UNIVERSITY OF SÃO PAULO
SCHOOL OF ECONOMICS, BUSINESS AND ACCOUNTING
BUSINESS DEPARTMENT
BUSINESS POS GRADUATE PROGRAM

A efetividade dos MOOCs através da experiência dos usuários

The effectiveness of MOOCs through users' experience

Fernando Antonio de Melo Pereira

Advisor: Maria Aparecida Gouvêa

São Paulo

2018

Administração

Prof. Dr. Marco Antonio Zago

Reitor da Universidade de São Paulo

Prof. Dr. Adalberto Américo Fischmann

Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Roberto Sbragia

Chefe do Departamento de Administração

Prof. Dr. Moacir de Miranda Oliveira Júnior

Coordenador do Programa de Pós-Graduação em Administração

FERNANDO ANTONIO DE MELO PEREIRA

A efetividade dos MOOCs através da experiência dos usuários

The effectiveness of MOOCs through users' experience

Doctoral Dissertation presented to the Business Department, School of Economics, Business and Accounting, University of São Paulo, as a requisite to the acquisition of a PhD Degree in Management Sciences.

Advisor: PhD Maria Aparecida Gouvêa

Corrected version

São Paulo

2018

FICHA CATALOGRÁFICA
Elaborada por Rafael Mielli Rodrigues – CRB-8/7286
Seção de Processamento Técnico do SBD/FEA/USP

Pereira, Fernando Antonio de Melo
The effectiveness of MOOCs through user's experience / Fernando Antonio de Melo Pereira. – São Paulo, 2018.
189 p.

Tese (Doutorado) – Universidade de São Paulo, 2018.
Orientador: Maria Aparecida Gouvêa.

1. Educação a distância 2. Administração 3. Administração de empresas I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 371.334

Abstract

MOOC (Massive Open Online Courses) is an emergent technology in education, a natural evolution of e-learning and it shows a disruptive way to learn. Although, many studies have focused on e-learning evolution, they ignore particularities of massive courses, which include the connectivity, evaluation process, virtual communities, economic and cultural impacts and other characteristics. This study sought to investigate the determinants of satisfaction and continuance of use in MOOC courses from the perspective of students. Through an online survey, 890 users of the most popular MOOC platforms made a self-evaluation of the performance and adherence to tools available in the courses. To generate a theoretical model, it is used a bibliometric and systematic review with meta-analysis, as well as methods of measurement scales, including Item Response Theory, random forest, common method variance and factorial analysis. The model was validated through the structural equation modeling. Among the main results, it is highlighted that the quality, usability and value impact on the performance of the users in the courses. In addition to performance, using interactivity and collaborative learning resources generates greater adherence of users to the teaching format of MOOCs. Both performance and adherence to MOOCs generate more satisfaction and a consequent desire to continue completing the courses on the teaching platforms. In addition, user engagement is crucial to generating satisfaction. The study contributes to the advancement and maturation of research on the theme, mainly at the empirical level and converging to the movement of MOOCs as a business model headed by universities. In the methodological field, the study contributes with the literature of quantitative methods for exploring emerging methods at measurement scales and complementing traditional methods. The originality of the study is based on the application of the particularities of the MOOCs in a context of continuance intention, in the exploitation of the studies on e-learning, in the evaluation of variables as payment and in the moderating effect of the engagement.

Keywords: Massive Open Online Courses, Disruptive Technology, E-learning theories, Effectiveness, Measurement Scale.

Resumo

O MOOC (Curso Online Aberto e Massivo) é uma tecnologia emergente em educação, uma evolução natural do e-learning e que apresenta uma forma disruptiva de ensino. Apesar disso, muitos estudos que focam na evolução do e-learning, ignoram as particularidades dos cursos massivos, que inclui a conectividade, o processo de avaliação, as comunidades virtuais, os impactos culturais e econômicos, entre outros. Este estudo buscou investigar as determinantes da satisfação e da continuidade uso em cursos MOOCs na perspectiva dos estudantes. Através de um survey online, 890 usuários das plataformas mais populares em MOOCs fizeram uma auto-avaliação sobre a performance e o uso de ferramentas utilizadas nos cursos. Para gerar um modelo teórico, foram utilizadas uma revisão bibliométrica e sistemática com meta análise, além de métodos de escalas de mensuração, incluindo a Teoria da Resposta ao Item, random forest, variância comum ao método e análise fatorial. O modelo foi validado através da modelagem de equações estruturais. Entre os principais resultados, destaca-se que a qualidade, usabilidade e o valor impactam na performance dos usuários nos cursos. Aliado a performance, utilizar recursos de interatividade e aprendizagem colaborativa geram maior aderência dos usuários ao formato de ensino dos MOOCs. Tanto a performance como a aderência aos MOOCs geram mais satisfação e consequente desejo de continuar realizando os cursos nas plataformas de ensino. Além disso, o engajamento do usuário é crucial para gerar satisfação. O estudo contribui no avanço e amadurecimento das pesquisas sobre o tema, principalmente no âmbito empírico e convergentes ao movimento dos MOOCs como um modelo de negócio encabeçado pelas universidades. No campo metodológico, o estudo contribui com a literatura de métodos quantitativos por explorar métodos emergentes em escalas de mensuração e se complementando a métodos tradicionais. A originalidade do estudo se baseia na aplicação das particularidades dos MOOCs em um contexto de intenção de continuidade, no aproveitamento dos estudos sobre e-learning, na avaliação de variáveis como pagamento e no efeito moderador do engajamento.

Palavras-chave: Cursos Online Abertos e Massivos, Tecnologia Disruptiva, Teorias do e-learning, Efetividade, Escala de mensuração

Summary

List of Figures

List of Graphics

List of Tables

1. Introduction

This chapter introduces the background of research theme and the main concepts related to the research area. A study proposal, gaps and a problem are discussed. After, general and specific objectives are presented. The sequence contains the expected contributions for research area, in theoretical, organizational and society scope.

1.1 Background and concepts

One of the biggest challenges of the twenty-first century is to migrate the information society to the knowledge society. The information and communication technologies (ICT) have allowed the learning process to break barriers of traditional teaching, such as time constraints, geographical distance and limited time (Behara, & Davis, 2015). This process of change only happens with the advancement of technology and the break of paradigms in education. The evolution of the Internet and ICTs are the backdrop to new educational technology-based solutions (Bond & Leibowitz, 2013; Little, 2013).

The moves alter dramatically both the higher education and the business world. At the epicenter of the changes is the e-learning, with a potential influence on higher education, offering low-cost options and high-quality content (Behara & Davis, 2015; Lin, Lin & Hung, 2015). Besides that, Spector (2014) points out that the biggest challenge in education is to make effective the use of new technologies while preparing students for productive lives in the twenty-first century.

The e-learning and higher education have developed a closer relationship from the rising of MOOCs (Massive Open Online Courses), becoming an important tool for the future of universities, along with the custom e-learning and game-based learning (Spector, 2013). The MOOCs are a promise for distance education and is considered a disruptive technology that is revolutionizing the education area in recent years (Alraimi, Zo, & Ciganek, 2015; Ventura, Bárcena, & Monje, 2014).

Belleflame and Jacqmin (2015) point out that the year of 2012 showed the emergency of a new player in higher education evolution. Many universities around the world are planning and offering MOOCs. In the first half of 2014, there was an increase of 91% in number of MOOCs offered, more than 2,600 courses were already being offered by universities and large organizations in the globe (Hew, & Cheung, 2014). According to the Open Education Europe

(Available in openeducationeuropa.eu) in February 2016, Europe offered 1,708 MOOCs, and 78 were soon starting. In USA, Lin, Lin and Hung (2015) affirm that approximately 33.3% of the students have chosen one or more online courses for learning.

E-learning, according to the American Society for Training and Development's is a broad set of applications and processes, which involves learning via web, computer-based learning, virtual learning and digital collaboration (Kaplan, & Leiserson, 2012). It is the current representation of distance education. Other themes can be associated with e-learning, such virtual communities, lifelong learning, digital inclusion, among others. The MOOCs are regarded as the fifth generation of e-learning.

Although there is no precise definition of MOOC, two features are highlighted: open access, it does not need to be a student enrolled in any educational institution, and scalability, that supports an unlimited number of participants (Siemens, 2008). Therefore, MOOCs are part of e-learning and are a hot topic in the field of e-learning (Bernhard et al. 2013).

In the context of MOOCs, there is great diversity among students. There are students interested in these courses to satisfy a personal desire or curiosity in the topic or subject. Some ones are just looking for a certificate to prove the skills they have, and others are obsessive to participate in the courses, treating as leisure or hobby (Hew, & Cheung, 2014). In point of view of the creators, a MOOC provide low cost and quality education to the masses, which justify this variety of motivations. In the last five years, very rich materials are being created, and this transformation in learning mechanism induce the students to study more and teachers to improve classroom skills, pushing universities into new pursuits (Kesim, & Altinpulluk, 2015).

In few words, following the basic features mentioned: MOOCs are online education platforms accessed for free by great masses. In this way, online courses of open access and with scalability are considered MOOCs. In this study, the definitions of e-learning and MOOCs presented guide the approach given to these issues during the research.

1.2 Gaps

With so many MOOCs to choose from, students are spoiled for choice. However, not all MOOCs are well received by students who enroll on them. There are courses better received than others (Hew, 2014). From this fact, it is important to evaluate what motivates the user experience in online learning. Researches about e-learning are consolidated and they reached a

mature level. By the other side, the MOOCs have many aspects that need development and correction (Kesim, & Altinpulluk, 2015).

The MOOCs can be used effectively to support a learning system with global dimension, but as a solitary educational technology, fail to add aspects already established by e-learning initiatives (Spector, 2013). When characteristics of e-learning success are combined with potential of MOOCs, this union can create a significant impact on a university and on education in general.

Measure these aspects under users' perception remains the most effective mean of evaluate services and consequently evaluate online education. Specifically, on higher education, Lindsey, Rhoads and Lozano (2015) emphasize that the benefits of MOOCs in this segment, act as a solution to overcome challenges and revolutionize higher education. And the authors continue, stating that little is known about the consequences of these online courses and which aspects arouse students' interest. Klobas (2014) corroborates the idea that the online open education is a real opportunity to increase information resources and knowledge, and also highlights the lack of measurement of the impact caused by MOOCs on all interested actors: students, universities and society.

The MOOCs constantly change and have particularities, such as evaluation system, interaction by social media, the form of entry, among others. Chang, Hung and Lin (2015) state that there are few studies that evaluate these characteristics. Therefore, although aspects of e-learning can be used to evaluate consistently the MOOCs, aspects of this teaching mode cannot be ignored.

Chiu et al. (2005) confirm that there are numerous researches that evaluate e-learning services; however, it’s a consensus in the literature that there are few studies that evaluate a MOOC under a management perspective focused on the user (Carr, 2012). Also, there are few studies that developed theoretical models focused on e-learning that can be applied in MOOCs studies. It is important to emphasize that the concept of MOOC follows the prerogative that MOOC is the newest stage of e-learning (Gelderman, Ghijsen, & Diemen, 2011). Thus, the development of theoretical models that include adaptive constructs to particularities of MOOCs is necessary.

1.3 Research problem

Being new and for being a phenomenon of learning, that has emerged in last years, the MOOCs are becoming popular in several countries, and now, more people are searching for

universities that offer these courses (Ma, Zheng, & Zhao, 2014. For being so different from traditional education and previous e-learning, terms such as online enrollment and dropout are being reconceptualized (Deboer et al. 2014). For example: 10% completion rate might seem a bad result for a course, but it seems to be a better measure to success to evaluate the rate of those ones who really wanted to complete the course. Alraimi, Zo and Ciganek (2015) say that the effectiveness of a MOOC is still an open question, and the completion rate is substantially lower than in other types of education mode.

Various methods and approaches that seek to understand the MOOCs, their enrollment and withdrawal have revealed that not all students seek to complete the course, and they also can interact with classmates by different ways and levels, accessing only forums, or just downloading and reading materials, or just searching for a group discussion on social networks, among other ways (Greene, Oswald, & Pomerantz, 2015).

Empirical studies that permeate the MOOC literature use theories related to adoption of technologies, with use of virtual learning environments (VLE) and with evaluation of user's experience that try some ICTs. Many studies focus on decision making with management purposes. In operational terms, these studies are seeking real-time measures of students' satisfaction and identify factors that can be improved and that strongly influence the retention of users (Abeer, & Miri, 2014).

The relations between students, between student and teacher and the students' perception about course, have consequences such as effectiveness of a course, from the point of view of management. Lin (2012) points out that satisfaction is one of those key measures allied with effectiveness. It is satisfaction that connects the intention or adoption with reuse or intention to continue doing courses.

The continuance intention is measured from the user experience, where a decision to continue doing courses happens after acceptance or initial readiness to try that course. These factors are an inheritance of the literature on post-purchase, acceptance of technologies and use of information systems (Lin, 2012). Thus, antecessor models have as theoretical foundations: marketing studies and IS (Information System) studies, focused on management.

This inheritance enables to develop with accuracy the determinants that define if a MOOC was effective and successful. Li, Duan, Fu and Alford (2011) conclude that an understanding of the multiple factors that affect student behavior, can improve knowledge for decision making, resulting in better learning systems. When these factors act together, they promote greater engagement, maximizing performance, and influencing cumulatively and recursively the student behavior (Lands, & Ramsay, 2015).

In this context, the problem formulated for this research is as follows:

What are the determinants that MOOCs users consider important to satisfactorily evaluate a course and decide to enroll in other courses?

1.4 Objectives

General objective: Analyze the effectiveness of MOOCs from the perspective of users in relation to satisfaction and continuance of use.

To achieve the general objective, the following specific objectives are proposed:

(i) Investigate the variables that compose user performance and evaluation aspects of a MOOC;

(ii) Validate MOOCs Adherence Index through methods of measurement scale;

(iii) Identify distinct groups of users by demographic variables and rating profiles of items;

(iv) Validate the theoretical model which measures MOOCs satisfaction and continuance of use through user's experience.

1.5 Justify

The ICTs have influenced many aspects of everyday life and have manipulated the field of education. With a fast development, driven by technological advances, ICTs are generating technological advances in distance education. The transformation of passive users to active users, the quality content production, driven and created in large-scale, represent the emergence of web technologies and applications for society (Kesim, & Altinpulluk, 2015). Chang, Hung and Lin (2015) corroborate this point of view, affirming that the technologies for education are important tools for social changes.

By a social angle, the rise of MOOCs is a tendency that is attributed to the initiatives of universities and other entities, with projects that offer high-level education to anyone with

Internet access, appearing as an essential tool for the democratization of teaching (Njenga, & Fourie, 2010). Since 2012, the huge growth of this teaching mode has generated social benefits such as flexibility in the learning process, break of geographical barriers, autonomous learning, self-discipline and other aspects, reflecting the economic and social importance of the issue to society (Loya et al. 2015).

This study aims to analyze the effectiveness of MOOCs from the perspective of users in relation to satisfaction and continuance of use. In order to reach this objective, a reliable and reproducible research instrument will be developed, capable to measure effectiveness and success of a MOOC course. The study focuses on the management aspects of a course, the quality of a tool, performance and satisfaction of users, and others aspects that show relevant. An efficient management of these courses ensures the popularity of MOOCs by improving enrollment rates, reduction of dropout rates, and sharing high level content, without cost. Klobas (2014), stresses the importance of management for the courses, because there are costs involved to promote a course with satisfactory quality, and to provide interaction between individuals with different interests and different levels of learning.

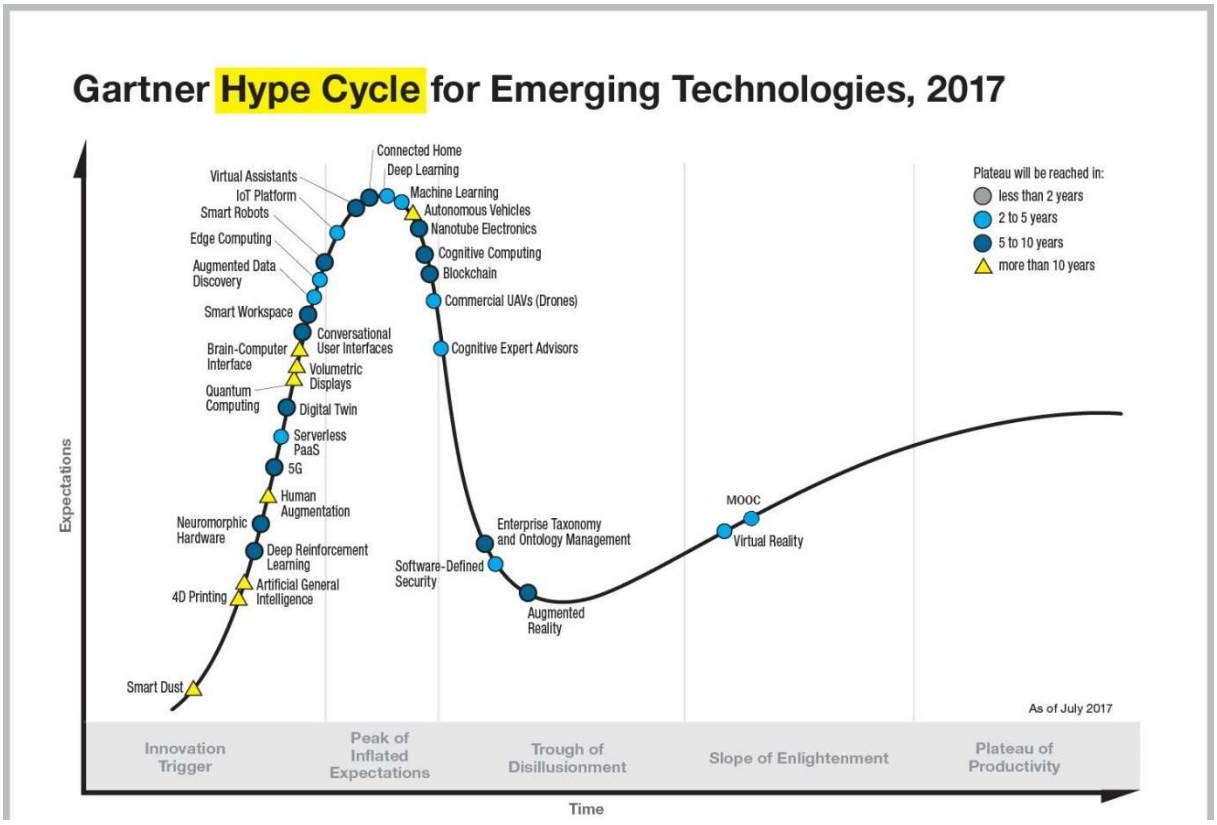
Another theoretical benefit is the use of constructs such as satisfaction and continuance intention of a service, that finds applicability in marketing studies, in use of information systems, use of public services, among others. In addition, the empirical validation of a theoretical model allows detailed examination of the issue from a quantitative perspective, from training indicators and effective use of analytical tools for decision making (Gomez et al. 2011).

Hype Cycle and MOOCs

To clarify the importance of this theme in 2018, it can be alluded the Hype Cycle of Gartner Group, which identifies emerging technologies and allows the allocation of the MOOCs in the Hype Cycle chart. The timeline of chart marks common periods in a line of time. The first period is a trigger, when the technology is a novelty and the commercial viability is not yet being proven. The second period is marked by a growth of technology, adding the expectations generated by the new technology. It's a period of enchantment and a peak of inflated expectations. The third period is a consequence of the second period and is named through of disillusionment. This stage occurs because the generated expectations were not fully met. In fourth phase, slope of enlightenment, the overstatement and disillusionment are replaced by a rational approach. By the end, in fifth phase, the technology reaches maturity.

From the studies that seek to define the current situation of MOOCs, there is a prevailing view that MOOCs are recently thought rationally. The researchers are realizing that there are success stories, but also in certain contexts, the MOOCs cannot work. Thus, the MOOCs are in the beginning of fourth stage, as can be seen in Figure 1.

Figure 1: Hype Cycle 2017 with MOOC



Source: Adapted* from Gartner Group, 2018.

