decade, promoting new ways of learning and teaching based on autonomous work and carefully designed curriculums. A key pedagogical component of such digital courses is learning paths, as a conglomerate of sequential activities that guides the process of learning. Learning paths have been widely described in online education, but little attention to the reasons of its success or failure have been investigated in MOOCs. In this study, we analyzed the relationship between following a proposed learning path in a coding MOOC and the performance in the course. We followed an analytic methodology for retrieving, cleaning, processing, and visualizing data from learning paths in our course, with a subsequent analysis based on concepts related to the theory of socio-cultural vision of education. We found that those students who followed the proposed learning path of the course showed a better performance in comparison to those who decided to navigate through the course in different paths. Such effectiveness of the proposed learning path is unraveled considering the socio-cultural concepts of participation and reification, demonstrating why the sequence of activities designed for the course creates a learning environment that facilitates the building of knowledge.

Implications for practice or policy

- In-depth inquiry of the relationship between performance and learning paths in MOOCs.
- Novel analysis of learning paths design under socio-cultural vision of education concepts.

Keywords: Online curricular design, MOOC evaluation, massive online education, active digital learning,

Introduction

A plethora of courses have been available through the Internet for many years, however, only recently a massive interest for online education has been growth among students and educators worldwide. Such renewal of interest is a consequence of the arrival of Massive Open Online Courses (MOOCs for short). According to Jansen et. al. (2015), MOOCs can be defined as: "online courses designed for large numbers of participants, accessible by anyone anywhere as long as they have an Internet connection, are open to everyone without entry qualifications, and offer a full/complete course experience online for free". Dissemination of MOOCs is outstanding, as generally thousands of participants are enrolled in a single course, with over 7,000 in the few past years (Shah, 2016).

Information and Communication Technologies (ICT) evolve frequently and provide innovation in multiple services and functionalities for the design and implementation of MOOCs. One of such innovations are the Learning Management Systems (LMS), which helps teachers to administrate, document, track down, report, automatize and deliver educational courses in e-learning environments (Tahiri et al., 2017). A key component of an LMS is the denominated learning path, which is defined by De Smet (2014) as "the LMS functionality to order a number of learning objects in such a way that they result in a road map for learners. Within a learning path, learning steps are structured in a general way (as a navigation map or a table of contents) or in a very specific sequenced way (e.g., complete first step 1 before moving on to step 2)" (p. 2). Learning paths encloses the potential to improve the interaction in the processes of teaching and learning between teachers and students, an interaction embedded in a digital environment (De Smet et al., 2016). Therefore, in this study we determine the effect of a determined learning path of a programming MOOC in the performance of students, with a subsequent analysis framed by concepts belonging to the socio-cultural vision of education theory.

Literature review

MOOCs in higher education

The wide-spread enrolment of MOOCs is a somewhat recent phenomenon starting in 2008, which obtained high acceleration around 2011 and 2012 (Bonk et al., 2015), offering a scalable, sustainable and viable alternative form for teaching and learning in higher education (Selwyn et al., 2015). Some stakeholders and educators actually consider this sort of online courses to be a keystone of transformation holding the potential of democratizing the higher education system (Koller, 2015). In fact, such transformation can be evidenced in Dillahunt (2014), whom identified that learners unable to afford formal higher education had higher completion rates if they were motivated by the obtention of a certificate. Moreover, it has been elucidated that just in 2014 students from 194 countries were involved in this sort of digital courses (Ebben & Murphy, 2014).

MOOCs have also faced multiple criticisms from professionals in higher education and in multiple sectors of education alike, nonetheless. Some researchers have declared that this type of courses was created looking for the benefit of universities, as thousands of students might be enrolled with a lower number of instructors and learner interactions, leading to a self-regulated and self-engaged learning environment (Koutropoulos et al., 2012). Learners often struggle in such environment though, as such structure and the absence of instructional support is not the typical learning scenario (Zhu et al., 2018). Therefore, the hot issues on MOOCs in present years are oriented to their differences from the presential teaching in form, teaching methods and ideas, and their potential effects in the current higher education models (Li, 2019). Furthermore, with the overwhelming number of MOOCs to choose from, it becomes particularly difficult to assess the implications in learning of MOOCs as a general field. As such, there are some courses that are more well received by students than others, leading to important questions on what makes a MOOC better than other at the time of promote engage in the online learning experience (Hew, 2016).

In all learning schemes, including online learning, student engagement has been considered an essential element for building knowledge and skill development (Guo et al., 2014). Previous research in presential learning have pointed out that student engagement manifest impacts in both short-term and long-term development. Regarding the former, findings indicate that engagement have a considerable influence in grades and conduct (Hill & Werner, 2006; Lam et al., 2012; Marks, 2000). Concerning the latter, engagement have been linked to individual academic accomplishment, socially appropriate behaviors and self-esteem (Hawkins et al., 2001; Maddox & Prinz, 2003). In the context of MOOCs, student engagement becomes more challenging as there is a large and diverse student body (Hew, 2016). There are multiple student scenarios, ranging from the ones who find engagement due to curiosity, personal interest or the possibility of an extrinsic value for work advancement (Breslow et al., 2013), to the ones who are unable to engage and dropout before the competition of the course. Hence, the quality of a MOOC should be determined by the learning experience and the achievement of objectives of learning, given that both can be good predictors of student engagement.