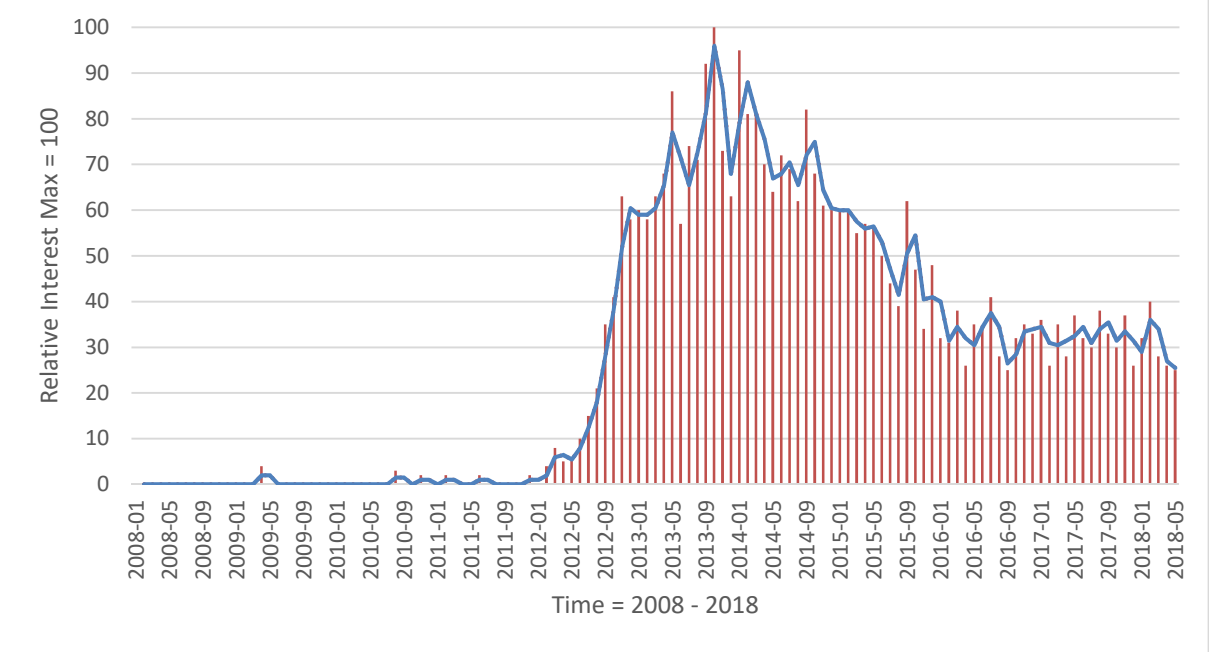
* Gartner Hype Cycle for emerging technologies plus the MOOC such one them.

By collecting data from Google Trends between September 2012 and September 2014, Klobas (2014) researched about MOOCs and generated a line graph to identify the current stage of MOOCs in the hype cycle. During 2012, there was a growth in searches about the theme, and continued to grow until 2014. The insights to be gained from this result are consistent with the recent publications on the subject, discussing the benefits for users, the value intangible to be generated for the user, the barriers to be overcome and the search for factors that support a model of effective learning (Terras, & Ramsay, 2015).

The same procedures of Klobas (2014) were replicated, but in this time, it's covered the period from 2008 to 2018. The year 2008 was included to show the period of low popularity

and to counterpoint the explosion of popularity in 2012. Graphic 1, generated from the Google Trends data, shows the trajectory of MOOCs over this period.

Graphic 1: Google trends for term: Massive Open Online Courses



Source: Elaborated by author with data obtained in Google Trends, 2018.

It is noticeable when observing the Graphic 1 that MOOCs began to gain popularity in 2012 and peaked in late 2013 and early 2014. The popularity scale generated by Google Trends obeys a scale of 0 to 100, with 100 being the maximum value of popularity for the period researched. By referring to the Hype Cycle and according to the current discussions on the topic, the trend line, created through the moving average measure, shows that the MOOCs are in the disillusion step, but gives evidence that it is beginning to come to the fourth stage: slope of enlightenment. The Figure 1 supports the position of MOOCs encountered in Graphic 1 and is coherent with what is being studied on the subject since 2012.

2. Theoretical Review

This topic seeks to present the key concepts associated with MOOCs, since their epistemological conceptualization to their main characteristics and the context of the emergence of MOOCs. It also aims to present the view of literature in certain points: criticisms about the subject, implementation barriers and, the business model that is being developed since 2012 by platforms and professionals.

2.1 Definition of MOOC

The MOOCs appeared to provide unique opportunities in education. To be open and free, an Internet connection is enough to take advantages of these benefits. Notable examples include Coursera, Miríada X, Future Learn, Khan Academy, Udacity, and many others. The main argument in favor of MOOCs lies precisely in the ease of anyone to learn, anytime and anywhere (Merino et al. 2015).

Since 2011, the MOOCs became reality and proved to be a revolutionary force that came to add to the traditional model of higher education (Gillany, & Eynon, 2014). Being one of the most prominent trends in education, the "M" of the term means massive, global, borderless. The "O" means it is open and free. The "OC" means online course, since it comes to distance learning with use of web technologies (Baturay, 2014).

The milestone that began the popularization of the term occurred in November 4th of 2012 when the New York Times published an article entitled "The year of the MOOC", by Pappano (2012). The enrollment and the number of courses in this period grew exponentially, according to Chen and Chen (2015). At the end of 2014, only Coursera, there were more than 10 million entries in 840 courses offered by 114 institutions. The average enrollment per course, exceed 20,000 people, with the most numerous reaching 230,000 users.

These numbers in the end of 2015 had risen to 119 institutions and over a thousand courses being offered, involving the participation of more than 13 million registered users. The impact of MOOCs not only attracts users, but also investors. It is estimated that by the beginning of 2016, more than 85 million dollars have been raised (Belleflame, & Jacqmin, 2015).

And the impact of MOOCs is not revolutionary only to be open and free, but also rekindle varied teaching modes in schools and mainly in universities. Besides that, they encourage new teaching methods, with automatic correction exercises, group attendance, videos with

theoretical content and use of social networks to promote discussion groups and exchange of knowledge (Merino et al, 2015).

The MOOCs are defined as a type of online education or as the current state of e-learning. The e-learning is the most usual configuration of distance education (Abeer, & Miri, 2014). The MOOCs are also allocated as an emerging technology and a disruptive technology in education (Alraimi, Zo, & Ciganek, 2015). In the social sphere, the MOOCs are a symbol of the so-called: knowledge society or the digital age, together with the popularization of the Internet and others ICTs (Steffens, 2015).

The MOOCs are also associated with long-life learning, which lies in the idea of people learning not only in educational institutions, but also outside of them, in formal and informal settings, giving equal importance to the collective study and individual study (Steffens, 2015).

Lin, Lin and Hung (2015) cite the open education as one of the principal concepts related to the subject, explaining that MOOCs have not only advantages associated with costs and physical distance, but also impact on the knowledge of those who want to learn. With this argument, say that the MOOCs are a superior model of education compared to earlier stages of e-learning.

The foundations of MOOCs reside in connectivism, being a philosophical theory created in the early years of the decade 2000. The generativism also helps to explain collaborative learning. Still on philosophical movements, Spector (2014) attributes characteristics of the courses, as deadlines or well-defined learning objectives, to the behaviorist approach. Phelan (2015) associated web-learning with constructivist and objectivist approaches, on account of unlimited access to information and resources via web, to develop the ability to share ideas with a large audience and therefore increase user satisfaction. The concepts of connectivism and generativism will be explained in section 2.3.

Synthesis: The MOOCs are free and open courses, offered via web, and comprehend the fourth phase of e-learning, being in this phase, currently. Is one of the main trends on education, especially on higher education. It is considered a revolutionary and disruptive technology, being also a symbol of the knowledge society in digital age. Their epistemological bases are connected with connectivism; however, the philosophy of MOOCs cannot be exclusivity explained through connectivism.

2.2 E-learning evolution

One of the most important aspects of digital technologies is that exists potential to support learning. The long-life learning, being a topic widely discussed in the last century, it is still being discussed under the focus of new forms of education. Since the early 90s, education has undergone many changes, potentiated by web technologies. Progress in socialization, knowledge and spirituality, came under discussion with the study of e-learning services. These cultural aspects cited by Steffens (2015) were responsible for the study of behavioral variables as determinants of the effectiveness of e-learning.

The first explorations of web technologies focused on paid courses accompanied by tutors. Soon, the first web-based learning initiatives appeared, as well as the first Virtual Learning Environments (VLE). The use of web technologies in education was called e-learning, being a new distance learning mode. However, Nicholson (2007) points out that web-based technologies are the third stage of e-learning.

Phelan (2015) corroborates Nicholson (2007) to define the e-learning as the use of information and communication technologies, which include various forms of educational technologies, such as learning via the web. Cheng (2012), bringing together the definitions of several authors, has a consistent definition as Phelan (2015), conceptualizing e-learning as a tool, that uses the Internet through electronic media to help organizations deliver study materials for the users, using web-based communication to support users and provide the service to anyone, anywhere and anytime.

According Nicholson (2007) and Pereira (2013), the e-learning can be divided into 4 stages, and currently, we are living the 4th stage with MOOCs, as shown in Table 1.

Table 1: Historical phases of e-learning

Phase	Era	Center of interest	Features
1°	1975 – 1985	Programming; Computer Assisted Learning (CAL)	Behaviorist approach to learning and instruction; use of programming to build and solve problems
2°	1983 – 1990	Computer-Based Training; Multimedia	Use of old models, such as CAL, with multimedia courseware. Projects more intuitive and educational software focused on user
3°	1990 – 1990	Web-Based Learning	Internet content delivery; limited interactions to the end user
	1995 – 2005	Web-Based Learning expansion; Virtual	Internet based on flexibility, increased interactivity, online courseware and end-user interactions

		Communities; Social Media	
4°	2005 - 2007	Mobile Learning; Distance Open Education	Learning from mobile devices; open distance learning courses, with web 2.0 use and social networks
	2008 - 2016	MOOC (Massive Open Online Course)	High quality content in platforms with open access and free

Source: Adapted of Nicholson (2007) and Pereira (2013).

*Created by Nicholson (2013), extended by Pereira (2013) and extended again.

As can be seen in Table 1, the phases are marked by the advance of new technologies and the maturing of web-technologies. Studies on the subject emphasize the direct relationship between the evolution of web technologies and the development of e-learning in their 3rd and 4th stage (Chien, 2012). The author also states that the focus of several studies is in the transmission of knowledge, from a technological perspective.

The e-learning can be synchronous or asynchronous. Synchronous refers to when the material is presented in real time; on the other hand, the asynchronous refers to classes stored and available at any time. A hybrid model or called blended learning, concerns the combination of the two (Phelan, 2015). Cheng (2012) also shows the classification of e-learning systems that support e-learning and are management systems, including Learning Management System (LMS), Learning Content Management System (LCMS), Learning Design System (LDS), and Learning Support System (LSS). These categorizations allow an e-learning service can be managed in 4 areas.

The e-learning is used in universities to promote distance learning courses, but they also adopt blended learning, taking advantages of technological solutions for education and thus, diversifying the types of education offered to students (Wong, Tatnall & Burgess, 2014). Some authors divide the use of e-learning services for companies and educational institutions, terming the e-learning in business as an e-business tool, that is, the use of online technologies to support activities. However, despite different objectives, the service is the same (Wong, Tatnall & Burgess, 2014).

The objectives of an e-learning course are to promote knowledge, develop behaviors and attitudes of students. This set of characteristics determines the effectiveness of the courses. These characteristics come from the user relationship with the course and also the advantages that user can get to do the course. Thus, Chien (2012) names four levels of evaluation of users, namely: reaction, learning, behavior and results. This approach has an educational direction.

Under a direction of Information Systems (IS), the evaluation is focused on the Virtual Learning Environment (VLE), and more: acceptance, use and re-use of technology. To achieve the success of e-learning, behavioral aspects and information technology infrastructure are equally important. Among the most important aspects, Chien (2012) presents two approaches that seek to define the factors that best determine the effects of e-learning use: approach of Putch and Lee, which list the factors of functionality, interaction and responsiveness; and approach of Lord and Volery, which list ease of use, interface and interaction. Other behavioral and structural factors may also influence the e-learning in some way, such as quality, self-efficacy, reliability and performance.

Synthesis: E-learning is a form of distance learning that provides an education service with the use of technology platforms. This teaching mode can be synchronous (real-time service), asynchronous (available content at any time) or hybrid, when use both. The MOOCs account for the current state of e-learning and its main objective is to promote knowledge through the effective management of the offered services and the provision of high quality content.

2.3 Epistemological view of MOOCs

Connectivism was introduced to the world through publication of an article called "Connectivism: a learning theory for the digital age". The main argument of article discussed about classic theories of learning, such as behaviorism, cognitivism and constructivism, and how these epistemological chains were outdated in the world we live in 2018, since the way we learn is different. Siemens (2005) defined the term as the integration of explored principles, which drinks in fonts of chaos theory, network theory and complexity and self-organization theories.

The term connectivism gained certain notoriety in 2008, when George Siemens and Stephen Downes developed what they called Connectivist Distributed Learning Model. In 2011, the production of educational videos was aggregated into online platforms supported by web resources. From there, the MOOCs not stopped growing (Baturay, 2014). Siemens explained that the term MOOC is inspired by the online RPG games, being called Massively Multiplayer Online Role Playing Games - MMORPGS (Greene, Oswald & Pomerantz, 2015).

According to Siemens, the advancement in ICTs dramatically changed the field of learning and the knowledge production. The theory or connectivist chain indicates that each individual is responsible for his/her learning, and that MOOCs can be considered extensions of a personal learning environment (PLE) and personal learning networks (PLN). The goal of connectivism is to connect learning theories to the digital age, under an educational point of view (Lindsey, Rhoads & Lozano, 2015; Kesim & Altinpulluk, 2015).

Clarke (2013) argues that Siemens and Downes seek to show and suggest what has changed on what we learn, how we learn and where we learn. In this statement, there are contrasts between connectivism and other theoretical chains, just for trying to add characteristics from various theories. In behaviorism, there is a stimulus response relationship. In cognitivism, learning is achieved by a structured process and, in constructivism, learners are engaged to participate in social and cultural exchanges. Thus, connectivism seeks to add these principles.

From this idea, there are authors who dismiss the concept of connectivism as a current of thought, however, do not rule out the idea that will be a theory in near future (Clara, & Barbera, 2014). Some authors argue that there are not enough criteria to consider the connectivism a theory of e-learning, but do not show what criteria are sufficient to be considered (Lindsey, Rhoads, & Lozano, 2015). What is fact, is that the connectivism has proven to be consistent in explaining the inclusion of emerging technologies in education and regardless of to be or not considered a theory or school of thought, contributes strongly to represent the epistemological field of this phase of e-learning (Clara & Barbera, 2014).

From a philosophical point of view, Steffens (2015) points out Siemens's work for some principles that guide learning through MOOCs: first, Siemens and Downes differentiate learning and knowledge. About learning, Siemens (2005) defines it as a process of connecting specialized nodes or source of information. To defend this concept, he adds that learning can reside in non-human devices; the ability to know is more critical than knowledge at the last five years; the ability to establish connections is a key skill; and decision-making is itself a learning process.

From this concept, Clara and Barbera (2014) defend the idea that the e-learning paradox applies to connectivism. The paradox is the gain of knowledge: how can anyone know anything? If the student did not know before, how can he recognize this something? If he is able to recognize, it is because some knowledge was already known before. Learning theories try to solve this paradox by different approaches.

There are two solutions to the paradox. The first formulated by Socrates and Kant is that the human being at birth already has everything you need inside mind to learn. That thought led to innatist psychology, with different approaches, such as Fodor's approach, which sees the mind as a complex computational tool, and Piaget's approach, which defines the development of the mind through balancing and centering.

A second solution, initiated by Hegel, was developed in psychology area, from the Zone of Proximal Development (ZPD), and after years, was incremented by Vygotsky. The idea lies in the matter of the mind that, by Vygotskyan vision, can be shared by two or more individuals so that this knowledge can be used together in the learning process (Clara & Barbera, 2014).

Apparently, connectivism seems to apply to the paradox, it is defined as a knowledge network. However, Downes points out that the standard of this network needs to be recognized by the individual, so, the paradox persists. The Vygotskyan solution seems to apply best to connectivism. But Clara and Barbera (2014) highlight a contradiction: since connectivism depends on the ability of the individual to recognize a pattern of connections to generate knowledge, how the connectivism explains the ability to recognize patterns?

Even with the proposed contradiction, there is a consensus that the interaction enhances the learning process through a standard connection. Deepening interaction, Lindsey, Rhoads and Lozano (2015) highlight the constructivist contributions in understanding the connections. Connections or established social network are contextualized and derived from a social, cultural, economic and physical environment. The notion of collaborative learning depends not only on the link between individuals, but also of the environment that surround them.

Lindsey, Rhoads and Lozano (2015) state that the Vygostskyan notion of online collaborative learning establishes a strong sense of community among users of MOOCs, this sense results in high levels of satisfaction, persistence, engagement and readiness to experience technologies. Knowledge is built through collective understanding of the contribution of ideas and debate, in formal and informal environments, which form a critical thinking.

On the one hand, there is consensus about the increase in the learning process through the connections, although there is still distrust about the source of knowledge in this process. Therefore, one of the newer chains that are becoming space is the generativism, developed by Carneiro. Steffens (2015) states that the basic argument of generativism is that OERs can be explained by classical theories such as behaviorism, even in a learning environment driven by technology. The generativism still criticizes focus of connectivism on the individual, since it defines the learning process as a generator of knowledge, and the individual's role is only to bring previous knowledge, which can be transformed and gain new meaning through new

experiences, that is, is changeable. From this point of view, collaboration between MOOCs users has a higher importance, if compared to prior knowledge of one individual.

Synthesis: Connectivism is a current of thought that guides the epistemological view of MOOCs by integrating principles of various classical schools of thought, to conceptualize learning and knowledge, highlighting the role of technologies embedded in the learning process. Despite the key role, critical to connectivism resides in two aspects: the failure to present a solution to the paradox of learning, and the fragile concept of interaction, by presenting it as a binomial and static relationship. From the connectivism, theoretical advances have been noticed, focusing on connectivism's weaknesses and focus on the evolution of MOOCs since its inception.

2.4 cMOOCs and xMOOCs

The first MOOCs were those associated with connectivist distributed peer learning model, and are called cMOOCs, being an abbreviation of connectivist MOOCs or Canadian MOOCs, because the creators of connectivism and the first MOOC are Canadian (Kesim & Altinpulluk, 2015). Basically, a cMOOC is a wide network of people and resources online, where course materials are derived from the experience of students during the course (Baturay, 2014). These courses differ from xMOOCs basically in pedagogical base, that underpins the cMOOCs and economic base that underpins the xMOOCs (Clara, & Barbera, 2014).

The term xMOOC is an abbreviation for extended MOOC. The teaching methodology in xMOOCs is based on behaviorism, which involves the transmission of information, peer assessment and computer-marked assignments. The x-MOOCs are designed for mass education through free or paid courses, and certified or not. Thus, most currently MOOCs, which are offered by organizations that seek profit as Coursera, EDX, Udacity, among others, are considered xMOOCs (Shen & Kuo, 2015).

The cMOOCs were established before, and with the development of MOOCs under management platforms that aim to profit, there was a need to differentiate the two types (Kesim & Altinpulluk, 2015). Siemens emphasizes that cMOOCs have the goals of fostering creativity, autonomy and learning in social networks. The model of xMOOCs focuses on duplication of knowledge, with the use of a traditional model of learning, which brings a pre-defined content by an instructor, bringing benefits for those who do not have easy access to quality materials.

Despite the differences, both types use the tools of social media and VLEs during the learning process, because they facilitate communication and interaction among users. Other technologies via web are also used for the same purpose (Shen & Kuo, 2015). The technologies facilitate both structures courses and both have the objective of ensuring the effectiveness of the courses, to ensure user satisfaction (Alraimi, Zo & Ciganek, 2015). There is also similarity in feedback from reviews, to have encouraging collaboration in the correction process, that is, the evaluation system is collaborative and can be inserted as another form of interaction and collaborative learning (Shen & Kuo, 2015).

Synthesis: The MOOCs can be divided into two categories, the cMOOCs, conducted under connectivism and interactive content created throughout the course, and xMOOCs, conducted under a traditional learning, with pre-defined content. Most courses are normally xMOOCs and are offered by platforms that aim profit, supported by major universities that offer high standard of quality in content.

2.5 MOOCs challenges

The MOOCs refer to a new teaching mode that provides learning to a wide range of people. Certainly, this type of education has valuable resources for students who can access them. It is easy to imagine the educational potentialities and economic interest in this learning model. In counterpoint, it is hard to imagine how this model can go wrong (Ventura, Bárcena & Monje, 2014). There are several researchers skeptical about these capabilities, certain that the great attention given by MOOCs generates high expectations of the institutions and companies who invest in it.

First, the MOOCs user profile is predominantly formed by graduate users, employees, from developed and developing countries, and men (Greene, Oswald & Pomerantz, 2015). Knowing this profile, is natural to think about the real impact of MOOCs on digital inclusion and quality content offering in geographic environments where there is no similar quality. This point causes many people deem to have less impact on educational community than meets the eye (Ventura, Bárcena & Monje, 2014).

In fact, the courses offered are coherent with what has been studied in large universities, in other words, the content of the courses is level over. The interest that awakens the users comes from a prior knowledge on the subject of the course, and it is natural that the most interested users, and with the set of prior skills to make it, are closer to universities that offer

high quality education. What can be understood as a challenge is how to make attractive courses for unemployed participants from developing countries and for more women (Steffens, 2015).

Changes in the dissemination of MOOCs are also necessary. The focus of advertising is the online access by marginalized people who need a little help to boost the entry into higher education and to increase the knowledge, which otherwise would not be possible to obtain (Spector, 2014). Disclosures of this nature are inconsistent with the user profile existing currently.

Another point is the argument that the completion rate of the courses is very low, ranging from 5% to 15%. In traditional courses, these indices represent a failure. So, regarding thousands and some courses with hundreds of thousands of users who sign up, there are few who can complete the courses. For some authors, this is not surprising, since there are passive users, with different objectives and that few ones have the major goal of completing the course (Greene, Oswald & Pomerantz, 2015; Perna et al 2014).

Alraimi, Zo and Ciganek (2015) support this view, stating that the completion rate is not the best measure to evaluate a MOOC and cannot be associated with their effectiveness. Radford, Coningham and Horn (2015) also corroborate this position, bringing information about the preferences of users. The authors report that the majority of people join to the course to have more knowledge about a familiar theme, each one with different objectives. Besides that, only 13% of subscribers are intended to receive a certificate. These ones can be considered active users.

Perna et al. (2014) argue about the definition of active and passive users regarding MOOCs. Although there is no agreement on these definitions, four user groups can be identified: those who register, but at no time enter the area of the course, those who access less than half of the materials of the course, those who are just explorers and seek specific materials and those seeking the certificate to access all materials, complete quizzes, meet the deadlines and comment in the forums. Perna et al. (2014) stress that it is a challenge for researchers and managers draw conclusions on evaluation studies, without knowing which group each individual belongs to.

One typology listing the goals can separate active and passive users, and from there, it becomes possible to check indexes for each user type and determine the success of a course. From typologies proposed, Greene, Oswald and Pomerantz (2014) point out that they can be found in various aspects of the course, from the curious, those who seek only information, a specific class, learn a little more about it, even those who sign up aiming to study all the modules and get the necessary points to secure the certificate.

Another point of debate is on social and educational merit, being questioned constantly, starting the growth of xMOOCs. The criticism is based on possible conflicts between the educational goals and objectives of capital invested, compared to production of open source platforms and meet the demands for courses (Klobas, 2014). Skeptics say that these constant conflicts can break higher education institutions (Abeer & Miri, 2014).

Some settings about the MOOCs are also batting, such as the wave of e-learning, suitable for digital and online age tends every day to become a superior experience to any type of education face to face (Clarke, 2013). Opponents resist this idea, stating that the modalities of distance learning in general are threats to higher education, and that may affect the area of research and accreditation (Abeer & Miri, 2014). Another very hitter definition is that this new phase of e-learning and long life learning is a disruptive technology, not to revolutionize the education area, but to be a substitute for traditional teaching model (Alraimi, Zo & Ciganek, 2015).

Under the psychology point of view, Terras and Ramsay (2015) address some challenges to be pursued by MOOCs, such as differences in abilities, preferences and user profiles, that can hamper the fulfillment of all learning needs; motivation and engagement are essential for better performance, understand how these feelings can be obtained is critical; there is high demand for more skills of monitors to meet the demands in formal and informal settings, it decentralizes the role of instructors and gives more responsibility to the monitors.

Spector (2014) presents another controversy, based on the content of the courses, saying that many MOOCs not undergo to evaluation and test of content before being offered, fame and charisma of an instructor may cause a bias in the evaluation of quality course, and there even greater deficiency of a MOOC can basically be a collection of knowledge and learning objects, which may be inconsistent to be transmitted that way. Spector (2014) points out that a typical MOOC cannot present these shortcomings.

Phelan (2015) suggests the notion that there are more effective areas with use of MOOCs and e-learning in general; however, there are areas where teaching face to face is more efficient and can achieve higher levels of engagement. The author maintains the idea that a greater diversity of teaching modes also diversifies teaching strategies.

Synthesis: In order the MOOCs to reach a point of maturity, it is necessary face challenges, which are evidenced by well-founded criticism, such as: the lack of diversity in the user profile; the low level of completeness and incoherent notion that most completion rate

means greater effectiveness; the conflict between educational goals and management objectives; the notion that MOOCs are substitutes of traditional teaching model and consequently are threats to universities; insufficient actions to plan and prepare a MOOC course, among other challenges.

2.6 MOOCs in 2018

The MOOCs are relatively new, and many people still have a certain distrust of the potential of this technology in long-term. A notion that still foments distrust and that endures from previous phases of the e-learning is that conventional classes would be replaced. In fact, the e-learning tools can increase access and can provide greater flexibility for students (Phelan, 2015).

Steffens (2015) argues that empirically, there is no doubt as to support learning, but the advantages of being massive are not so clear. In universities, there were large investments in Open Education Resources (OER) and provision of online courses, establishing an industry that is growing exponentially in recent decades (Clarke, 2013). These investments are driven every day by the emergence of new interactive and collaborative technologies.

OER can be defined as educational resources, being offered online and free to educators, students and self-learners, to facilitate learning (Lindsey, Rhoads & Lozano, 2015). The MOOCs add the OER and are developed with similar interface to digital libraries, offering specific courses and topics in list, to search for a subject, in videos, video's transcription, slide presentations, among other resources (Zhuhadar, Kruk & Daday, 2015).

Spector (2014) argues that these investments follow a natural course of investments in distance education. The education involves systematic development of knowledge, skills and attitudes. These three characteristics cause persistent changes in individuals. The relation with technology occurs because the technology is understood as a systematic application of tools that facilitate learning.

The MOOC designation was given by offering an online open course about connective knowledge or connectivism, developed by the University of Manitoba, by George Siemens and Stephen Downes, with 2200 participants around the world. The course consisted in connectivist teaching model, where learning is understood as knowledge acquired through the use of technological resources and individuals networked (Margaryan, Bianco & Littlejohn, 2015), or also as a persistent change in human performance or potential performance as a result of experience and interaction during the course (Steffens, 2015).

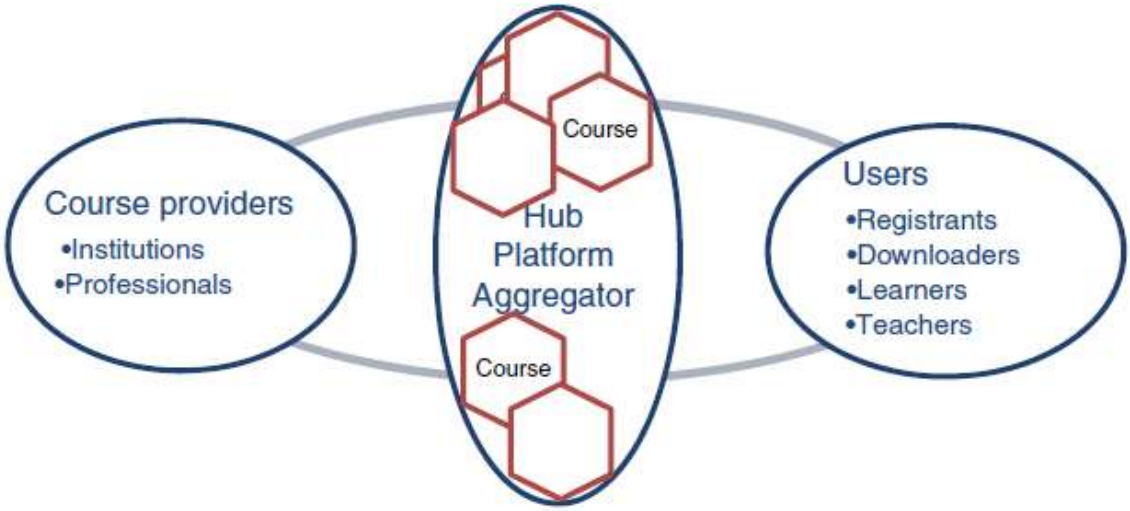
The first MOOCs were decentralized, and they had a nonlinear structure, focused on exploration and conversation, rather than the broadcast content. This constructivist mode of MOOC began to popularize in mid-2012, when large universities started offering MOOCs via commercial platforms, such as Coursera, Miríada X and Udacity. From the growth, these courses were becoming hyper centralized, with multiple-choice tests, linearity in monitoring the content and monitoring of progress topic by topic.

The offer was intensifying, until it consolidated the importance of MOOCs and majority of people have the understanding that MOOCs represent the greatest advancement in higher education at the beginning of XXI century (Margaryan, Bianco & Littlejohn, 2015). It is consensus that the content offered has best quality, because the course is planned for the typical student of the great universities and can suit the programs of disciplines as complement or replacement.

For universities, there are advantages to diversify teaching face to face, under the educational point of view. Besides being a trend in major universities, enriches learning with intelligent use and innovates in learning, because universities can develop different approaches to blended learning, and take advantage of the proliferation of technologies and software's that allow us to offer advanced MOOCs, which can operate with efficiency and together with tools, such as Twitter, Youtube, Google Drive, Google Academic, Wikipedia, Evernote, Moodle, Slideshare, Prezi, Facebook, Pinterest, among others (Clarke, 2013).

It is necessary to identify who are the actors who make the MOOCs, those who are connected to a MOOC from planning and implementation until those that determine your success. Klobas (2014) seeks to identify all the participants of this process, and divides these actors into two groups, as shown in Figure 2. On the one hand, there are the course providers, comprised of technical professionals that assist in assembling the course, and also formed by institutions providing financing and academic support. On the other hand, there are users, consisting of classes of passive and active users; in addition, teachers and monitors complete the second side. In the middle, there is a platform where the course is hosted.

Figure 2: Actors who made MOOCs



Source: Klobas (2014).

Universities are an integral part of the MOOCs market, and consequently the MOOCs are an integral part of higher education; it is through it that the educators provide high-quality content. Ossianilson (2012) points out that even confirming the important role given to MOOCs, there are few indicators to monitor the quality of courses in the educational field and management level. Wong, Tatnall, and Burgess (2014) add that the management of e-learning services should be based on balancing academic management and administrative management, taking into account the readiness level of students to adopt new technologies.

Cai and Zhu (2012) present numerous indicators that may be barriers to the motivation of students to join a course in e-learning. These barriers are common to MOOCs and other e-learning services. Among them may be cited: lack of dedication of time, high volume of materials to be studied, language barriers, exacerbated difficulty of topics, among others.

In addition to the actors who play the MOOCs, Scott (2014) supports the idea of considering spatial aspects, which should no longer be treated as a theoretical construct and become a social reality. This new vision allows universities to manage courses more accurately measuring the real impact of the absence of geographical limitations and time, and going to consider global alliances with companies and other institutions.

Universities must make strategic decisions to: promote own MOOC, ensure that the content with larger impact on society is being transmitted in MOOCs platforms, invest in blended learning to refine teaching at the university, emphasize the qualities of teaching face to face for the students, among other actions (Clarke, 2013).

Outside these decisions, it is necessary to consider that not all Open Online Courses (OOCs) are massive. It may be more interesting to offer a course at a lower cost in certain areas,

such as learning technologies for certain age groups. Courses of this nature can be more sustainable and less costly to meet the participants and also with the potential to generate revenue through additional services such as accreditation. Some courses are also called Social Open Online Courses, still called the Scalable Open Online Courses (SOOCs), which focus on the development of collaborative social learning strategies, without the need to be massive, expanding the teaching strategies and management not encompassed by MOOCs (Porter, 2015).

Other MOOCs or SOOCs have the possibility of not being open. A variety of these e-learning modalities are the Small Private Online Courses (SPOCs), where access is restricted to a limited group of users, which can range from hundreds to thousands. These adjustments to the MOOCs result of the different strategies of universities to sustain the continuity of e-learning service, and also to monetize the service (Porter, 2015).

Synthesis: there is high demand for consumer collaborative technologies and access to the best universities content; however, there is a distrust of the best way to promote a MOOC. A consensus was being formed regarding the advantages promoted by MOOCs, from its inception to the intense use of these online services. The actors that participate in MOOCs are divided into two groups: technical teams and institutions on the one hand, and users of educational platforms on the other hand, comprising students and teachers. For universities, there are different strategies for how to offer a MOOC to ensure the motivation of users, and to obtain international visibility, or even to profit from the courses.

2.7 MOOCs Business Model

The emergence of OERs, the OOCs and consequently MOOCs became reality on educational area. Through open platforms, free to enrollment, with open and integrated with social networks curriculums, the MOOCs operation mode differs as to cMOOCs and xMOOCs (Kesim, & Altinpulluk, 2015). The cMOOCs have low ability to monetize their services; however, the xMOOCs have possibilities of exploring a business model (Ma, Zheng & Zhao, 2014).

The MOOCs are considered one of the most intriguing and hyped developments in higher education, especially about the question of what is the business model of MOOCs? There is high demand, no doubt, but it is controversial how this e-learning service can be sustainable. In

the current context of the MOOCs market, some models may be promising, with economic feasibility to implement these courses.

In first place, the higher education is experiencing a new financial reality. Previously, the government was responsible for funding much of the budget of the universities, and it has been decreasing over the years; moreover, the need for investment in technology in universities has only increased, and consequently, increased the costs dramatically. These characteristics are causing radical changes in the business models of universities. The online courses, best exemplified by MOOCs, have democratized higher education, together with the demand for a workforce that need to continuously recycle. The reality presents a scenario of global educational competitiveness, with services offered by universities, and even by other companies (Clarke, 2013; Behara & Davis, 2015).

Behara and Davis (2015) identified three major problems of this new scenario: including the need to obtain return on investment in online education, promote changes in education delivery and still deal with the dissatisfaction of employers who may become dissatisfied with the performance of users on the workplace, giving low income on the teaching model. Indeed, it is difficult to make changes in a model that has remained unchanged for centuries. The impacts with these changes can be catastrophic if poorly planned (Clarke, 2013).

There is still no clear business model for MOOCs. The providers are still exploring potential revenue opportunities (Porter, 2015). Leaders in the market which offer xMOOCs, have a revenue sharing arrangements, making agreements with universities, which pay a fee for support and develop the courses, profits can be generated by specific course, charging for subscriptions and certificates. The courses offered follow a pattern established by the platform on which they are hosted (Baturay, 2014).

The business model for universities seek to obtain revenue only enrollment and paid certificates, with the addition of some funds and donations. However, universities are showing negative income, with losses between $ 700 million and $ 1.5 billion in operating costs (King & Sen, 2013). The question that remains is: who will pay for MOOCs in the future?

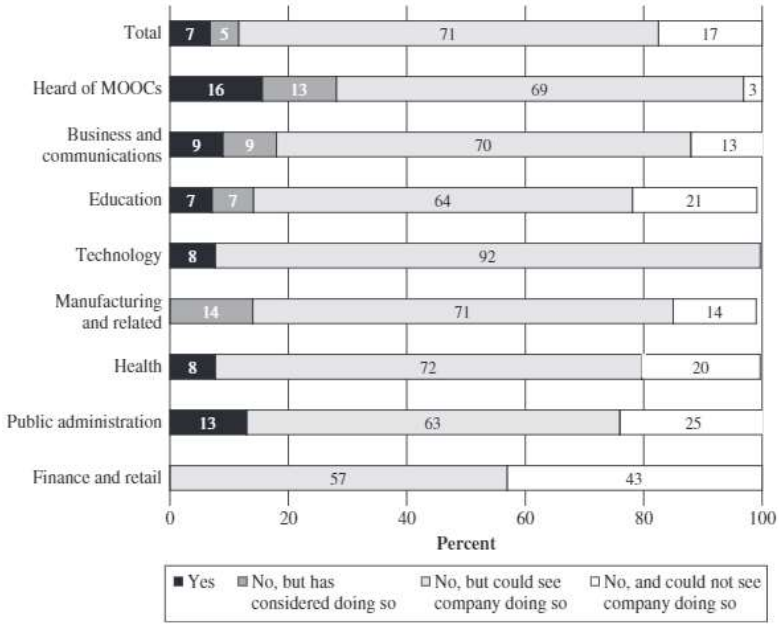
It is necessary to explore the different products that MOOCs offer, which can be sold together or separately. It is necessary also think about what can be monetized and what cannot, to be considered a MOOC (Porter, 2015). From a purely economic point of view, the MOOCs platforms present a very low marginal cost per student; so, invest in a platform removes the need for universities to make constant payments to the platforms. Universities can also get government financial benefits, while delivering to society a digital inclusion tool in which

everyone can access and get quality content, regardless of the social profile that the individual presents (Belleflame & Jacqmin, 2015).

What is clear to Clarke (2013) and Terras and Ramsay (2015) is that universities should continue to provide classroom courses with the same volume, and increase the supply of blended learning, while maintaining educational quality and the ideals that underpin the essence of university. Regarding MOOCs, they should promote marketing campaigns with advertisements that represent reality, and they should direct the focus of MOOCs for certain niches, such as employees, by providing courses that can enhance employees' basic training. Other strategies that have been used are avail the credibility of the university, including in certificates paid, the mark of the university, and also, offer combined courses in a program or specialization in a subject (Terras & Ramsay, 2015).

About these strategies, Radford, Coningham and Horn (2015) report the research conducted at Duke University in the early 2014 with managers on the intention to use of MOOCs for professional development, to improve the performance, engagement and productivity. The search results are shown in Graphic 2, with the percentage rates. According to Chart 2, about 7% of organizations use MOOCs for professional development, 5% have considered using them, and 71% see their companies making use of MOOCs for training in the future. When asked whether they had heard about MOOCs, only 3%, that already became familiar with the subject, said they have no intention to use MOOCs, in the hypothesis the company to adopt practical training. The results are also presented by the area of the company in which the employee works, having most effective use in public administration, with 13 percent.

Graphic 2: Preference for use MOOCs for professional development



Source: Finding and Developing Talent Survey, by Radford, Coningham and Horn (2015).

Clarke (2013) states that these initiatives and strategies are still uncertain because the MOOCs are still under development. What is known is that a MOOCs business model must meet those who offer the platform (Coursera, EDX, Future Learn) and those who create courses (universities). These two providers are closely linked and depend on each other so that the continuity of MOOCs may be viable in the future (Porter, 2015). Radford, Coningham and Horn (2015) complete this argument by claiming that the determinants of the effectiveness of MOOCs have not yet been clarified, and organizations need to experiment with innovation to develop different ways to ensure benefits to stakeholders in MOOCs.

Synthesis: Among the cMOOCs and xMOOCs, the second type of MOOC has more chances to monetize the e-learning service offered. Higher education is undergoing dramatic changes in its business model and MOOCs are an integral part of a new model still uncertain. In addition to the existing revenue sources with enrollment and paid certificates, different monetization strategies should be experienced by providers of MOOCs.