MOOCs and learning paths

Learning paths rely in a sort of building blocks normally identified as learning objects. Kay and Knaack (2007) describe them as “interactive web-based tools that support the learning of specific concepts by enhancing, amplifying, and/or guiding the cognitive processes of learners”. Learning objects within a learning path might display a variety number of features and functionalities, such as appearance, context, animation, content, structure and behavior; hence, professors and instructors often decide the use of learning path based on their multimedia nature (De Smet et al., 2016). Highly interactive learning objects might guarantee engagement among learners, as they fell attracted to dive into the content and activities through multimedia channels (Mayer, 2005). As a learner navigates through the learning path, s/he is able to display cognitive processes that select images or sounds, creates mental representations and integrates experiences with previous knowledge. Such processes support cognitivist learning as promotes the development of abstract schemas of thinking, which subsequently increase the learning performance (Davis et al., 2016).

Learning paths play a key role in curricular alignment of MOOCs, given that its design directly impacts the levels of engagement of students. Curricular alignment guides the curriculum design of a course by articulating learning objectives, stated as clear actions that learners need to put into practice. These objectives are then developed through teaching and learning activities designed under pedagogical

principles and teaching methods, ultimately providing well-designed learning paths to learners (De Smet et al., 2014). It has been discussed that a successful curricular alignment based on learning paths, combined with evaluation tasks designed to assess comprehension and to deliver adequate feedback, lead to the successful completion of a MOOC (Guo et al., 2014).

Some investigations have been developed in e-learning and intelligent tutoring communities regarding the impact of learning paths in digital education (Kock & Paramythis, 2010; Rozinat & van der Aalst, 2008), however, little is known about its impact in MOOCs (Davis et al., 2016). Empirical evidence has shown that multiple reasons lead learners to decide to do not adhere to proposed learning paths in MOOCs such as lack of time, learning styles, engagement, among others (Guo & Reinecke, 2014), but there is no further analysis of the impacts of such decision in performance. Therefore, becomes evident the absence of in-depth investigation, framed by well described educational concepts from well-established learning theories, about the reasons behind the success or failures of learning paths design in MOOCs.

A socio-cultural vision of education under learning paths

Learning is defined as a process where knowledge is created through transformation of experience (Kolb, 1976). As such, experiences lead to a cognitive transformation in the learner's mind through reflection and theoretical support, providing active schemes for the building of knowledge. Conforming to the socio-cultural vision of education, learning must not be envisioned as an individual and isolated process in which the learner is able to enroll in a learning environment as a "receptor", and somehow create knowledge by passively interacting with a pedagogical scheme or through the repetition of specific actions (Vygotsky, 1978). Instead, such theory of education proposes that knowledge is created through active and complex interactions (Radford, 2008), as well as processes of recontextualization (van Oers, 1998). On the one hand, Wenger (1998) ideas explain that learners must be involved in two interactions, participation and reification. The former involves the actively interacting and creating identity in a community of learning, meaning that the learner must be involved in pedagogical actions that s/he could take active part of the learning process. The latter addresses that learners should continuously be immersed in the process of transforming abstract information into real artefacts, so they can build their own constructs through experience. By dynamically performing participation and reification, students would be able to build significant learning. On the other hand, as described by van Oers (1998) recontextualization can be seen as the process in which the actions and meanings of a person become less and less determined by the aspects from the context in which the those were originally learned. Hence, significant learning can be called as fully developed when students are able to *recontextualize* actions and concepts from an original situation to new defiant scenarios.

As such, the success of a learning path can be conceived as how much the pedagogical scheme of it is able to deliver multiple and dynamic moments of participation and reification, as well as deliver carefully designed processes of recontextualization. Such actions arise as critical in virtual environments, as generate them might result in a challenging task taking into account the absence of face-to-face interaction. Nevertheless, given that MOOC students navigate through learning paths in each course week, week after week, professors must ensure the curricular alignment of the course contemplates such actions. Therefore, the purpose of this paper is to describe the experiences of a case study MOOC of Java programming, focusing on the deployment of learning paths and explaining its success framed by the ideas of the socio-cultural vision of education.

Research questions

1. Does the structure of a learning path affect the success rate of students enrolled in a programming MOOC?
2. How socio-cultural concepts explain the success of a determined learning path?

Method

Setting

The case study chosen to determine the influence of the learning path in participants in the MOOC *Introduction to object-oriented programming in Java*, provided by Universidad de los Andes through

Coursera since August 2017. This MOOC aims to address some well-known difficulties that students have while learning to program, such as motivational problems and lack of balance between two key aspects involved in programming: algorithmic thinking and coding. The MOOC intends to contribute to the overcoming of these difficulties by means of 1. An active learning approach based on problems, 2. A balance between abstraction and programming skills development, and 3. A pedagogical methodology highly supported by Information and Communications Technology (ICT) tools to help students learn to program in a structured way. This MOOC seeks to provide a highly interactive way to learn how to program under the object-oriented programming paradigm, which allows students to validate not only their theoretical knowledge but also their coding skills. What is innovative about this course is that it gives real-time feedback to students on their performance on specific programming tasks, as well as provides feedback to students on how well they code, what mistakes they made, and why their solution is incorrect. Therefore, students understand their difficulties and correct their code as many times as they need as principal axis of a learning path.