3. Systematic and bibliometric review

This chapter presents the part of literature review focused on selecting studies that provide the theoretical support necessary for creation of the theoretical model of this research. In this way, the concepts of the review types, the literature review process adopted, and the results of the papers analyzed are presented.

3.1 Concepts about literature review

The research propose is to develop a reproducible scale specifically for MOOCs services. To achieve this objective, a discussion on what is being studied on MOOCs in 2018 and the evolution of these studies over a recent period (ten years) is promoted. A bibliometric and systematic review with meta-analysis was conducted to assist in the identification of items, constructs and methods of analysis more useful and viable.

To present a research model, the literature review should look into the studies, building a knowledge about the evolution of the theme over time, as the general and specific aspects. A model can be defined as a language that allows to observe the form of hypotheses and then test them empirically. The creation of hypotheses is a result of a literature review. Regarding MOOCs, it is looking for a model that focuses on the perception of the users, with a use of behavioral variables.

There are three types of literature review: narrative or traditional, bibliometric and systematic. The narrative review is the traditional literature review, used to describe the history and theme development, as well as to discuss the topic at the theoretical level. However, Pocinho (2008) emphasizes that narrative review does not provide quantitative answers to specific topics, mainly for not having sources or method selection well exposed. The concept and stages of literature review are described as it follows.

<u>Defining bibliometric review</u>

An important step in all research is to disseminate the results found so that other researchers can confirm, replicate to other social realities and build an improved reality from this (Sjostedt, Aldberg & Jacobsson, 2015). In order to operationalize this premise, it is

necessary count, measure, compare, quantify and analyze data in a general way. With the large volume of scientific researches, publishing and disseminating knowledge has become bigger and complex (Pendlebury, 2011). In this context, bibliometric analysis fits in and is gaining more space in scientific researches, because it has a selection method well defined and high synthesis capacity.

The bibliometrics can be defined as a set of laws and empirical principles, being used to analyze the path followed by publications on the subject (Pritchard, 1969). With a bibliometric research, it is possible to characterize the current state of research on the topic, summarize the accumulated collective knowledge and identify gaps that can be exploited (Spinak, 1998). In a few words, the bibliometric analysis produces statistics from data, describing the productivity or the impact in form and frequency of publications (Sjostedt, Aldberg & jacobsson, 2015).

In bibliometric research, there are three types of results that can be generated from a search of articles in indexers. These results can be categorized in: general characterization of the work, with the identification of the authors, journals and indexers; evaluation of the content, involving the identification of sub-themes and related areas; and evaluation of references, including varying results about citations and identify seminal studies on the subject. The results raised were treated by descriptive statistics, assigning frequency analysis.

A controversial point in bibliometric research is to define what constitutes large-scale publication and what justifies the use of bibliometric analysis. From previous experiences Sjostedt, Aldberg e Jacobsson (2015) aggregate indicators based on citations and affirm that the sum of aggregate products does not exceed more than 50 articles. However, the authors point out that there are specific needs, depending on the research, and the important thing is that it is feasible to analyze all the selected articles. Pendlebury (2008) also appoint that human judgment does not overlap with what quantitative analysis offers. The final decision on which articles to select and which ones will be best used does not escape a human judgment and the relation with the purpose of the research.

The same author has listed ten rules about analysis of publications and citations. Of the ten rules, some of them facilitate the understanding and justify the joint use of systematic review with the bibliometric review, such as the selection of sources. In this case, the selection of indexers, where the articles search will be done. This first step is a practical application of Bradford Law, which refers to the dispersion of scientific periodical literature, that states that if scientific journals are sorted in crescent order of articles productivity on a particular subject,

they may be divided in a core of journals more particularly dedicated to the subject, or in several groups or zones, containing in each group papers that share resemblances.

The rule four of the ten proposed by Pendlebury (2008) concerns the most cited publications. In this study, this rule is followed in an adapted form, because the importance of a paper is given, not by the number of citations, but by the studies with quantitative approach and that present theoretical models validated from a research instrument presented in the article. The judgment of the paper quality invariably goes through a subjective evaluation, but with the use of systematic review, the evaluation can be done in a standardized way and with rules well defined.

Defining Systematic review

By the definition of American Psychological Association (APA, 2018), the systematic review comprises texts in which the authors define and clarify an academic restlessness, synthetize and summarize previous studies and inform to readers the state in which a certain investigation area is located. It also identifies relationships, contradictions and gaps in literature, as well suggesting solutions for these problems (Koller, Couto & Hohendorff, 2014). In the systematic review, from the perspective of Koller, Couto and Hohendorff (2014), the review is carried out in a similar way to conducting a survey, in which the observations are the studies. In the traditional review, evaluations of published materials are made, however, without making use of a systematic process to select the studies.

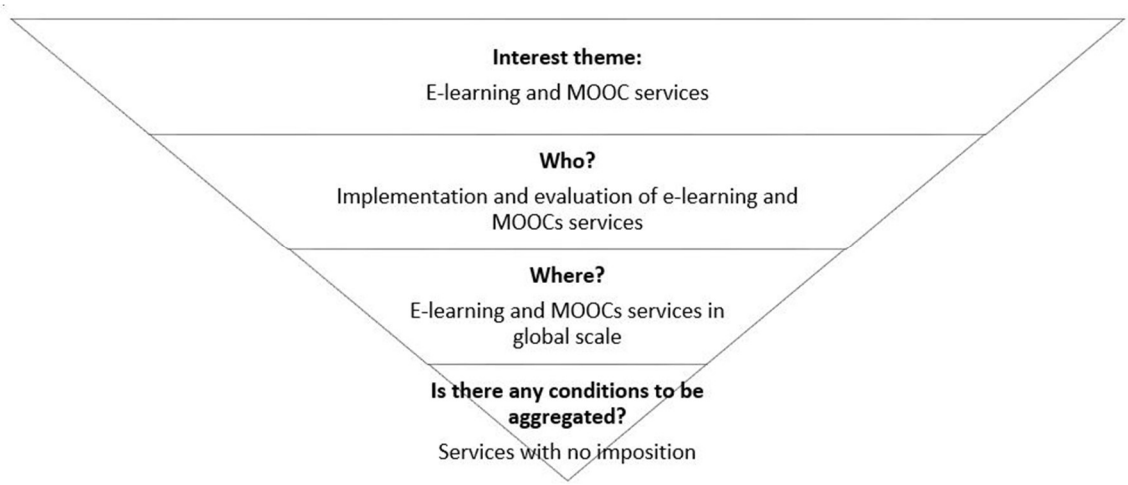
Tjahono (2010) justifies the use of systematic review by the need to identify trends, approaches, tools, technique, benefits and combinations of Six Sigma, the subject of authors' review. As a collection strategy, indexers were first identified under a temporal cut, to start the search for papers. In addition to the Tjahono (2010) process. The review of Leonidou and Leonidou (2011), from marketing and management publications in the period between 1999 and 2008, listed three eligibility criteria in the selection of scientific articles relevant to a study, comprising a search for journals, books and conferences. Finally, the papers must have to write in English, with the addend that this last criterion depends on each theme.

The reason for the use of review at Leonidou and Leonidou (2011) includes the characteristics of the authors who write on the subject, the evaluation of methodological procedures and the need to identify subtopics for the area. Similarly, Schaltegger, Gibassier and

Zvezdov (2013) also performed a systematic review with the same purposes, with the difference of using a data triangulation, by a snowball sampling, avoiding repetitions of papers.

Before to conduct the systematic review, it is important to delimit the topic so that the selected publications serve to the interests of researcher, in other words, so that publications be useful. According to Koller, Couto and Hohendorff (2014), it is possible to elaborate an inverted pyramid presenting the topic of interest; what is sought about the theme or who are the subjects of the research; where, defining the scope of the search and; if there is any aggregate condition to be used in the search. The delimitation of this review is showed as it follows.

Figure 3 Delimitation of systematic review



Source: Elaborated by author, 2018.

* Structure and topics of interest suggested by Koller, Couto and Hohendorff (2014).

Choose of review type

The process of choosing the type of review involved a comparison of the types of review and what each type can provide to researcher. Comparing the three types, it is possible to understand when and which situations to use each revision, what are advantages and disadvantages of each type and to evaluate in which situations it is necessary to use more than one type. Table 2 was adapted from Pocinho (2008), with contributions from Kitchenham (2007), adding contributions from Pendlebury (2008).

Table 2: Types of revision

Items	Narrative	Systematic	Bibliometric
Generic goals	Discuss the theme at a theoretical and contextual level, with a multidisciplinary approach	Evaluate the quality of the studies Summarize evidence Inference from evidence	General characterization of study Identification of subtopics and content aspects Evaluation of references
Question	Broad	Specific	Broad or Specific
Source	Often unspecified or with implicit criteria	Explicit and malleable search strategy	Explicit and embedded search strategy
Selection	Often unspecified or with implicit criteria	Criteria-based selection applied uniformly	Filter based selection
Evaluation	Free	Reproducible, with in-depth content evaluation	Reproducible, with generalist synthesis
Approach	Qualitative	Quantitative, qualitative or qualitative/quantitative	Quantitative
Inference	Based on the results context	Based on descriptive and inference analysis	Based on descriptive statistics
Allows meta-analysis	No	Yes	Yes

Source: Adapted from Pocinho (2008).
* With contributions from Kitchenham (2007) and Pendlebury (2008).

According to Table 2, the decision on which review type passes through the objectives of the review, the research question (review), the search strategy or in which sources seek the research, the selection criteria, the evaluation process, the approach to synthesis, possible inferences and whether meta-analysis is allowed.

- Generic goals: refers to the purpose of reviewing the literature. The review has an end in itself and may or not be added to a larger survey. In case of the review be a part of a bigger research, the review has the objective to obtain a product to be used in a larger research, and this information influences what type of review contribute in obtaining the desire product. In this research, the product of literature review is a theoretical model to be tested.

- Question: the question can be broad or specific. The magnitude that separates these two levels depends on each theme or topic studied.
- Source: the source is a criterion of choice of the review that refers to search strategy. The search can be embedded, full of rules that cannot be broken or may not have rules well defined, being flexible in search strategy.
- Selection: as well as the source, is a part of the process of obtaining the studies to be analyzed in review. This criterion concerns how selection is performed.
- Evaluation: the evaluation can be in depth or just generalist, being able to be more in depth in the narrative and systematic revision, by allowing to adopt subjective criteria and that can shorten a subject.
- Approach: Regarding the approach, the systematic review is the only one that can have a mixed approach, being able to be quantitative and qualitative.
- Inference: it concerns the possibility of inferring from the analysis and the way inferences are made.

Finally, the meta-analysis can be conceptualized as a set of statistical methods used to present quantitative results from several studies, synthesizing and summarizing the empirical knowledge contained in them about a topic (Littell, Corcoran & Pillai, 2008). By the definition of Kitchenham (2007), the meta-analysis or synthesis of data involves collection and summarization of the primary studies. A synthesis may be non-quantitative; however, a descriptive analysis normally can be complemented with a statistical test.

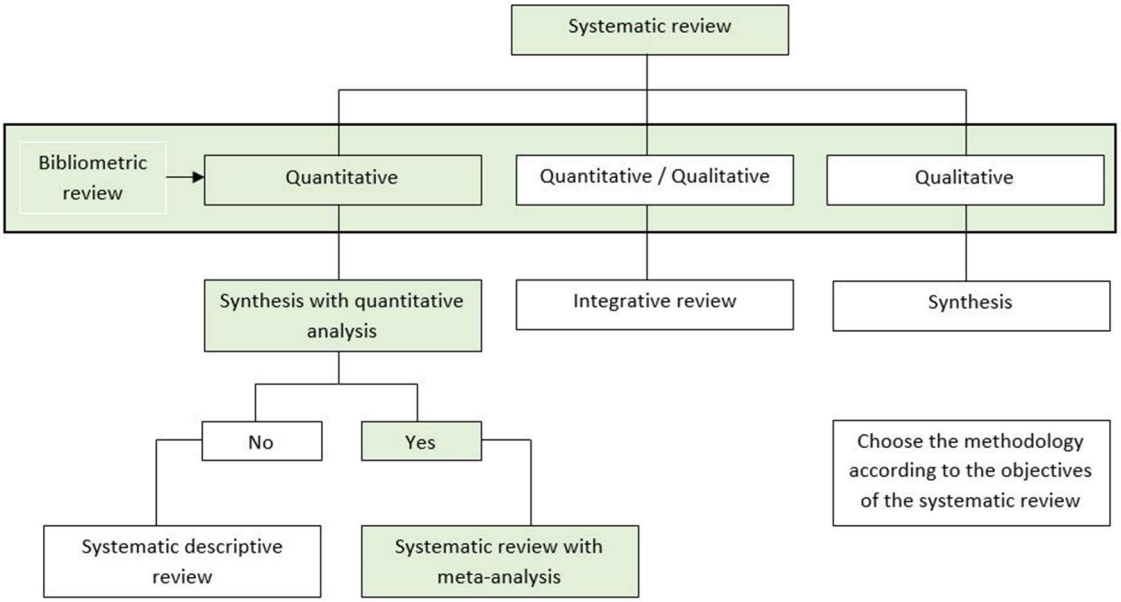
The statistical methods can be used to analyze central trends and variations in results across studies, “and to correct for error and bias in a body of research” (Littell, Corcoran, & Pillai, 2008, p. 2). The results of the studies can be converted into metrics, and from there using the measures in the statistical scope, crossing data, creating graphics, testing hypotheses with parametric and non-parametric tests, among other procedures. The synthesis of the data can be described in the planning of the review, identifying its purpose and possible variables to be summarized.

In this research, the meta-analysis acquires greater importance, by directly assisting the construction of a research model to be tested empirically and analyzed with a quantitative approach. In this case, the theoretical support that embassies the model goes beyond a narrative synthesis.

Thus, the option for this research was the use of systematic review with meta-analysis together with the bibliometric review. The study by Guanilo, Takahashi and Bertolozzi (2011)

presents a diagram that represents the choice of the type of review. With an adaptation, Figure 4 uses the author's diagram adding the bibliometric review in revision scheme.

Figure 4: Review type of research



Source: Adapted* from Guanilo, Takahashi and Bertolozzi (2011).

* The bibliometric analysis was inserted in the scheme proposed by the authors.

Synthesis

The choice of the type of review to be used in a study depends on certain criteria, involving the purpose of the review, as well as its main question. In addition, it is necessary to understand how the search and selection process define the results generated and how the results can and should be presented. The presentation is made through a synthesis, including qualitative and quantitative approaches, involving descriptive and inference analysis, so that the objective is reached. The meta-analysis may or not be used, depending on the need for each review. In the case of a review with the objective of obtaining a theoretical model to be empirically validated, the better option is a scheme with systematic and bibliometric review, using meta-analysis to obtain high accuracy.

3.2 Bibliometric and systematic review process

The bibliometric review is a method of revision that if well executed, diminishes the chance of failures in the collection process, besides presenting transparency in the search

method and selection of papers. However, for the formation of a research model, the structured script of bibliometric review is not sufficient to extract the necessary information. Specific aspects of the study of MOOCs are needed in support and assistance for the creation of the research model. For this reason, a systematic review is used together with bibliometric review.

The reason for choosing systematic review over traditional review is centered on the objective of the research, to search impartially for studies that theoretically support a research model. The narrative review has a tendency of selectivity in the literature, due to its partial and subjective character (Pai et al. 2004), being a possible bias acquired still in the conception phase of research model. Generate a bias of this magnitude in this step, disturbing the remaining steps and is irreparable after data collection.

Systematic review guarantees impartiality through paper selection with well-defined criteria, targeted reading and focus on the problem. Pai et al. (2004) cites some of the advantages of the systematic review, such as: search for studies following a question or problem, clear and reproducible qualification criteria for the selection of studies, critical evaluation through a script and synthesis of results according to a predetermined method.

Among the various reasons for using systematic review, Kitchenham (2007) lists the most common ones, such as summarizing existing evidence in the literature, providing a framework that allows new research activities on a topic or, to examine empirical evidence that contradicts or confirms hypotheses and helps to formulate new.

In the study of Hauge, Conradi and Ayala (2010) stages are defined about systematic review, resembling the steps of bibliometric review, however, with the view of having a specific problem; in other words, there is a need for information to be obtained. The steps of review, by Hauge, Conradi and Ayala (2010), following the Cochrane methodology, were adopted. Adapting to approach focused on behavioral studies and entering the steps of systematic review, bibliometric review, steps performed on literature review follow in Tables 3 and 4. The bibliometric review permeates several stages of the systematic review, mainly in phases three, four and five, involving the method of collection, selection of studies and documentation of the studies for later analysis.

Table 3: Systematic review protocol

Step	Operationalization
(1) Relevance	This is an emerging technology in education Lack of studies with quantitative approach to the subject Understood under a management perspective Low use of e-learning studies in the study of MOOCs
(2) Problem	From a management perspective, the MOOCs are being effective for users?
(3) Strategy collection	Bibliometric review • Search for indexers area of management and related areas with high impact journals • Selection of indexers • Search by index
(4) Selection of primary studies	Bibliometric review • Division by two themes: e-learning and MOOC • Keywords chained • Strings • Filters • Guided reading
(5) Critical evaluation of the studies	Reading full text Search script with specific information on the subject
(6) Analysis of results	Results of bibliometric review Results of specific information on the subject
(7) Conclusions	Summary of results (meta-synthesis)

Source: Elaborated by author, 2018.

According to Table 3, after setting the information that sustain the problem (1) and present the need for information (2), begins the collection (3), which also marks the beginning of bibliometric review, study selection criteria are the continuity of the steps of bibliometric review (4). Then it is made a critical assessment of studies by reading (5). The results (6), seek to meet the need for information, and finally, interpretations are made from the results (7), comprising the call meta-synthesis.

The bibliometric review also follows a specific form identifying the search terms, strings, filters, cutting time and outlines to the search strategy, until the studies that will be used in the research. Table 4 shows the form of bibliometric review, resulting in the number of studies selected for this research.

The primary source is the Capes directory, which brings together the major world indexes; in order to access the primary source, it is necessary to select the indexes, where the pursuit of studies is performed. The selection of indexers saves the researcher time, since it identifies indexers which include periodic areas of knowledge that permeate the research topic.

The search strings refer to the keywords used in the search, thought and chosen from prior knowledge of the researcher on the subject and also the search experience in the primary source. Two subjects were surveyed to specifically understand the MOOCs, but also e-learning, as a general theme. It also includes on strings, the time of publication. For e-learning, the period from 2007 was selected, and for MOOCs, from 2008 was considered, with the rise of MOOCs.

Table 4: Bibliometric research form

Themes: E-learning and MOOCs					
Primary source		Objective			
Capes platform		Index selection			
Research strings		Description			
Keywords for e-learning		E-learning and E-learning theories			
Keywords for MOOCs		Massive Open Online Courses and Higher education			
Time cut for e-learning		2007 to 2017		E-learning after consolidated	
Time cut for MOOCs		2008 to 2017		MOOCs rising	
Filters (primary sources)		Number after filter (e-learning)		Number after filter (MOOCs)	
Keyword 1		11836		2343	
Articles		10935		1562	
Revised by pairs		9509		1027	
Language: English		8621		958	
Keyword 1 and Keyword 2		1235		358	
Filters (secondary sources)					
Full text available		859		346	
Lecture: title, abstract and keywords					
Secondary source (by index)		Objective: identify articles		Objective: download articles	
		E-learning	MOOCs	E-learning	MOOCs
1	Cambridge	36	31	0	1
2	Emerald	64	11	10	3
3	Gale Cengage Learning	34	0	3	0
4	Oxford Journals	89	5	2	2
5	Sage Pub	48	48	0	2
6	Scielo	14	0	0	0
7	Science AAAS	0	4	0	2
8	Science Direct	256	68	24	21
9	Scopus	150	15	0	7
10	Springerlink	8	0	2	0
11	Wiley Online Library	160	160	8	0
12	Taylor&Francis	2	4	2	0
Total		859	346	51	38
		1205		89	

Source: Elaborated by author, 2018.

According to the Table 4, two separate searches are made in Capes platform; each one begins with the first keyword, where search filters are applied, the number of items to use each filter is reported. The first filters are: the study is an article; the article has been peer-reviewed and is written in English. The last filter of the primary sources is the inclusion of the second keyword, with the first interlocked. Using the second keyword it is allowed to select the articles focusing on the topic and eliminate those that just mention the subject.

Secondary sources involve two filters: the first one selects those in which the text is fully available in the Capes platform. The fact of being fully available in Capes indicates that when the article is selected in the index, it will also be fully available. Then the title, keywords and abstract are read; thereby, adherence to the theme can be better analyzed.

Finally, the search is directed at each index, searching for the items that were selected in each index to perform the download. Of the 859 articles selected in the search for e-learning, after reading titles, abstracts and keywords, 51 articles were selected. Of the 346 articles selected in the search for MOOCs, after reading titles, abstracts and keywords, 38 were selected for full reading. With 89 selected articles, the complete read of the articles was performed, being a guided reading, through a script of bibliometric and systematic research.

The bibliometric study generates results that together allow the viewing of the progress of studies on the subject. Under a quantitative approach, it is possible to identify a year publication peaks, research centers by nationality of the department and university researcher, indexers and journals with greater focus on the theme, historical data quotes, besides the formation of publishing networks co-authorships and citations.

As the limitations of this method for selecting high impact studies, it highlights two features: first, the algorithm used in each index is not known by the researcher, so indexers with more efficient algorithms may report search results more relevant. Another feature that limits the method is that the researcher trusts a priori on the ability of the authors of the studies to properly choose the keywords of their studies.

Despite the limitations, the advantages of bibliometrics overcome the limitations. First, by highlighting the impartiality in the search for studies: to find studies that disagree with each other or with contrasting results is more frequent in this type of search. Second, there is a standardization in the search method, which ensures process reliability and greater transparency in the choice of the most striking studies. Finally, regarding studies that are not selected, it happened by a plausible justification for not meeting the criteria used in bibliometric research.

About systematic review, unlike bibliometric research, the search script information in the study is not standardized, it depends on each theme. When the researcher starts reading the

most relevant articles determined by the order of appearance in the indexers, he can specify search criteria based on content. Thus, a specific search script for that study is formed with criteria that seek specific information on the topic that the articles have in common.

Synthesis: bibliometric and systematic review are methods to conduct a literature review, applicable in any area of knowledge. They are methods which standardize the search system by means of predefined steps. Two searches were considered, due to the nature of the topic and the need to generate a research model using theoretical contributions of e-learning and MOOCs. A selection of articles was performed, and from selected studies, a search script is designed to generate results, and the script of predefined bibliometric review, and the script of the systematic review, formed from the first readings. Eighty-nine papers were read in full and were used in the research to create a research model.

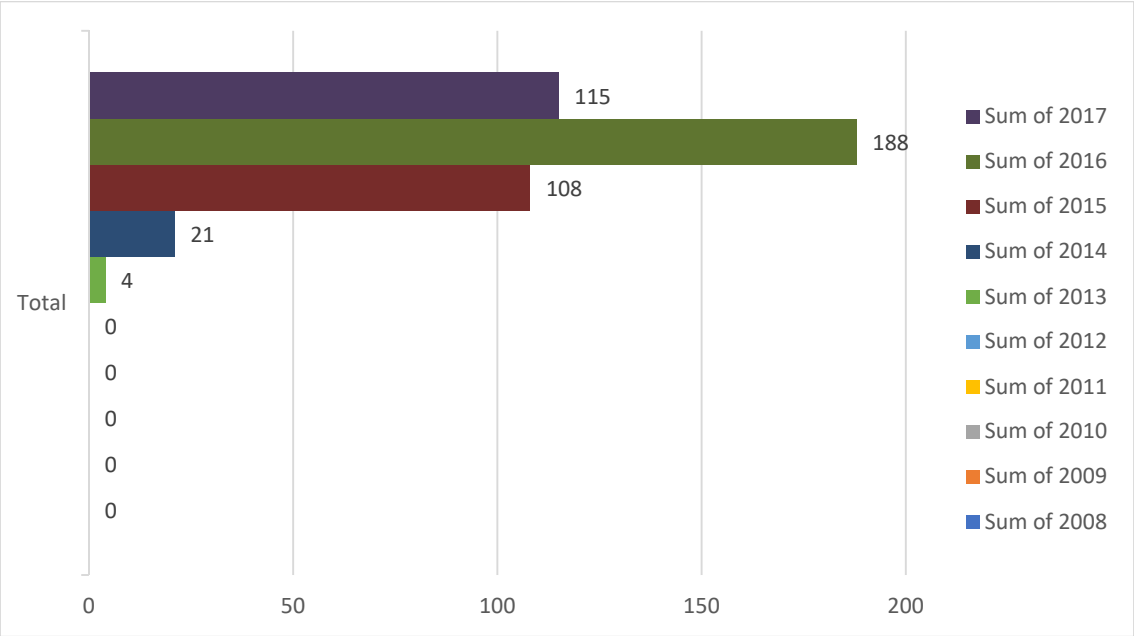
3.3 Results of systematic and bibliometric review

Together with the eighty-nine papers selected by the studies, 346 studies were selected until the last step, which consists an analysis of title, abstract and keywords. At the final of readings, only 38 were selected. Even if the most of studies were discarded, they can provide important information about the MOOCs literature. It is important to point out that the objective of this bibliometric review did not seek the most cited studies, but rather those that could provide more support in theoretical model.

Although they have not been used in other stages of the literature review, they can inform the current state of the publications about the subject. The same could be done about e-learning studies, however, the focus of this research is on MOOCs. In addition, MOOCs are a natural extension or branching of e-learning. Furthering the study of e-learning in other ramifications, such as corporate distance learning or e-learning for elementary and middle schools, would not gain real benefits for this research.

To assist the synthesis of the studies, we used the VosViewer, a visualization software for bibliometric networks. The research on MOOCs was extended to other indexers present in the Web of Science. In all, some more papers were added, being found 384 studies about MOOCs. The search used the same parameters previously performed and presented at bibliometric research form. Thus, it was possible to verify by the year of publication, the growth of published papers on the subject, as showed in Graphic 3.

Graphic 3: Publications about MOOCs since 2008 to 2017



Source: Elaborated by author, 2018.

It is noticeable the absence of papers published before 2013, one year after the birth of the first MOOC in the Coursera platform and the year that marks the increasing popularization of MOOCs. Starting in 2015, the number of publications increased sharply, peaking in 2016, with 188 papers published. In this way, it is possible to infer that the theme has been studied more frequently since 2015 and has aroused the interest of many researchers.

Besides noting the increase in the number of publications, the number of citations was also analyzed. The most cited papers are usually the seminal papers on the subject. The fifteen most cited papers are presented in the Table 5, containing the title, authors, year of publication, number of citations and average citations per year.

Table 5: Most cited papers about MOOCs

Title	Authors	Year	Citations	Mean by year
Literature and Practice: A Critical Review of MOOCs	Laverde, A.; Hine, N.; Martinez-Silva, J.	2015	20	6,67
Precise Effectiveness Strategy for analyzing the effectiveness of students with educational resources and activities in MOOCs	Merino, P.; Valiente, J..; Hoyos, C.; Sanagustin, M.; Kloos, C.	2015	18	6
A Study on the Pedagogical Components of Massive Online Courses	Rivas, M.; Figueira, E.; Campos, J.	2015	17	5,67
Are MOOCs Promising Learning Environments?	Bartolome, A.; Steffens, K.	2015	17	5,67
Designing MOOCs for the Support of Multiple Learning Styles	Gruenewald, F.; Meinel, C.; Totschnig, M.; Willems, C.	2013	27	5,4

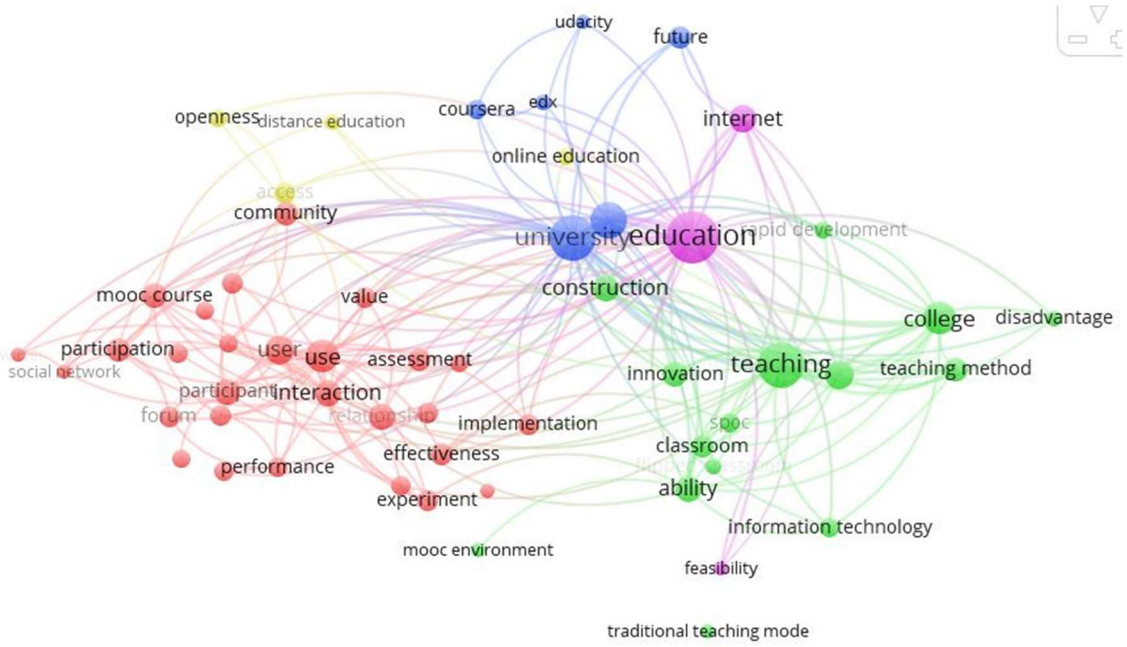
Analysing the Impact of Built-In and External Social Tools in a MOOC on Educational Technologies	Hoyos, C.; Sanagustin, M.; Kloos, C.; Parada, H.; Organero, M.; Heras, A.	2013	25	5
What Drives a Successful MOOC? An Empirical Examination of Criteria to Assure Design Quality of MOOCs	Yousef, A.; Chatti, M.; Schroeder, U.; Wosnitza, M.	2014	20	5
Teaching entrepreneurship using Massive Open Online Course (MOOC)	Al-Atabi, M.; DeBoer, J.	2014	19	4,75
Design, Motivation and Performance in a Cooperative MOOC Course	Castano, C.; Maiz, I.; Garay, U.	2015	14	4,67
Analysis and Implications of the Impact of MOOC Movement in the Scientific Community: JCR and Scopus (2010-13)	Meneses, E.; Cano, E.; Gravan, P.	2015	14	4,67
Peer and Self-Assessment in Massive Online Classes	Kulkarni, et al.	2013	23	4,6
Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization	Xing, W.; Chen, X.; Stein, J.; Marcinkowski, M.	2016	9	4,5
Democratizing education? Examining access and usage patterns in massive open online courses	Hansen, J.; Reich, J.	2015	13	4,33
Evolution of the Conversation and Knowledge Acquisition in Social Networks Related to a MOOC Course	Penalvo, F.; Benito, J.; Gene, O.; Blanco, A.	2015	10	3,33
MOOCs: The Challenges for Academic Librarians	Barnes, C.	2013	9	1,8

Source: Elaborated by author with data of Web of Science, 2018.

It is noticed that the most cited studies have the year of publication concentrated in 2015 Some papers seek to review what has already been studied about MOOCs, in other words, it has revisions about MOOCs. Other papers seek to understand the role of MOOCs in certain contexts and the influence of related themes, such as pedagogical factors or the influence of social networks on MOOCs. Going to a practical approach, other studies aim to understand or measure the impact of MOOCs on different realities. It can be inferred that there are already surveys that measure the impact of MOOCs and there are also searches for papers that question the real role of MOOCs in distance learning. Since the most cited papers have less than three years, it is possible that there will be further many aspects of MOOCs to be studied, extending the findings of the literature on the subject.

To check specifically the terms covered in the 384 papers and how frequently they are assessed, a network was created with the help of VosViewer, associating the most recurrent terms in the titles, abstracts and keywords. In addition to the associations, the software also groups the terms that usually appear together in the publications. The network has sixty-seven terms and five clusters, as can be seen in Figure 5.

Figure 5: Network about MOOCs: title, abstract and keywords

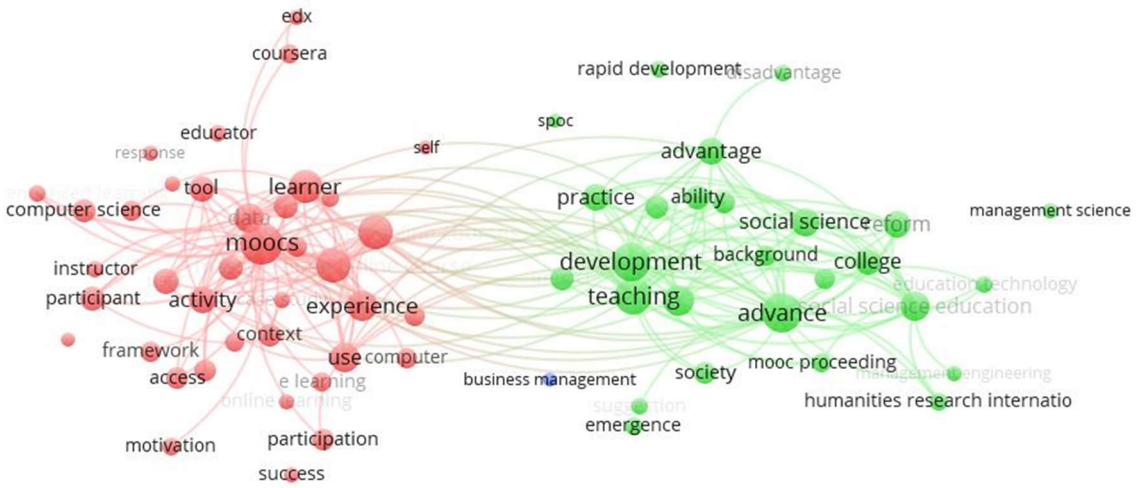


Source: Elaborated by author, 2018.

According to Figure 5, the most recurrent term is "education", followed by "university". Education and Internet are terms commonly found together in papers, forming a cluster. The university cluster brings together the names of platforms, such as EdX, Udacity and Coursera. There is a cluster with the most cited word being "teaching". This cluster has terms related to pedagogical factors and related to the use of information technologies in the teaching environment, bringing together words such as "classroom", "college", "information technology" and "innovation". The cluster with the highest number of terms has as its main representative the term "use" and adds related terms such as "experiment", "assessment" and "implementation". In addition, it also includes possible antecedents that explain the use of MOOCs, such as "interaction", "value", "effectiveness" and "performance".

Going deeper the analysis, the search for terms was extended to the full text of papers with the objective of verifying associations that would bring more insights and help to understand what is being studied about MOOCs. Analyzing the full texts of the 384 papers, 3 clusters were found. The network can be seen in Figure 6.

Figure 6: Network about MOOCs: full text



Source: Elaborated by author, 2018.

The presence of a small cluster is perceptible, with only the term "business management". It is possible to infer that the studies on MOOCs at a corporate level study the business model of the MOOCs and do not share similarities with the other studies on MOOCs. The other two clusters are represented in large numbers. The right-side cluster presents terms related to the theory of MOOCs and the impact of MOOCs on society, with a more theoretical approach. The cluster on the left side has terms commonly associated with empirical research, which measures or uses data to learn more about MOOCs.

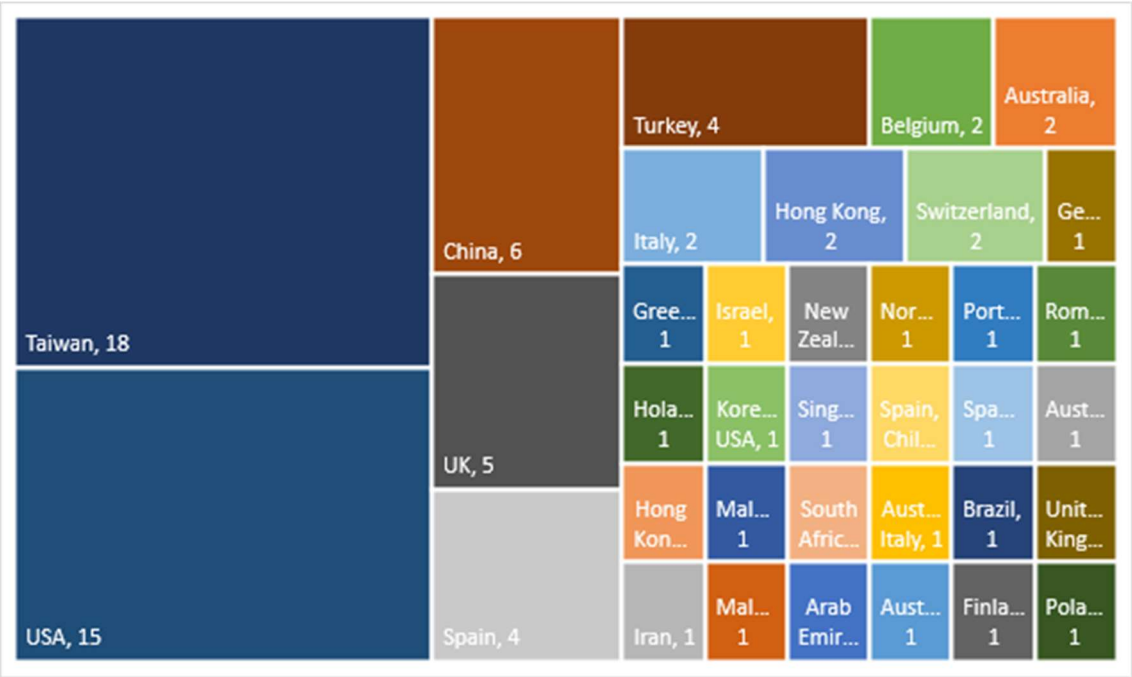
In general, the eighty-nine studies selected for this bibliometric research are concentrated in the clusters headed by the use of MOOCs and those clusters that represent the empirical studies. The remaining papers, however unused, provide an overview of publications on the subject, dividing efforts into empirical studies and researches that seek to explain the role of MOOCs in society.

Analysis of papers selected by the study

Focusing specifically on the eighty-nine papers selected in bibliometric phase, some information was collected on the publications to characterize the studies. The first one is the country of the authors, being data collected from the professional address of the first author of each paper. Identifying the country of origin allows mark poles of research on the topic. According to Graphic 4, the country with the highest number of first authors comes from

Taiwan with eighteen papers, followed by the USA, with fifteen papers. In Europe, there is also great representativeness dispersed in some countries, such as Spain, UK, Italy, Belgium and Turkey.