The learning path designed for this MOOC starts by introducing theoretical concepts to learners. Then, students' transit from theory to practice in a dynamic way throughout coding vide-lectures. Subsequent coding challenges are presented to participants with real-life problem situations at different difficulty levels (basic, medium and advanced), where they are encouraged to test and improve their coding skills. Afterwards, the teacher engages the learners in the building of a relatively complex solution for a real-life problem via software development. Finally, a process of evaluation is carried out to obtain information about the level of theoretical and practical knowledge that students have acquired. It is important to highlight that such learning path is encompassed with active learning elements all the time. Examples of tools used to support active learning are CupiExámenes (<https://cupiexamenenes.virtual.uniandes.edu.co/>), CupiTips (<https://cupitips.virtual.uniandes.edu.co/#/>) and an in-house developed eBook (<https://universidad-de-los-andes.gitbooks.io/fundamentos-de-programacion/content/>). CupiExámenes is a web application that enables the creation, solution, and automatic grading of programming assessments, based on carefully designed study cases. CupiTips is another web application in which students can interact with the most common mistakes made while programming in Java. For each identified mistake, it is possible for the student to consult the associated topic, the description of the mistake and several examples. Students are able correct the mistakes in the code and obtain feedback in order to create a reflection learning environment. Lastly, the eBook arises as a complementary tool within the learning process in which students are encouraged to explore theoretical concepts in depth.

Overall, the MOOC consists of five modules each to be developed within 1-2 weeks. Most of the modules were developed with the same didactic structure, consisting of a motivational video, a recommended reading from the eBook, a group of well-organized and sequential theoretical video-lectures, three programming challenges and a theoretical and practical assessment. In total, the course comprises 81 videos, 33 readings, 12 challenges (distributed as one basic, one medium and one advanced over the four first modules), four automatically graded theoretical exams, four automatically graded practical exams, five peer- review assignments, and 5 forum discussions (1 per module). Module five does not include challenges or assessments, as it is focused on integration of knowledge. Considering that interaction in the learning process, understood as a planned opportunity for knowledge construction with others within a community for practice and learning, has been claimed as critical for digital environments (Darren, Soren, Suter, 2005), two pedagogical spaces has been dedicated for it. First, there are academic or technical consultation forums and the open-reflection forums created for each module, where peers interact both among them and with the course instructor. Second, students dive into peer reviewed activities, in which they can apply what they have learned following specific guidelines designed by the teacher.

Data sources

Two different data sources were used to extract data for analyzing the results of course in regard of the research questions established. On the one hand, MOOC data measures were collected from all MOOC activities: videos, readings, challenges, practical and theoretical assessments, peer- review assignments, and visualization of concept maps (only developed in modules 1 and 2). Combined, this resulted in a total of 24 variables. The order of MOOC activities (timestamps) was also retrieved. This information is available in the platform as Coursera CSV files, which include all the data that Coursera stores about the course and user's activity. On the other hand, students' performance on programming, measured as grades of challenges and practical assessments on CupiExámenes were used to obtain insights in the effects of the

proposed learning path. CupiExámenes information consisted of a grades database (PostgreSQL relational database) and an activity log (an unstructured text file of approximately 3 GB that stores entries for every activity graded by the tool).

Data pre-processing

Once the raw data was gathered, we focus on preprocessing the information so it could be structured correctly to then load it into the analytics tools. In this phase, an ETL (Extraction, Transform and Load) process was performed (Vassiliadis, 2009), in which the raw data was extracted from the sources, cleaned and consolidated using software developed on Java, Python and Tableau. The resulting data was stored on CSV files that were later loaded into the analytics tools. Hence, data from *CupiExámenes*' database was extracted through an in-house developed Java software, in order to clean, standardize and store information about the grades obtained by all students in each practical activity in CSV files. Then, using Python software, the IDs for every user were obtained from the log to map *CupiExámenes* users to their corresponding Coursera ID. By using Tableau software, the generated files were consolidated with the Coursera exports through the students' ID (especially with the course progress, that stores the timestamp for every attempt a student has done for each activity) to generate consolidated CSV files that were used as input for the analytics tools. Data from modules one to four was consolidated through this process. Data from module five was discarded since it does not include challenges or assessments. The preprocessing process is illustrated in Figure 1.