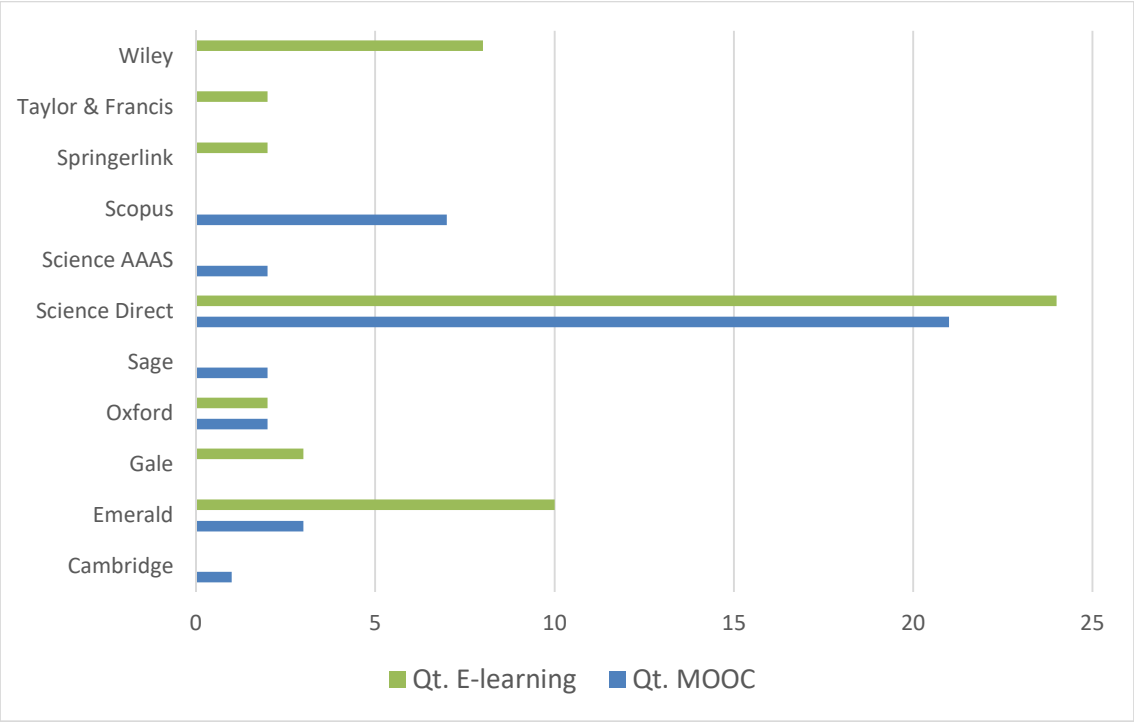
Graphic 4: Country of respondents



Source: Elaborated by author, 2018.

Another point analyzed was the representativeness of papers by indexer and by theme. Of the fifty-one papers on e-learning and thirty-eight on MOOCs, most of them focus on Science Direct, having twenty-four and twenty-one papers respectively, as shown in Graphic 5. Some indexers were represented only by papers on e-learning, being the Wiley Online Library, Taylor & Francis, Springerlink and Gale. In the same way, there are indexers that only present articles on MOOCs, being Scopus, Science AAAS, Sage and Cambridge. Thus, Science Direct is the indexer with more published material on both subjects.

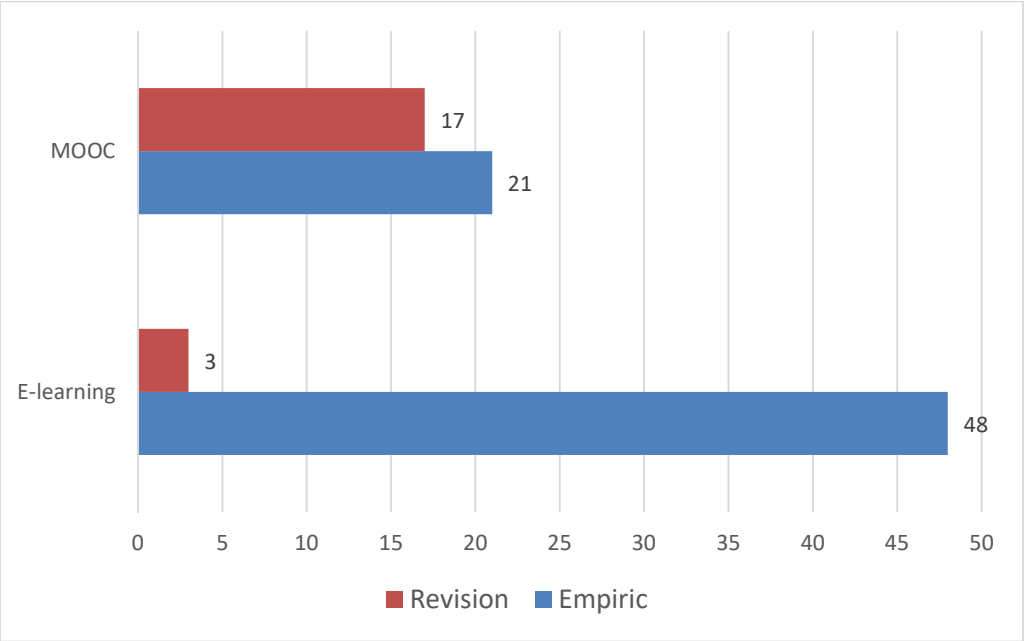
Graphic 5: Indexers by theme



Source: Elaborated by author, 2018.

Another point analyzed concerns the type of research, if it is a theoretical or empirical type. According to Graphic 6, the representativeness by theme presents differences: while papers on MOOCs are balanced between theoretical and empirical papers, in the research of e-learning, the majority is empirical. This result elucidates the different moments about the state of the art of each theme in 2018. Having most publications of the empirical type suggests that the study on e-learning is already consolidated, not necessarily exhausted. MOOCs, on other hand, are on the way to consolidation, with the number of empirical papers managing to overcome theoretical papers.

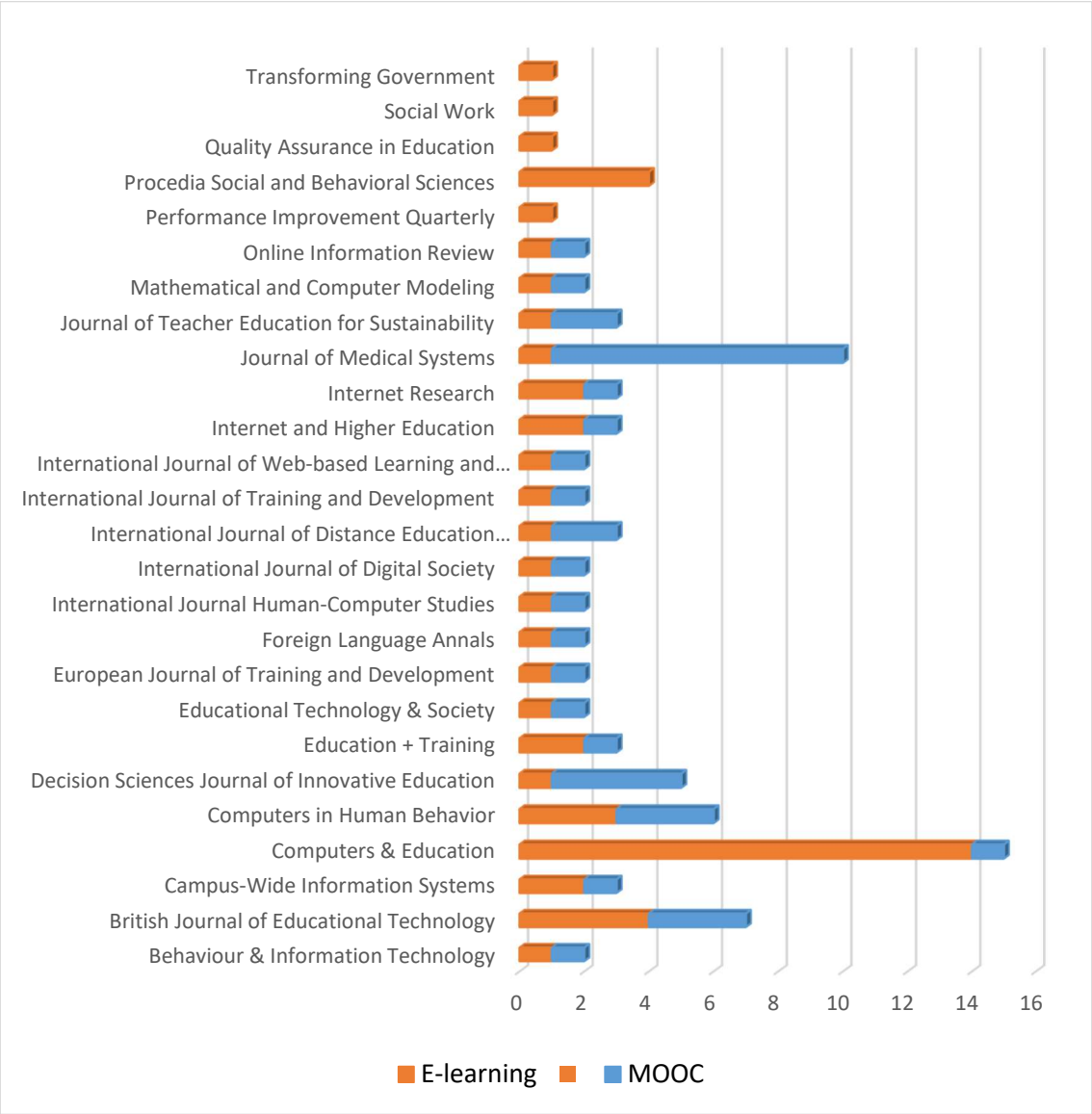
Graphic 6: type of study



Source: Elaborated by author, 2018.

In addition to checking the indexer, the periodical in which each paper was published was identified. Graphic 7 shows journal titles and representativeness by topic. It is interesting note the presence of periodicals with focus on other areas, showing the multidisciplinarity of the themes. Besides that, there are journals that only publish about MOOCs, but, in most studies there is a continuity of publication on e-learning and MOOCs in the same journal. The journals with the highest number of articles published on the subjects are Computers and Education, Computers in Human Behavior, Journal of Medical Systems and the British Journal of Educational Technology.

Graphic 7: Journals by theme

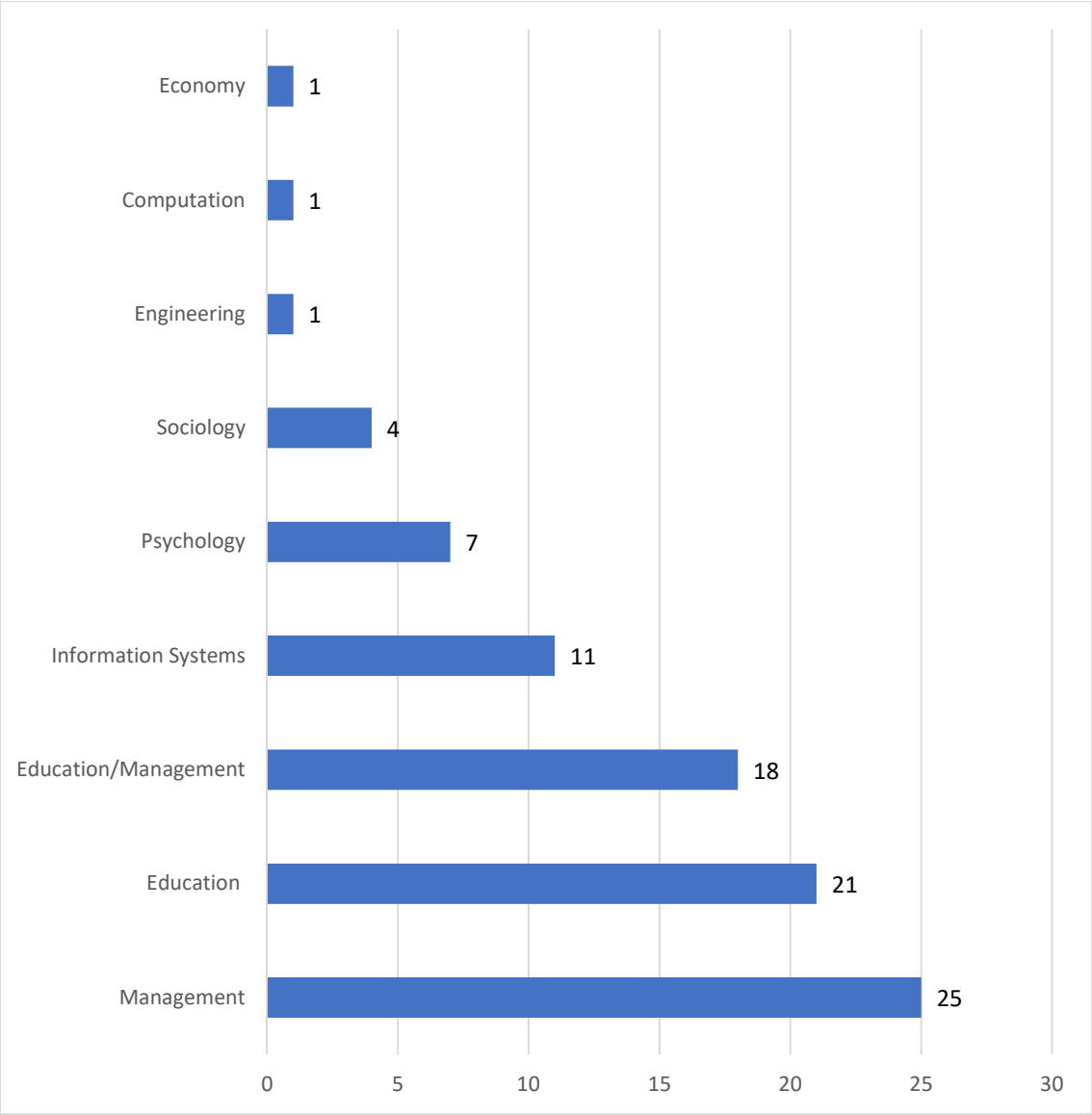


Source: Elaborated by author, 2018.

Advancing the content of the papers analyzed, the areas of each study were identified. Although the names of the journals indicate that there are distinct areas publishing about the themes, when analyzing the content of papers, the area of study is better identified. According to Graphic 8, most of the papers selected come from the management area, followed by the education area. Soon after, there is a group of papers that cannot dissociate the areas of management and education. These results were expected, since studies focuses on the implementation, use and success of e-learning and MOOC services, and these interests are usually the focus of researchers in these areas. However, the use also involves researches in

information systems, being represented by eleven papers. Other areas such as psychology, sociology, engineering, computing and economics have also been identified.

Graphic 8: Content frequency



Source: Elaborated by author, 2018.

Another element evaluated was the theoretical scope of guiding theory of each paper. They were identified at revision topic of each paper. In case of use of various theories, the most important or the most relevant to the study was associated to the paper. The most used models are presented in Table 6. Also, in Table 6, the object of study was identified. The theory and study object are integrated elements, because the theoretical scope supports the analysis about the study object.

Among the most widely used models, TAM (Technology Acceptance Model) and EDT (Expectancy Disconfirmation Theory) are the most used, followed by TRI (Technology Readiness Index), TPB (Theory of Planned Behavior), UTAUT (Unified Theory of Acceptance and Use of Technology) and Information Systems Success Model. Regarding the study object, most studies focus on applied researches in higher education. This term was also used such a keyword of bibliometric analysis. It is possible to identify some study objects with theoretical perspective, such as acceptance of e-learning, e-learning continuance, emergent technologies and MOOCs business model.

Table 6: Theoretical model and study objects most used

Theories	Qt.	Study object	Qt.
TAM	13	Higher Education	20
EDT	12	Employees	13
TRI	7	MOOC users	12
TPB	4	Acceptance of e-learning	8
UTAUT	3	E-learning continuance	7
IS Success Model	3	Emergent Technologies	5
		MOOCs Business Model	4

Source: Elaborated by author, 2018.

Elements of studies with quantitative approaches

Regarding methodological aspects, the sample size for quantitative studies was measured. Of the eighty-nine papers, sixty-one present a strictly quantitative approach. Of these, twenty-seven studies reported sample size, with an average of 226 observations. The largest sample was encountered in Sawang, Newton and Jamieson (2012).

Still in the quantitative studies, the dependent variables and independent variables of the studies were identified, in order to verify the most recurrent constructs in the theoretical models. Table 7 lists the dependent variables with the highest occurrence rate. According to Table 7, "satisfaction" and "intention to continue" using e-learning services and MOOCs are the most recurrent, with twenty-six and nineteen occurrences respectively. There is also the small-scale use of "intention to use" and "adoption", propitious for collection at moments before and during use. In addition to these, some constructs act together with satisfaction, such as "performance", "readiness", "engagement", "success" and "effectiveness".

Table 7: Dependent variables most recurrent

Dependent variables	Qt.
Satisfaction	26
Continuance intention	19
Intention of use	7
Adoption	3
Performance	3
Readiness	2
Engagement	2
Effectiveness	2
Success	2

Source: Elaborated by author, 2018.

To list the independent variables, the online resource called “Imgur” was used. This tool creates a word cloud from a text file, containing terms separated by comma. Since the diversity of independent variables is high in the selected studies, the cloud prioritizes the most recurrent ones, increasing the text font, while gathering many words in the same image. Figure 7 shows the result for the sixty-one studies with quantitative approach.

Figure 7: Independent variables most recurrent

Source: Elaborated by author, 2018.

According to word cloud, "quality" is the most recurrent independent variable, followed by attributes. These attributes are related to most commonly theories used in replication studies. Other variables with great representativeness are the "interaction", the ease of use, with the

combination of "ease" and "use", "value", "attitude", self-efficacy, with the combination of "self" and "efficacy" and finally "usability".

Evaluation of quality of papers with quantitative approach

The absence of important information was perceptible during the readings of full text of papers. The presence of complete information is a sign of quality. Likewise, if a study has a high quality, it is expected to be useful for the research's purpose. Thus, the quality of each paper was analyzed by a scale between zero and ten, observing objective and subjective elements that should be presented in papers.

Thus, for sixty-one studies with a quantitative approach, some important aspects were checked whether they were presented in the papers, being: (1) the indication of the software used; (2) the presentation of the sampling plan or just the sample size; (3) information about the choice of scale type used in the research instrument and (4) information on the model tested, whether it is an adaptation, replication or a new model. These variables are categorical, using "yes" or "no" codified by zeros and ones. The quality note presents subjective elements, such as the capacity for synthesis, clear objectives and well-presented results.

Based on these information, a statistical test was performed to answer the following question: are there differences between the presence of relevant information in a paper between papers classified as lower quality and higher quality? It is important to emphasize that lower quality category is in fact compared to those of higher quality, but still maintain a high standard of quality, according to the journals in which they were published.

The data were coded and imported into the R integrated development environment, used for statistical calculations, graphical analysis and programming in language R. The R provides a wide range of statistical techniques, including classical tests, modeling, classification, grouping, mining, among others. Being widely used by statisticians, it has been gaining popularity in the applied areas of statistics (Muenchen, 2012).

The variable for grouping, that is quality, was divided into two categories, being 0 - lower quality and 1 - higher quality. The studies that obtained note equal or below to five, were classified at lower quality category. Studies that obtain note above five, were classified at higher quality category. According to Walpole, Myers, Myers and Ye (2009), when performing a hypothesis test, it is necessary to evaluate the type of variable and the number of groups that we are trying to compare. In this case, four paired tests were performed. For each test, one of the variables, such as software, was compared with the quality variable, being both categorical.

For this configuration, Walpole et al. (2009) indicate the use of two-sample test for two proportions. Test to verify that the two proportions are equal. "In general, we want to test the null hypothesis that two proportions, or binomial parameters are equal. That is, we are testing p1 = p2, against one of the two alternatives p1 < p2, p1 > p2 or p1 ≠ p2" (Walpole, Myers, Myers, & Ye, 2009, p. 231). In this case, the test is uni-flow, that is, the test has the alternative p1 > p2. The critical regions follow the critical points of the standard normal curve, so the critical region corresponds to 1.64.

The test adopted maintain similarities with the contingency test. The samples of each group are compared by the ratio of the checklist equal to one (element present in the paper) of each variable in each group by the number of papers analyzed in each group. Following this formula.

$$p_1 = \frac{x_1}{n_1} - p_2 = \frac{x_2}{n_2}$$

Being:

x_1= number of information present on item x of papers in group 1: low quality;

n_1= number of individuals in group 1: 31 studies

x_2= number of information present on item x of papers in group 2: high quality;

n_2= number of individuals in group 2: 27 studies

By generating $p_1$and p_2, it is also generated a p defined by:

$$p = \frac{x_1 + x_2}{n_1 + n_2} \qquad (1)$$

The value z to test p is defined by:

$$z = \frac{p_1 - p_2}{\sqrt{pq(\frac{1}{n_1} + \frac{1}{n_2})}} \qquad (2)$$

With: q = 1 - p.

The test was conducted in R, and the results are presented in Table 8, containing the presence of software, sample, scale and model information from smaller and higher quality papers with quantitative approach. Besides that, the results of the test with p-value and status of hypothesis are presented. Finally, the average of the information checks was calculated, and a general test was performed.