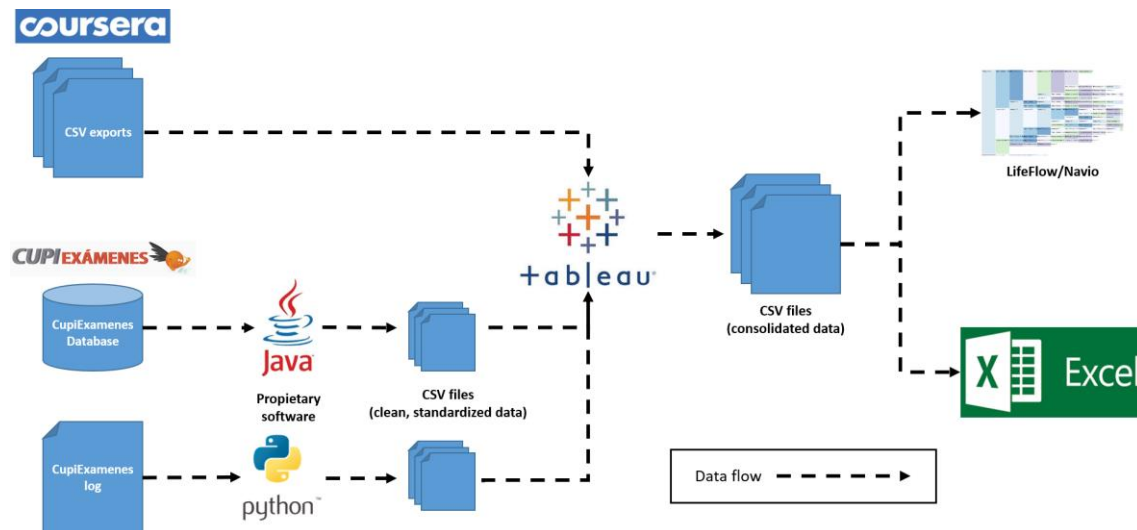


Figure 1. Preprocessing phases. ETL (Extract, Transform and Load) process for the data.

After data preprocessing consolidation, the volume of data obtained for the analyses in number of records was 346,062 when considering all students, 113,382 for paid users of the course and 43,197 for completers. This information also showed that each student does, on average, 162 attempts throughout all activities in the course.

Tools for data analysis

Two approaches were developed to explore the preprocessed data. In the first one we used multiple tools for data visualization such as EventFlow (Monroe et al., 2013) and Navio (<https://navio.dev/>). EventFlow, which is an evolution of Lifeflow (Wongsuphasawat et al., 2011), is a visual analytics tool for temporal categorical data developed at the Human-Computer Interaction Lab (HCIL) of the University of Maryland. The tool was designed to support summarization, exploration and querying of event sequences across time, derived from the categorical properties of the data. introduced clever visualization techniques that summarize thousands of records into an Icicle-like tree structure (Krustal & Landwehr, 1983). The tree aggregates all the possible sequences present in the records, making the visualization able to represent them as aggregated bars arranged in the vertical axis. Moreover, Navio is a visualization widget that summarizes large multivariate data into a single tabular representation, which also allows for dynamic queries. The data

is presented initially in the natural order it comes in, but the user can easily sort it by attributes. Additionally, once an interesting pattern is identified, the user can create range filters or categorical filters. The user can also combine filters with additions or negations. Once each filter is applied, a new Navio is presented to the user, showing the subset of data filtered, while keeping the original one for context. Then, the user can perform nested queries just with simple visual interactions and while keeping the context visible.

Regarding the second approach, we soon realized that EventFlow was good for the initial explorations, since the data was composed of a series of time-based (categorical) events in the course, as well as the main analysis was to understand and compare the overall paths followed by the students. However, we encounter an absence of functionalities for allowing users to quickly explore and query the original data in a dynamic way. To alleviate such lack, a custom in-house prototype was built that combines a basic implementation of the tree summarization technique of EventFlow, with the query flexibility of Navio. In a notebook-based platform for prototyping in JavaScript, when data is loaded the tool displays a visual summary of it using Navio. Users are subsequently able to query for specific subsets of the data. Each time a query is made, the tool displays a summary of all the common paths for that specific query, using a custom made EventFlow summary visualization (without including the time parameter). This strategy allowed us to compare paths for different subsets of the data and enabled us to generate and quickly test hypotheses. Furthermore, inspired in other projects such as Similian (Wongsuphasawat & Shneiderman, 2009), a simple querying system was implemented, which finds all the paths that include a particular event without filtering out all the other events (e.g., show all the activities taken by students who took the Module 2 exam).

Data analysis

Tableau and Excel were used to obtain grades' statistics and create diagrams, while LifeFlow/Navio strategy was used to report the students' learning paths. To analyze such paths, two student groups were defined given common characteristics of MOOCs: students who finished the course and students who does not. Each one of them was also classified according to their average practical activities' grade (theoretical grades were ignored as they did not show a significant difference between students). Additionally, learning paths were analyzed for each course module.

After loading the CSV file with the information of all students, a filter on the average (filter by grade) was applied to obtain those above 4 over 5. A second filter (Filter by module) was used to isolate data for module 1, obtaining a total of 50,915 records. It was then necessary to order them by timestamp to obtain the learning paths, as displayed in Figure 2.

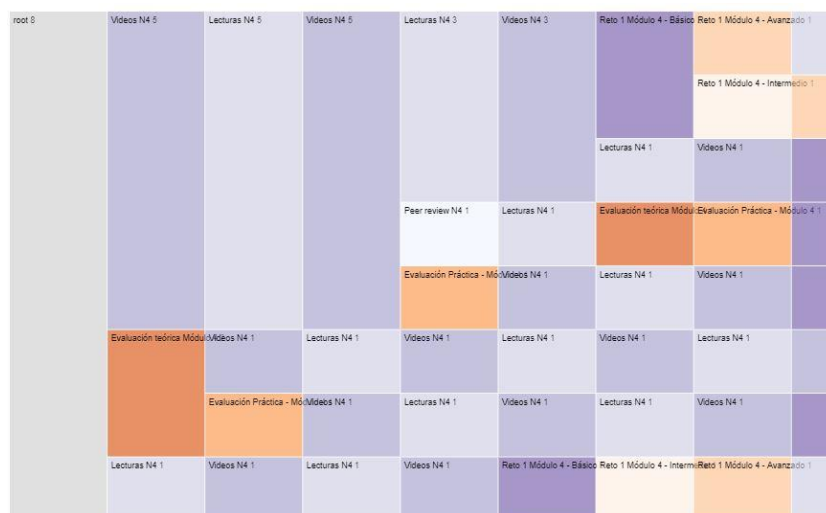


Figure 2. An example of resulting learning paths in our LifeFlow prototype.

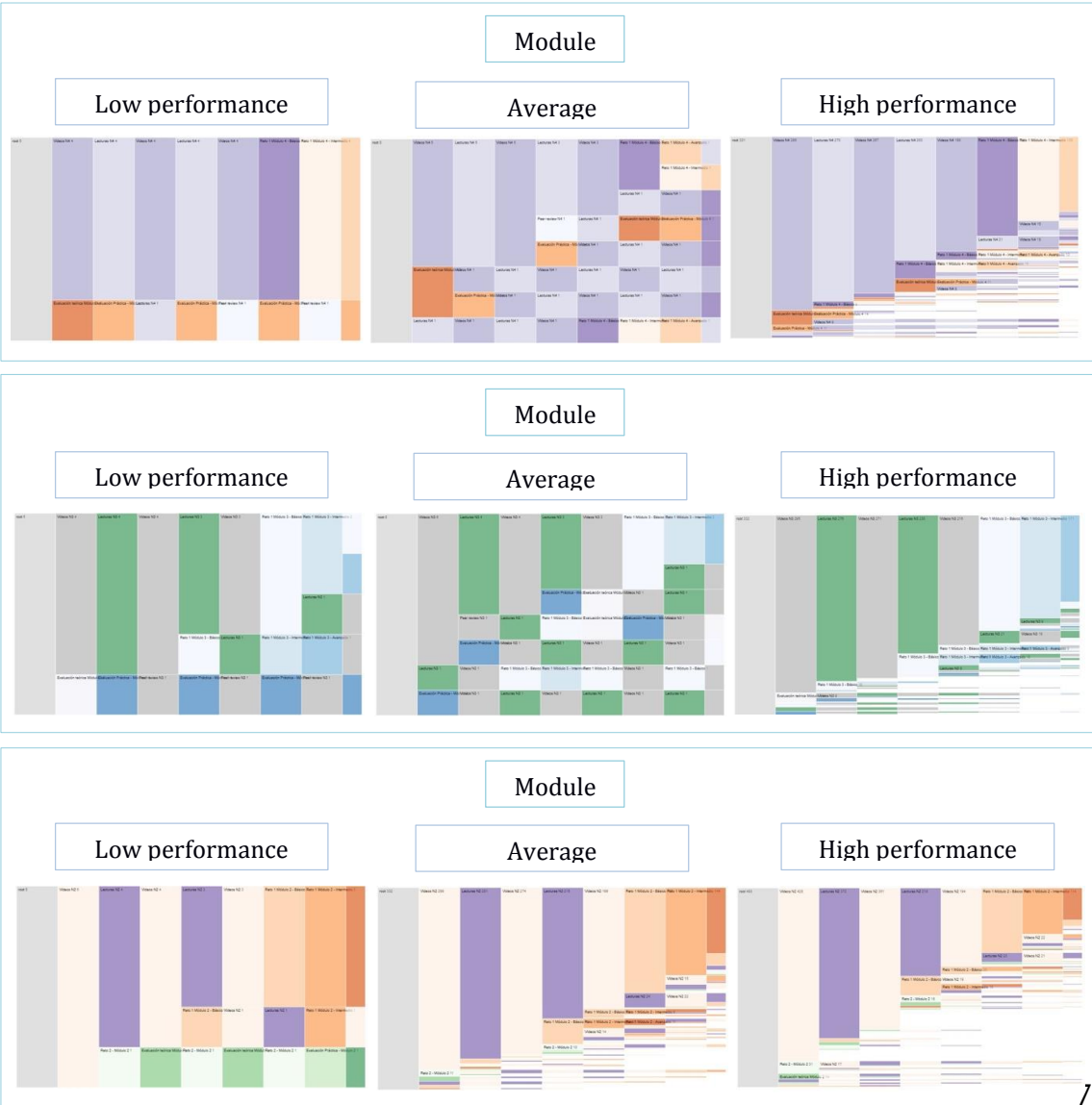
Each of the bars in the learning paths represents an activity or an aggregated activity to simplify the analysis, as in the case of videos. The bar's height represents the number of students following that particular path. For instance, the root in figure 2 displays the number 453, meaning that number of students started module

1 with an average grade above 4.0. From those students, 372 continue their paths by taking the entrance assessment, and 355 continue to watch a video. Such analysis allowed us to determine the most common paths for each specific group of students and to identify emerging patterns.

We were also interested in the analysis of student’s performance on challenges, practical and theoretical assessments. These measures were calculated for all active students who took challenges, as well as for the paid students who developed challenges, theoretical and practical assessments on each module of the course. Finally, we decided to explore the behaviors of the students with the best and worst results who has complete the course. Their particular paths were studied with a twofold purpose: looking for emerging patterns that could be related to their performance and compare them with the paths analyzed for the above-mentioned groups.