Table 8: Hypothesis testing

Variable	Low quality	High quality	Z	p-value	Status
Software	20	20	-1.1010	0.1355	H0 not rejected
Sample	7	21	-4.3693	0.0000	H0 rejected
Scale	2	9	-2.7161	0.0033	H0 rejected
Model	22	23	-1.6480	0.0497	H0 rejected
Mean	12.75	18.25	-2.2329	0.0128	H0 rejected

Source: Elaborated by author, 2018.

According to Table 8, only the software did not have the null hypothesis rejected, indicating that this information does not help differentiate between low and high-quality papers. For the other variables and general test, the value of statistical tests was above 1.64 in module, with p-value below 0.05. Negative values indicate that studies classified as high quality are those that present more relevant information than those were checked.

It is possible to infer that a paper that provides information about the sample, choice of scale and information about the type of model to be tested, are those papers classified as high quality, and consequently may be more useful as used by other researchers. These results help in the next step, regarding the creation of the initial research model, since it provides a selection of more useful papers.

3.4 Research model and hypotheses

The research propose is to develop a reproducible scale specifically for MOOCs services through a model of effectiveness in MOOC courses. To achieve this objective, a discussion on what is being studied on MOOCs currently and the evolution of these studies over a recent period is promoted. A bibliometric and systematic review with meta-analysis was conducted to assist in the identification of items, constructs and methods of analysis more useful and viable. Finally, the topic 3 presents the model of research, including the initial search instrument.

To present a research model, the literature review should look into the studies, building a knowledge about the evolution of the theme over time, as the general and specific aspects. A model can be defined as a language that allows the model to observe the form of hypotheses and then test them empirically. The creation of hypotheses is a result of a literature review. Regarding MOOCs, it is looking for a model that focuses on the perception of the users, with a use of behavioral variables.

In this study, the constructs quality, usability and value are operationalized under the theoretical scope of the EDT, by Oliver (1980), which in turn gave rise to the ECT. In this theory, rooted in marketing studies, there is a cognitive model that supports the relationship between constructs. Starting from consumer's perception in before and after purchase moments, seeks to explain consumer's behavior, relating satisfaction and expectations (Pereira et al. 2015; Chiu et al. 2011). Thus, customer's satisfaction is co-determined by the disconfirmation of expectation (Chiu et al. 2011).

Satisfaction, as a dependent variable, is that one that determines whether there will be repurchase of the product or service. In e-learning studies and in the last five years in MOOC studies, the EDT is widely used, as it uses the satisfaction (related to the service) and continuance to use (such as repurchase) as dependent variables, as well verifying the expectation with the e-learning service through different independent variables.

In studies that use the EDT with a quantitative approach, identified in the bibliometric study, many of them consider the constructs commonly used when applying EDT, being more frequently quality and usability, that are constructs idealized in other theories, such as TAM and TPB. Other studies use the relation between expectation and satisfaction to apply in various e-learning services. One advantage of using disconfirmatory constructs is that it is possible to analyze two distinct moments in time, without having to do two data collections. In addition to these studies, five of them use EDT directly, being the guiding theory for replications and adaptations of various studies, besides serving as a basis for the construction of new models, as is the case of studies of Chiu et al. (2005) and Pereira et al (2015).

Among the constructs operationalized based on EDT, those with the highest frequency are quality, usability and value, being used by Chiu et al. (2005), Chou et al. (2010) and Pereira et al. (2015). The study by Cheung and Lee (2011) explores the quality, based on EDT, and includes technical attributes of the learning environment. The study of Lin (2011) uses EDT to adapt a model based on usability, usability disconfirmation and attitude. Hung and Cho (2008) also use other constructs, such as self-efficacy and compatibility.

Despite the use of common constructs to other theories, the most frequent one's present better results in the quantitative models tested by these studies. Another point in common between these studies, is the understanding that the set of constructs lead to perceived performance. Finally, the disconfirmatory constructs, capable of verifying expectations with the e-learning service, are used in some studies, each one with different arguments for use, depending on the object of study or the unviability of a data collection with longitudinal cut.

Effectiveness is based on the fulfillment of goals and the feelings expected at the end of a well-conducted process. The term training effectiveness, used by Chien (2012), goes even further by relating effectiveness as an end point resulting from the use of expected skills, behaviors and attitudes. In this study, the point of arrival is the completion of a MOOC course, accompanied by feelings such as satisfaction and continuance intention. Thus, the effectiveness of the MOOCs comprises the feelings measured by the constructs operationalized in this research.

Defining quality

The quality is a diffuse and multifaceted construct, which presents different approaches over time and has use in several areas. The quality analyzed in e-learning services and MOOCs originates from the marketing literature and to two decades in the behavioral studies of IS/TI users. The construct in the research theme is usually attributed to the SERVQUAL model and the SI Success Model. Roca, Chiu and Martínez (2006) had already contemplated it when proposing a decomposed TAM model, a model specific for e-learning services. In 2005, Chiu et al. (2005) had also used quality including disconfirmatory mode.

Recent researches, such as Ismail et al. (2012) divide the quality into three axes, using IS Success Model by DeLone and McLean (2003), being information quality, service quality and system quality. From this context, the concept of quality for the theme can be defined as a key element to differentiate services and obtain competitive differential and refers to the efficiency and effectiveness with which products and services are delivered (Pinho & Macedo, 2008).

Going specifically to the distance learning, the quality is conferred by indicators such as completeness of information, response time, layout, navigation, among others. It is important to obtain quality so that users (students) can be exceed expectations and consequently achieve better performance and be more satisfied (Cheung & Lee, 2011). The main theoretical models maintain that quality is the main construct that leads to satisfaction, as proposed by Parasuraman, Zeithaml and Berry (1988) with the service quality model (SERVQUAL) and later with the IS Success Model of Delone and McLean (2003).

In the context of MOOCs, quality makes perfect sense, as MOOCs have been developed to deliver high quality teaching. Measuring the quality of a MOOC as a teaching platform, allows for improvements, control and implementation of MOOCs with a higher success rate (Ventura, Bárcena & Monje, 2014). Thus, the quality in this research is a construct used in

disconfirmatory mode, based on the EDT, also being a component of the perceived performance. The indicators used in the quality construct are show as it follows.

Table 9: Quality variables

Construct	Indicator	Variable description
Quality (QUA)	QUA1 – Layout	The layout and interface of the course area are friendlier than expected
	QUA2 – Navigation	About the ease of navigating the course area, it was better than expected.
	QUA3 - Comfort	About the comfort of using a virtual learning service, it was better than expected
	QUA4 – Completeness	About the information offered in the course area, I felt that they were complete
	QUA5 – Content Quality	The quality of content offered is better than I expected
	QUA6 – Support Readiness	On system support for any issues, I felt it was better than I expected

Source: Adapted from Chiu et al. (2005), Cheung and Lee (2011), and Pereira et al. (2015).
*Items from these studies were adapted for the context of MOOCs.

Defining Usability

The usability is commonly associated with the TAM model, and can be evaluated in both moments, before and after use, thus allowing the comparison between expectation and performance. Usability refers to degree to which the users of the technology and the service offered considers the meeting of their needs (Lin, 2011). The most recurrent item in this construct is ease of use, often being placed as a construct itself. But when it comes to e-learning and MOOCs services, various indicators can integrate usability, such as the system support user productivity, the transition of content showing the organization of the information shown on screen, the features that the system offers, the ease with which the individual acquires skill in handling the platform, among others items.

All the studies evaluated in the bibliometric research that use the TAM model as theoretical scope make use of this construct. With a designation of usability, there are six studies, being those by Liaw (2008), Chou et al. (2010), Teo (2010), Cheung and Lee, (2011), Lambropoulos, Faulkner and Culwin (2012) and Pereira et al. (2015). As a common point, the indicators aimed at ease of use or usability are associated with user familiarity with technology and user-driven attitudes to become more skillful.

Among the six studies cited above, that of Lambropoulos, Faulkner and Culwin (2012) is the one that most explores this construct, proposing pedagogical usability indicators in an e-learning service, including system instructions, frequency of use, alignment with the course objectives, functionalities, attractiveness, among others. In this study, it is clear the approximation and importance of usability in services that offer teaching.

In the context of MOOCs, usability must comprehend technical aspects of the on-line platform, without distancing itself from pedagogical aspects. In fact, they do not need to be separated. Measuring usability through users' perceptions allows a guided offer of platform functionality, support for motivation so the user can learn more and can also can help with improvements that make it easier for the user to acquire system management skills in a short time (Lambropoulos, Faulkner & Culwin, 2012).

Some studies such as Cheung and Lee (2011), Chiu et al. (2005) and Pereira et al. (2015), among others, applied the usability in the theoretical scope of EDT. Thus, usability is measured in a disconfirmatory mode, leading along with other constructs to a perceived performance or directly to user satisfaction, as it was in the study by Cheung and Lee (2011). The purpose of such studies was to evaluate satisfaction by attending or not expectations related to access, navigation resources, speed to acquire skills and ease of use.

In fact, usability, in the teaching platform, follows the same principles of usability in information systems and games, with the pedagogical factor permeating the best selection of indicators. Thus, usability in this research is a construct used in the disconfirmatory mode, based on EDT, and a component of perceived performance. The indicators used in the usability construct are shown as it follows.

Table 10: Usability variables

Construct	Indicator	Variable Description
Usability (USA)	USA1 - Easiness	On the use of the teaching platform, I felt it was easier than I expected
	USA2 - Skill	On the use of the teaching platform, I felt that I became skilled faster than I expected
	USA3 – Features and activities	About the features and activities offered in the course, they are better than expected
	USA4 - Tools	About the navigation tools in the course area, I felt they are better than I expected
	USA5 – Content transition	Accessing the content in the correct order was easier for me than I expected
	USA6 – Support productivity	The teaching platform helps increase my productivity, even more than I expected.

Source: Adapted from Liaw (2008), Chou et al. (2010), Teo (2010), Cheung and Lee, (2011), Lambropoulos, Faulkner and Culwin (2012) and Pereira et al. (2015).

*Items from these studies were adapted for the context of MOOCs.

Defining Value

Value is an abstract concept; it refers to a user's view of the technology or technology based on service that is delivered. Value is a construct that can have a variety of designations, such as benefits, utility and feelings of accomplishment of the individual for being using a service (Pereira et al. 2015). In TAM and EDT models with this concept are widely used. In TAM model it is called usefulness.

The construct is supported by the literature because it greatly influences the levels of satisfaction and consequent continuance to use (Chou et al. 2010). In addition to the intense use and adaptation of the construct through the studies based on TAM and EDT, the value still presents flexibility in its composition of indicators, adapting to other theoretical bases and yet reaching high levels of impact on the dependent variables (Lin, 2012; Cheng, 2012).

One of the studies that extends the use of the value construct is that by Boe, Gulbrandsen and Sorebo (2015), who use value as the main construct to measure incentives related to information and communication technologies, using as theoretical bases the Agency Theory (AT) and Information Systems Continuance Theory (ICST).

Boe, Gulbrandsen and Sorebo (2015) argue that value perception requires previous experience with service or use of technology, being a use-based emotional response. This argument supports the use of the construct in the disconfirmatory mode, as it is based on an evaluation after use. However, the use in the disconfirmatory mode, despite collected one time, refers to two moments, before and after the experience. Other studies have shown that the value also shows a great impact when evaluated before use, reaching similar levels when measured after use (Chiu et al. 2005; Pereira et al. 2015).

In the context of the MOOCs, the value must follow the same essence already contemplated and consolidated in e-learning studies, focusing on possible benefits and feelings of accomplishment of the individual to be attending the course. The accomplishment feelings may or may not be related to the learning factor, since all feelings raise satisfaction rates and are directly related to perceived performance. Thus, the value in this research is a construct used in the disconfirmatory mode, based on EDT and a component of perceived performance. The indicators used in the construct value are shown as it follows.

Table 11: Value variables

Construct	Indicator	Variable Description
Value (VAL)	VAL1 - Achievement	I feel more fulfilled than I expected in this course
	VAL2 – Needs	I feel that I met more needs than I expected during this course
	VAL3 – Follow Trend	I feel that by attending this course, I most follow a trend than I expected
	VAL4 – Intelligence sense	I feel smarter than I expected to attend this course
	VAL5 – Independence	I have an autonomy in choosing contents and planning my studies in the teaching environment, greater than I expected
	VAL6 - Entertainment	I can enjoy myself along the course, more than I expected

Source: Chou et al. (2010), Lin (2012), Cheng (2012), Boe, Gulbrandsen and Sorebo (2015) and Pereira et al. (2015).
*Items from these studies were adapted for the context of MOOCs.

Defining Interactivity

Interactivity is a construct operationalized in academic research in the last ten years. Chang, Hung and Lin (2015) have stated that interactivity or interaction tends to be the major trend in the development of MOOCs. The construct refers to how people relate and influence each other. In the context of MOOCs, interaction represents the maximum potential of MOOCs in recruiting students and developing specific skills.

Under the theoretical aegis, the need for social awareness is a central point of the social constructivism. Learning through conversations can build knowledge, so learning derives from social interactions (Lambropoulos, Faulkner & Culwin, 2012). Based on this theoretical basis, in the context of MOOCs, the tools of interaction can be used in their full capacity, that is, the maximum collective construction of knowledge can be extracted through interactivity.

The interactivity can be still associated with social networks, being resources used in the recent history of education as pedagogical tools (Ventura, Bárcena & Monje, 2014). Social Network Analysis (SNA) and Visualization Interaction Tools (VIT) studies allow detecting communication and relationships between people and groups (Lambropoulos, Faulkner & Culwin, 2012). In the context of MOOCs, social networks are also widely used, along with other interaction resources, such as emoticons, avatars, on-line status, chats, forums, lives, discussion rooms, participation graphs and collective documentation of knowledge.

In the study by Chien (2012), the interactivity of the system with the user/student was measured as the antecedent of the training effectiveness construct. Thus, interactivity was measured as an antecedent of learning effectiveness. Similarly, Liaw (2008) also used interactivity as an antecedent of what was called developing effective e-learning. Despite the timid participation of the construct in e-learning studies, for MOOCs, the concept and use of interactivity is the same, changing only the magnitude of the influence of interactivity in the learning process.

Thus, interactivity is treated in this study as a construct, which makes part of effectiveness in MOOCs, focusing on the tools and resources available in the MOOCs, including the use of social networks and interaction resources available in different educational platforms, without ignoring the learning factor in their scope. The indicators used in the construct interactivity are shown as it follows.

Table 12: Interactivity variables

Construct	Indicator	Variable Description
Interactivity (INT)	INT1 – Hypermedia	Using multiple medias, throughout the course, facilitates communication and learning
	INT2 – Reading forums and chats	I read forum and chat messages frequently
	INT3 – Writing forums and chats	I write messages in forums and chats frequently
	INT4 – Stimulate participation	It is stimulating to participate in the forums and talk with other course students
	INT5 – Doubts and deepening	I always use the interaction resources to ask questions and learn more
	INT6 – Interaction outside the platform	I seek to participate in groups created in social networks by teachers or other students

Source: Adapted* from Liaw (2008), Chang, Hung and Lin (2011), Lambropoulos, Faulkner and Culwin (2012) and Ventura, Bárcena e Monje (2014).
* The items for this study were inspired by these studies.

Defining Collaborative Learning

Increasing students' participation in activities is a common practice, regardless of the teaching mode; participation can be increased, but the same cannot be said of learning (Lambropoulos, Faulkner & Culwin, 2012). The collaborative learning (CL or CeL) is a form of knowledge construction, based on collective learning and having knowledge as product. Collaborative learning is based by the same theory of interactivity.

In MOOC courses, collaborative learning happens in a particular way compared to other distance learning modalities. Firstly, students have heterogeneous backgrounds and scholarship, secondly, the level of content presented indicates different times of learning for each individual. Therefore, according to Gillani and Eynon (2014), it is important to understand how MOOCs participants will learn together, through such disparity of characteristics. It is also important to balance the use of collaborative learning with individual readings and tests.

Zhuhadar, Kruk and Daday (2015) proposed a new, more connected and multi-platform MOOC, focused on the learning process through collaborative learning tools. The proposal consists in a semantic MOOCs management system, being a prototype of a medium-sized platform implemented at the University of Kentucky. The platform comes from the concept of collaborative filtering. Following a similar precept, Biasutti (2011) sought to measure cooperative on-line learning, that is, with mutual collaboration among students.

In the context of MOOCs, collaborative learning already presents native and recurrent tools in MOOCs, such as the correction of activities among students and the freedom to share knowledge, to propose new ideas and new ways of studying (Paechter, Maier & Macher, 2010). Like the interactivity, the learning factor permeates the indicators of collaborative learning, focusing on the use of resources and tools of the MOOCs, without disengaging from teaching. Collaborative learning in this research is a component of the effectiveness of MOOCs. The indicators used in the construct are shown as it follows.

Table 13: Collaborative Learning variables

Construct	Indicator	Variable Description
Collaborative Learning (COL)	COL1 – Correction of activities (Learning)	I believe that the correction of activities also being made by students adds value to the learning process
	COL2 – Correction of activities (empathy)	I feel comfortable correcting other students' work
	COL3 – Freedom to propose new forms to study	I feel free to suggest new ways of studying
	COL4 – Freedom to indicate materials	I feel free to suggest supplementary materials that can deepen the content of the course
	COL5 – Teamwork	I can learn more by helping and being helped by other course students.
	COL6 – Interest to share	I like to share my opinion or what I have learned with other students
	COL7 – Knowledge shared	I try to find out what the other students understood about the content of the course and I share with them what I understood

Source: Paechter, Maier and Macher (2010), Biasutti (2011), Lambropoulos, Faulkner and Culwin (2012), Gillani and Eynon (2014) and Zhuhadar, Kruk and Daday (2015).

* The items for this study were inspired by these studies.

Defining Engagement

As discussed in previous topics, the MOOCs present distinct characteristics of the traditional teaching mode and previous distance learning mode. In MOOC courses, With the low rate of completion of MOOC courses, student effort prevails, that is to say, the individual characteristics weigh on performance and user satisfaction. In this context, there is the use of engagement, which refers to individual characteristics that enhance and increase the student's chances of achieving better performance.

Greene, Oswald and Pomerantz (2015) used the Implicit Theory of Intelligence to address student's engagement in the academic activities of a MOOC. This theory holds that skills are malleable and can be developed. In the study, some hypotheses were pointed out, such as relevance of MOOC to users, previous experiences with MOOCs, commitment and motivation. Similarly, Perna et al. (2014) addressed other indicators that relate to user milestones during the course, checking the progress in the course by the users, and thus, verifying the persistence and dedication in the courses.

In the study by Sawang, Newton and Jamieson (2012), the authors emphasize that individual characteristics of student are key factors for performance and user satisfaction. From this premise, the study addresses the openness to change, based on the already consolidated in e-learning studies, self-efficacy, and may also be related to IDT. Openness to change concerns the openness that student gives to face new challenges. Until recently, in 2012, MOOCs were new and most of the students were attending the first MOOC for the first time. In 2018, openness to change can be updated, being integrated into the engagement construct.

Considering the above, engagement takes on greater importance in MOOCs than in other types of distance learning, since MOOCs require the free will of the user entering the course and also requires individual characteristics such as motivation, persistence and dedication to achieve high levels of performance and consequent satisfaction with the course. In this way, engagement is a mediate construct of effectiveness and performance to satisfaction. The indicators used in the study follow a different structure from the others, due to different needs in the use of scales more adequate for each indicator.

Table 14: Engagement variables

Construct	Indicator	Variable Description
Engagement (ENG)	ENG1 – Adaptability in distance learning	Taking a distance course is easy for me (Yes or No)
	ENG2 - Persistence	Frequency of access readings and evaluations (selection box)
	ENG3 – Self Reported Commitment	Perceived commitment throughout the course (selection box)
	ENG4 – Dedication	Hours per week dedicated to the course (multiple choice)
	ENG5 - Assiduity	Days per week in which the students accesses the course area (multiple choice)

Source: Adapted from Sawang, Newton e Jamieson (2012), Perna et al. (2014), Greene, Oswald e Pomerantz (2015).
* The items for this study were inspired by these studies.