Results

Statistical analysis made over the grades indicated that all students who presented the theoretical assessments obtained higher results (4.49 average), therefore, such results were discarded as a filter for learning path analysis. Additionally, we found a notable difference in the performance of students in practical activities (challenges and assessments), leading us to select these activities as a filter for determining student’s learning paths. Furthermore, all students were divided in three groups by the average graded obtained in practical activities: *i)* high performance (students who obtained a grade of 4.0 or higher), *ii)* average performance (students who obtained a grade between 3.0 and 4.0) and *iii)* low performance (students who obtained a grade of 3.0 or below). To illustrate the learning path analysis results in terms of those groups, the diagrams obtained for modules 2, 3 and 4 are illustrated in Figure 3.



Overall, learning path analysis from modules 2, 3 and 4 allowed us to obtain interesting patterns observed in students. First, data indicates that most of the high-performance students followed the proposed learning path, which included the competition of the three level of difficulty challenges (basic, medium and advanced, in that specified order), with the subsequent development of theoretical and practical exams. We found that some students had to use more time going back and forth with videos and readings, but they eventually moved on to the development of challenges and assessments as the rest of the group. Second, some students belonging to the average group showed a similar behavior in their learning paths as those with high performance. However, the latest ones displayed a high variability in their learning paths. Third, we established that most of the students with low performance followed the proposed learning path for videos and readings, nonetheless they only developed the challenges catalogued as of basic difficulty, leaving unresolved medium and advanced challenges. Additionally, those who developed the basic challenge often returned to videos and readings before retaking the challenge again. Finally, data analysis showed that in every performance group analyzed there are some outliers who, in each of the modules, developed the theoretical assessment first and then proceeded with videos and readings, to sometimes retake the assessment. In this case we found there is a higher rate of dropouts.

Regarding the case analysis of the best and worst completers, their learning paths results were almost identical to their corresponding performance group patterns (Figure 4). We found a particular and significant difference, nevertheless, as the number of attempts on activities along the course result in more than 300 extra attempts in the case of the worst completer in comparison with the best. This data shows that a high-performance student does not need to retake too many activities to complete the course as someone with a low performance does, a phenomenon also observed in the groups' learning paths.

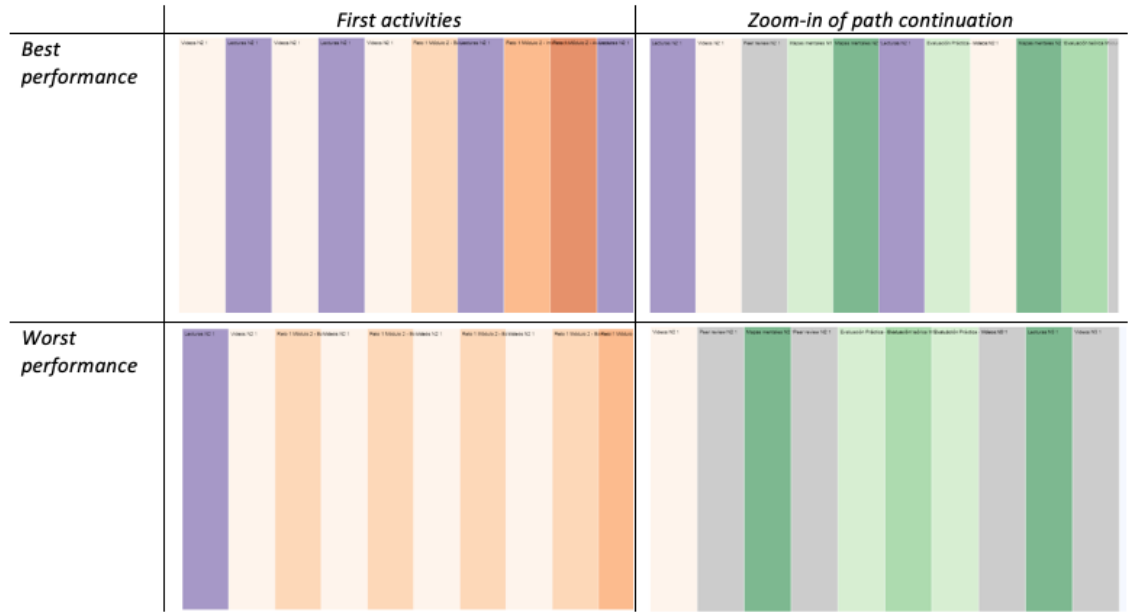


Figure 4: An example of learning paths from the best and worst completers in module 2.

Discussion

The research strategy in here described was elaborated to establish whether the deployment of a careful designed learning path positively influences in the learning outcomes of the programming MOOC. Following a well-structured quantitative approach, data indicates a strong difference in performance between those students who follow the proposed learning paths along the development of the MOOC and those who attempt to follow another self-made route for the competition of the course. The results obtained in this study provides a strong insight of how our pedagogical design successfully generates an educational

construct that improves how students approach to online education, which derivates in a better understanding of programming concepts and an efficacious development of programming skills throughout the MOOC.

Such success of the proposed learning path can be explained under participation and reification concepts proposed by Wenger (1998) in the socio-cultural vision of education. Regarding the former concept, participation is enriched throughout all the propose learning path as students navigate from theory to practice getting involved in multiple challenges, as well by being encouraged to test and debug for improving coding skills. Moreover, facing real-life problem situations leads to participation processes given that an active pedagogical atmosphere is created, where students are able to analyze, inquire, discuss, and evaluate multiple scenarios with alternative complex solutions. Consequently, students were able to interact with multiple forms of participation during the MOOC, promoting action and connection between theoretical programming notions and coding practices.

In terms of the latter concept, reification was fostered in students via a three-fold approach. First, the act of transform ideas into code itself is a powerful way of reification, giving that fragments of code becomes objects that represent the understanding of students, which is manifested through specific computational commands and concepts. Therefore, the different levels of difficulty of exercises, carefully designed by professors, helps the students to move from basic to complex cognitive constructs via reification processes. Second, students reify as they were encouraging to draft step-by-step the solution proposed for the real-life problem situations exposed in the course. Third, examinations allowed the learners to challenge their knowledge and represent it as objects in defiant scenarios, triggering complex processes of reification in scenarios under pressure.

Overall, students who followed the proposed path of learning display a better performance throughout the course. Such performance arise as a consequence of the multiple moments in which participation and reification were synergically generated by the course design. As proposed by Wenger (1998), when participation and reification are present learners are able to build significant meaning, leading to a more enriched learning experience. Moreover, students were exposed to multiple recontextualization scenarios as they successfully travelled through the established learning path. Taking into account the well-design set of exercises in each of the levels of the course, learners were able to recontextualize knowledge in new contexts as an effective way for increasing their expertise and improving their skills (van Oers, 1998). Furthermore, data analysis showed that our learning path improves the processes of teaching and learning between students and teachers in a MOOC, a key aspect in digital environments (De Smet et al., 2016). Each of the blocks of the course and its paths favored engagement and active interaction, increasing abstract schemes of thinking and stimulating building of knowledge (Davis et al., 2016).

Conclusions

Under MOOCs design, learning paths have become key curricular components for facilitating the transit of students and enrich their experience throughout the development of the course. Multiple pedagogical components must be taken into account to elaborate an adequate structure in a learning path, envisioning this educational tool as an enhancer in a learning digital experience. As such, we proposed and implemented a learning path for a programming MOOC in order to evaluate its impact in students' performance. Data indicates a strong relation between following the suggested learning path and performance in the course, showing that the pedagogical structure envisioned by the course professors successfully leads to programming learning and computational skill development. Additionally, data proved that those students who did not fully followed the learning path established struggled in accomplishing the objectives of learning of the course.

A deep analysis based in concepts from the socio-cultural vision of education allowed us to expose the reasons that lead to success our learning path. Those students who dedicatedly followed the learning path suggested for the course, had the opportunity to experience multiple moments of participation and reification articulated in the curriculum as precise activities to develop. Such active pedagogical actions offered high quality moments in which learners were able to build knowledge and recontextualize concepts and skills in different scenarios. By the explicit use of well-established educational concepts, we handle to evaluate how a MOOC learning path facilitates the navigation of students through objectives of learning.

Research studies as the one herein presented encloses the potential to improve learning paths not only in the field of programming and computer systems but in any field, as it holds a latent prospective for positively influencing curricular design and active learning. Therefore, we encourage professors to include socio-cultural concepts in their course schemes for enhancing the pedagogical experiences offered in MOOCs.

Limitations to the study and Future Research

The analytical power of the approach described could be enhanced by a larger sample from future cohorts for a longitudinal analysis. This would provide enough data to provide a definitive final statement about the scope of the approach herein described. A larger sample might find significant differences among learning paths that could enlighten improvements for the established pedagogical approach used in our programming MOOC. Moreover, this study has been focused on a specific programming MOOC designed in our institution. To obtain a more comprehensive picture of the power of the learning path proposed, the analysis of learning paths in other programming MOOCs might enhance the analysis of the scope and limitations of our proposal, as well as could provide a more general framework for the successful inclusion of learning paths through not only programming MOOCs, but courses designed almost in any area.

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request

Funding

Not applicable

Acknowledgments

Not applicable

References

- Bonk, C., Lee, M., Reeves, T., & Reynolds, T. (2015). *MOOCs and open education around the world*.
Breslow, L., Pritchard, D., BeBoer, J., Stump, G., Ho, A., & Seaton, D. (2013). Studying learning in the worldwide classroom. Research into edX's first MOOC. *Research & Practice in Assessment*, 8, 13–25.
Davis, D., Chen, G., Hauff, C., & Houben, G. (2016). Gauging MOOC Learners' Adherence to the Designed Learning Path. *Proceedings of the 9th International Conference on Educational Data Mining*, 54–61.
De Smet, C., De Wever, B., Schellens, T., & Valcke, M. (2016). Differential impact of learning path based versus conventional instruction in science education. *Computers & Education*, 99(1), 53–67.
De Smet, C., Schellens, T., De Wever, B., Brandt-Pomares, P., & Valcke, M. (2014). The design and implementation of learning paths in a learning management system. *Interactive Learning Environments*, 1–21.
Dillahunt, T., Wang, B., & Teasley, S. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *The International Review of Research in Open and Distributed Learning*, 15(5). <http://www.irrodl.org/index.php/irrodl/article/view/1841/3070>
Ebben, M., & Murphy, J. (2014). Unpacking MOOC scholarly discourse: A review of nascent MOOC scholarship. *Learning, Media and Technology*, 39(3), 328–345.
Guo, P., Kim, J., & Rubin, R. (2014). How video production affects student engagement: an empirical study of MOOC videos. *In Proc. L&S*, 41–50.
Guo, P., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs.

- L@A 2014, 21–30.
- Hawkins, J., Gou, J., Hill, K., Batting-pearson, S., & Abbot, R. (2001). Long term effects of the Seattle social development intervention on school bonding trajectories. *Applied Developmental Science*, 5, 225–236.
- Hew, K. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology*, 47(2), 320–341.
- Hill, L., & Werner, N. (2006). Affiliative motivation, school attachment, and aggression in school. *Psychology in the Schools*, 43(1), 231–246.
- Jansen, D., Schuwer, R., Teixeira, A., & Aydin, C. (2015). Comparing MOOC adoption strategies in Europe: Results from the HOME project survey. *The International Review of Research in Open and Distributed Learning*, 16(6).
- Kay, H., & Knaack, L. (2007). Evaluating the learning in learning objects. *Open Learning*, 22, 5–28.
- Kock, M., & Paramythis, A. (2010). Towards Adaptive Learning Support on the Basis of Behavioural Patterns in Learning Activity Sequences. *Intelligent Networking and INCOS '10*, 100–107.
- Kolb, D. (1976). Management and the Learning Process. *California Management Review*, 652–673. <https://doi.org/10.2307/41164649>
- Koller, D. (2015). The hype is dead, but MOOCs Are marching on. *Interview by Don Huesman, Knowledge@ Wharton, Last Modified January, 5*.
- Koutropoulos, A., Gallagher, M., Abajian, S., de Waard, I., Hogue, R., Keskin, N., & Rodriguez, C. (2012). Emotive vocabulary in MOOCs: Context & participant retention. *European Journal of Open, Distance and E-Learning*, 15(1). <http://www.eurodl.org/?p=archives&year=2012&halfyear=1&article=507>
- Krustal, J., & Landwehr, J. (1983). Icicle Plots: Better Displays for Hierarchical Clustering. *The American Statistician*, 37(2), 162–168.
- Lam, S., Wong, B., Yang, H., & Liu, Y. (2012). Understanding student engagement with a contextual model. In *Handbook of research on student engagement* (pp. 149–172).
- Li, Y. (2019). MOOCs in Higher Education: Opportunities and Challenges. *Advances in Social Science, Education and Humanities Research*, 319, 48–55.
- Maddox, S., & Prinz, R. (2003). School bonding in children and adolescents: conceptualization, assessment, and associated variables. *Clinical Child and Family Psychology Review*, 6, 31–49.
- Marks, H. (2000). Student engagement in instructional activity: patterns in the elementary, middle, and high school years. *American Educational Research Journal*, 37(153–184).
- Mayer, R. (2005). *The Cambridge handbook of multimedia learning*.
- Monroe, M., Lan, R., Lee, H., Plaisant, C., & Shneiderman, B. (2013). Temporal event sequence simplification. *IEEE Transactions on Visualization and Computer Graphics*, 12(12), 2227–2235.
- Radford, L. (2008). The Ethics of Being and Knowing: Towards a Cultural Theory of Learning. In *Semiotics in Mathematics Education*. Sense Publishers.
- Rozinat, A., & van der Aalst, W. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, 33(1), 64–95.
- Selwyn, N., Bulfin, S., & Pangrazio, L. (2015). Massive open online change? Exploring the discursive construction of the ‘MOOC’ in newspapers. *Higher Education Quarterly*, 69(2), 175–192.
- Shah, D. (2016). By the numbers: MOOCs in 2016. *Class Central*.
- Tahiri, J., Bennani, S., & Idrissi, M. (2017). diffMOOC: Differentiated Learning Paths Through the Use of Differentiated Instruction within MOOC. *IJET2*, 12(197–218). <https://doi.org/10.3991/ijet.v12i03.6527>
- van Oers, B. (1998). The Fallacy of Detextualization. *Mind, Culture and Activity*, 5(2), 135–142. https://doi.org/doi:10.1207/s15327884mca0502_7
- Vassiliadis, P. (2009). A Survey of Extract–Transform–Load Technology. *International Journal of Data Warehousing & Mining*, 5(3), 1–27.
- Vygotsky, L. (1978). *Mind in society* (E. H. U. Press (ed.)). Cambridge, Ma.: Harvard University Press.
- Wenger, E. (1998). *Communities of practice: Learning meaning and identity*. Cambridge university press.
- Wongsuphasawat, K., & Shneiderman, B. (2009). Finding comparable temporal categorical records: A similarity measure with an interactive visualization. *2009 IEEE Symposium on Visual Analytics Science and Technology*, 27–34.
- Krist Wongsuphasawat, John Alexis Guerra Gómez, Catherine Plaisant, Taowei David Wang, Meirav Taieb-Maimon, and Ben Shneiderman. 2011. *LifeFlow: visualizing an overview of event sequences*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1747–1756.

<https://doi.org/10.1145/1978942.1979196>
Zhu, M., Bonk, C., & Sari, A. (2018). Instructor Experiences in Designing and Delivering Interactive
MOOCs in Higher Education. *Online Learning*, 22(4). <https://doi.org/10.24059/olj.v22i4.1495>