Causal relations

To compose the research model, it is important to determine the paths that the determinants of satisfaction and continuance to use do. Hew (2014) in his research has gained qualitative insights into MOOCs that can help scale these paths. Through Self-Determination Theory (SDT), Hew (2014) measured what the MOOCs students consider important, in order of importance, the following aspects were listed: problem-centric learning with clear expositions, instructor accessibility and passion, active learning, peer interactions and using helpful course resources.

These elements charge great resemblance to the constructs presented in this research, which in turn are supported by quantitative approach studies selected in the literature review. In fact, the categorizations and results found in the selected studies are in line with the indicators of the research model. From the studies on MOOCs, five of them, including Hew (2014), have empirically investigated MOOC users' opinions in a qualitative way: Abeer and Miri (2014), Chen and Chen (2015), Ventura, Bárcena and Monje (2014) and Franklin (2015).

In these studies, the dependent variables, being satisfaction and continuance to use are always present, and may have designations with similar function, such as success and effectiveness. In any case, all of them measure a positive feeling of user and the tendency to continue using e-learning and MOOC services. The satisfaction as used in the research is a heritage construct of marketing research and behavior of technology users. Measuring the

satisfaction of a product or service is a constant challenge, and involves many determining factors (Chyung & Vachon, 2013).

In bibliometric research, the use of satisfaction and continuance is a constant, regardless of the theoretical matrix or study object. Whether the study is about adoption of innovation in e-learning, adoption of a MOOC such a learning option, evaluation of the implementation of a new course, evaluation of students' performance or determinants of success, the satisfaction and continuance to use are always presented. Among the most commonly used theories, the IDT, TAM, EDT, TRI and others, such dependent variables also fall within the scope of various proposes, can be an adaptation, a replication or a new model.

About the engagement, it is a construct considered by many educators as an aspect of teaching and learning by influencing student retention. Being an abstract construct, it can have varying definitions and take on different roles throughout each study (Hew, 2014). Often, engagement is confused or operationalized as motivation, but while motivation seeks to measure the reasons of physical and cognitive actions, engagement is the observation of the manifestation of motivation. In this context, engagement acts as a mediator between the aspects of adherence of the MOOCs (MAI) and perceived performance and satisfaction. Considering the dependent variables, engagement is also an agent that can directly influence the intention of continuance to use.

Being specific in use of EDT, Sawang, Newton and Jamieson (2012) emphasize that the complexity or difficulty is an important factor to determine the success of an e-learning service or MOOC, and that it is difficult for users and researchers to understand what forces that can lead to disengagement and loss of satisfaction. Therefore, perceived performance can be an alternative to complexity, by uniting indicators that converge to an experience that aggregates knowledge, in this case, the high perceived performance.

The IDT model supports the closeness between complexity and performance, pointing that service use is as significant as difficulties, and that both walk together. In the TAM model, users evaluate extrinsic technology factors or use motivators, to explain acceptance; in this case, the performance permeates such motivators. In the context of EDT applied in e-learning services and MOOCs, user satisfaction is caused by a base of individuals who judge and evaluate a technology service (Sawang, Newton, & Jamieson, 2012). In general, when performance is operationalized, the possible difficulties are already contemplated and can be perceived if the indexes of perceived performance, engagement and satisfaction are low.

As well as performance, the particular aspects of MOOCs such as interactivity and collaborative learning also addresses the difficulties associated with a MOOC course. If for

performance, the theoretical support comes from EDT and other models of technology acceptance, such as TAM, IDT and TRI, for the so-called effectiveness of MOOCs, the theoretical bases come from the different learning styles.

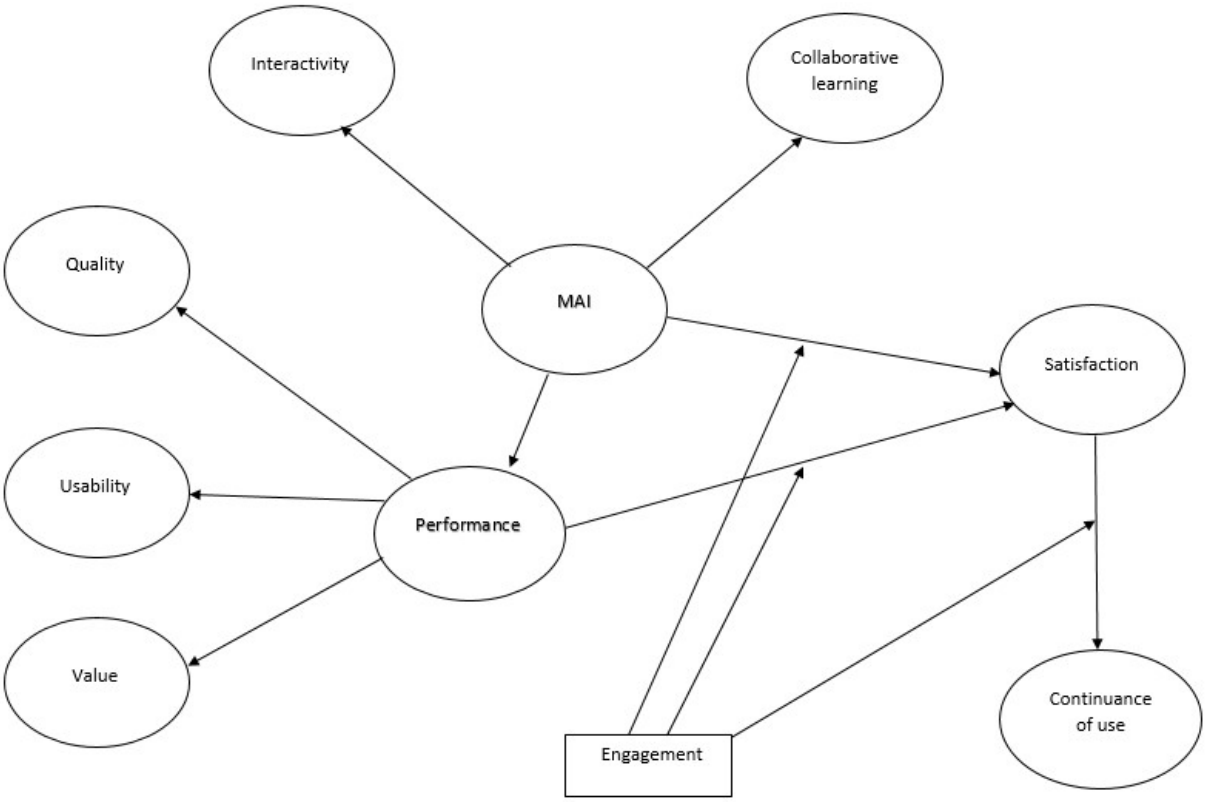
Chang, Hung and Lin (2015) point out that because of the heterogeneous nature of individuals' different ways of learning, this concept is a characteristic that has effects on MOOCs. Chyung and Vachon (2013) show that levels of satisfaction can be influenced by different factors from course area, by the content and actions of teachers and mediators, and these factors are precisely related to interaction and collaboration. Different ways of learning are so present in MOOC courses, that explain so much disparity in the evaluation of a MOOC course from student by student.

In this context, many factors contribute to the adoption, satisfaction and continued use of e-learning and MOOC services. The researches in the area are extensive, and despite the issue not be exhausted, it can be said that is a well-explored topic (Li, Duan, Fu & Alford, 2011). The MOOCs, being a special type of e-learning service, to generate a measure of success or effectiveness may have as determinants: features that are particular for MOOCs and the multiple aspects relevant to the use of technology, the same way that are used in the literature of e-learning (Wong, Tatnall & Burgess, 2014; Terras & Ramsay, 2015).

Derived from studies of MOOCs, there are collaborative learning and interactivity, two constructs which along with the performance contribute to the formation of the MAI (MOOCs Adherence Index). The MAI, itself, can present some results relevant to the study, such as the degree of effectiveness of MOOCs to the user, based on their own perception of the user experience with MOOCs.

The model of research is derived from studies about e-learning and presents the following constructs: performance, a construct formed by the constructs, quality, usability and value; satisfaction and continuance of use, as dependent constructs; and engagement, a mediator-moderator construct; in other words, this construct assumes role of moderator variable, assigning low and high scenarios of engagement, and also acting as a mediator variable, mediating the relationship between MAI and performance with satisfaction. The Figure 8 shows the model with causal relations.

Figure 8: Initial model of research



Source: Elaborated by author, 2018.

Regarding mediation assigned to the engagement, it is expected that the direct causal relationship between MAI -> Satisfaction and Performance -> satisfaction, be reduced or canceled by the mediator effect. The Sobel and Aroian tests determine the level of significance of the mediating effect, seeking to reject the null hypothesis that the mediating effect does not exist. Thus, two competing models can be assigned: in the first, there is the moderator effect of engagement, and in the second, the engagement has only direct relation with the levels of satisfaction and continuance of use. In both cases, it is expected that the indirect effects are low or zero.

4. Methodological procedures

This topic describes the theoretical and methodological basis of the research, the sampling plan, the history of the development of the research instrument and the data collection process, as well as discussing the data analysis plan, including the statistical procedures adopted and the flowchart of the study.

4.1 Research Characterization

This research is guided by logical empiricism. The empiricist orientation is widely diffused in studies of the area of administration, especially in the subjects related to finance, marketing, consumer behavior, information technology and technology acceptance.

Empiricism was devised at the Vienna Circle, where questions about being and existence were raised. A new philosophy of science called logical positivism was proposed. The logical positivists redefined science as the study of meaningful statements about the world. For a statement to be meaningful it has to be verifiable, which is known as the verification criterion. It means that it should be possible to determine the truth of a statement (Scholten, 2017a).

A strict rule of positivist logic is the non-acceptance of unobservable entities, creating problems in several areas. In scientific terms, unobservable variables are indispensable for theoretical construction. From this impasse, a moderate or adjusted version of logical positivism has gained ground, the so-called logical empiricism (Scholten, 2017a). Karl Popper argued that we can never conclude or prove an argument with observations, but we can conclusively refute it with contradictory evidence, leading to the maxim: a statement is meaningful only if it is falsifiable.

In this context, Malhotra (2006) and Vergara (2008) point out the tendency to use only one research instrument to collect primary data, establishing relations between variables based on the reality studied. These relations in turn are based on statistical hypotheses, which in turn are dismemberments of theoretical hypotheses. The so-called hypothetical-deductive method combines induction and deduction, requiring falsifiability and confirmation as the provisional support of a hypothesis (Scholten, 2017a). Thus, this study can also be characterized as empirical, of an original scientific nature and following the hypothetical-deductive method.

The analytical approach of the research is quantitative, being the most suitable to reach the proposed objectives. Based on the descriptive and inferential statistics, specifically with the use

of multivariate analysis, the theoretical model containing the hypotheses is tested and validated empirically. Finally, the study presents a cross-section on time, by measuring the face of a social reality at a given moment, stressing that the use of disconfirmatory constructs, despite rescuing a past reality, does not express a longitudinal character.

4.2 Sampling plan

The study covers students / users of MOOC courses offered free of charge or payment without the imposition of legal entities, who have already attended or are attending at least one course. The MOOCs are offered by numerous universities and public and private entities through platforms, containing a catalog of courses, divided by areas and with a teaching area per course. Anyone with internet access can enroll and follow the course. However, the possibility of obtaining a certificate depends on the course and the platform, can be free or paid.

The type of sampling in this study is non-probabilistic sample by quotas. At first, it was researched which MOOC platforms with the largest number of subscribers, then to get in touch with representatives of the platform or directly with those responsible for the courses. These contacts did not work, due to the teachers' refusal to include the research instrument in the course area. Among the justifications of teachers, the most recurrent was that the platform already has a user evaluation tool or teachers already had a simplified instrument developed by the teachers themselves.

A second strategy to reach users was to enter the course area and obtain the social networking address of the course. With this strategy, access to users was facilitated. However, it would be necessary to select courses by platform and select social networks to disseminate the research instrument. The platforms were selected by the number of subscribers and were present in large numbers in social networks. Among the chosen are Coursera, EdX, Future Learn, Udacity, MíriadaX and Iversity. As for the social networks of each course, Facebook is the most recurrent and with greater participation of users. In addition, on Facebook, the courses are organized by private groups, with input upon request to administrators. In this way, the research tool was shared in the Facebook groups of each course.

The choice of courses to conduct the research was a proportional selection of the number of courses in each platform and within each platform, with a proportional selection of courses in each area. In this way, the sample was obtained from the sharing of the research instrument

in four hundred private groups of Facebook, each group being part of a course of one of the MOOC platforms selected for this study.

The search instrument link with an initial text was shared in each course group on Facebook. In a first round, four groups were randomly selected to obtain fifty responses, composing a pretest sample. After descriptive analysis, the research instrument was reviewed by five experts in the field. The wording of three questions was changed because they were dubious and because they presented more than one probability distribution. A second round of pretest was conducted in other four randomly selected communities. Again, the data were analyzed, evaluating the violation to the normal distribution, presence of outliers and variability.

After the two rounds of pretest, a final version of the research instrument was obtained, proceeding with the collection in the other 392 groups. The collection process began on January 15, 2018, ending on January 31, 2018. The collection process was guided by the sample calculations.

The research instrument was formulated from the theoretical matrix formed after bibliometric and systematic analysis. The instrument consists of a questionnaire applied in the form of an online survey, being filled out voluntarily and random choice of respondents; is composed of forty - seven closed questions, thirty - seven questions using a metric scale of agreement by semantic differential of 0 to 10, being 0 - fully disagree and 10 - fully agree. Such a scale has advantages over others, such as the Likert scale, because it substantially reduces the response time of the questionnaire, provides more information than scales with fewer points, and facilitates the measurement of behavioral variables, because it more easily meets the assumptions of multivariate statistical techniques (Hair et al. 2011). Of the forty-seven questions, forty-two were presented in the topic of the research model, the five others are responsible for identifying the demographic profile of the sample.

4.3 Sample size

The sample calculation in this research is relevant as it statistically represents if the portion of the reality collected is sufficient to explain the phenomenon studied (Nebojsa, 2014). The population in this study is infinite because of the difficulty in measuring how many users there are in the main MOOCs platforms. Furthermore, it is not known the minimum or maximum

percentage of population elements that can be accessed, precisely because an unknown portion of the users of a platform does not have Facebook.

For the calculation of the sample, G*Power 3.1 was used. It is a simple yet powerful calculator, which features a priori and post-hoc sample tests. In the sample calculations, the level of confidence and statistical power are calculated, in addition to the size of the effects of exogenous variables on endogenous variables, according to studies by the developers, such as Faul et al. (2009).

The size of the effects (f^2) demonstrates the changes in R^2 in the relation between predictor and dependent variables. The f^2 has been used as a measure of reliability in structural equations via partial least squares (PLS) and as an adjustment measure for sample calculations with confidence level and statistical power set. The estimate of size of the effects can be calculated by the ratio between the partial R^2 and the unexplained portion of variance. In the a priori sample calculation procedure for use in factorial and regression analyzes, f^2 can be defined as the ratio between the explained and unexplained variance, with the sum of the variances being equal to 1, as presented next.

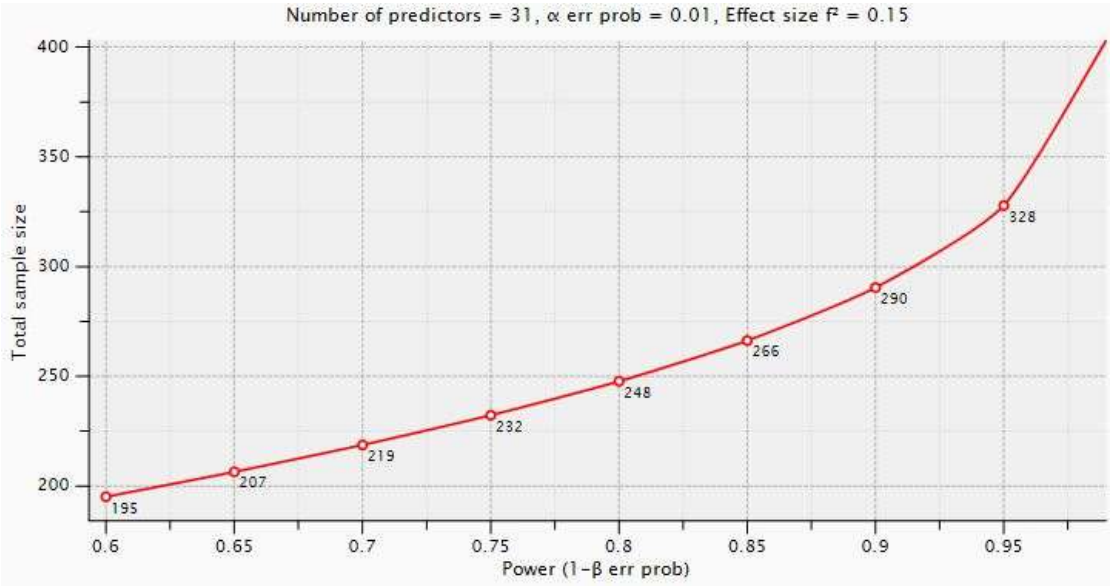
$$f_2 = \frac{V_S}{V_E} \text{ with } (V_S + V_E = 1) \qquad (3)$$

Where V_S is the portion of variance explained and V_E is the residue of the variance. Thus, V_S can be estimated with conventional values of f^2. According to Faul (2017), a small f^2 is 0.02 and an average value is 0.15. In G*Power, the effect size chosen was 0.075, being found by testing different values of V_S above 90% as to follow.

$$V_S = \frac{f_2}{1 + f_2} \qquad (4)$$

In the calculation of the minimum sample, performed a priori, the input values were stipulated for confidence level, at 99%, with sampling error at 1% and 31 predictor variables. The results can be visualized in Graphic 9. These input values represent the complexity of the model and the level of precision required of the parameter estimates.

Graphic 9: Sample required - a priori

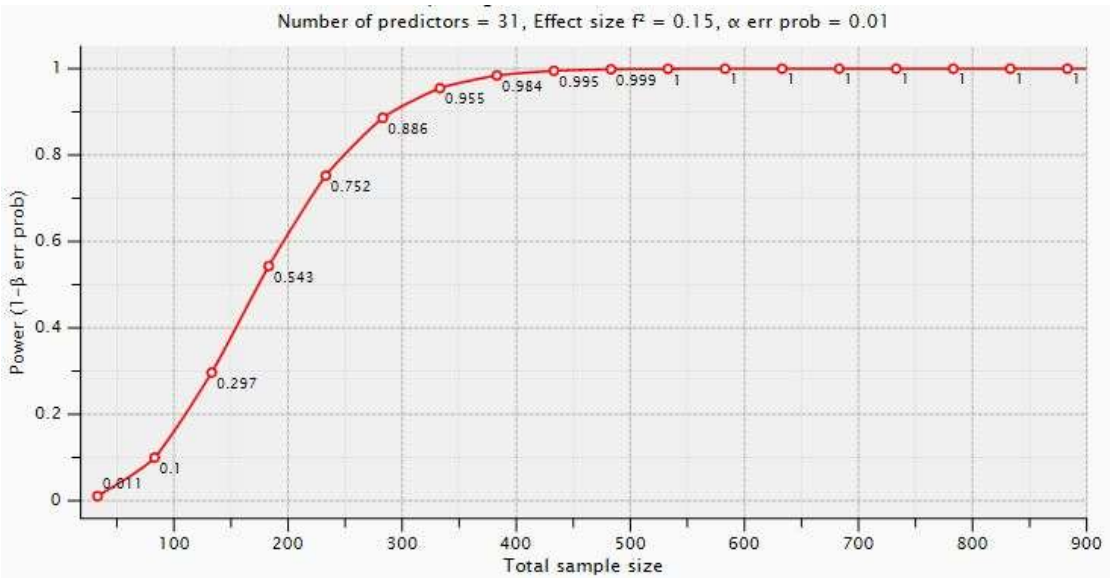


Source: Research data, via G*Power, 2018.

According to Hair et al. (2011), an adequate statistical power should be at least 0.8 or 80%, being ideal to obtain 0.95. Statistical power is presented for different sample sizes in Graphic 9. To obtain 0.95, 328 observations are required. With the use of moderating variables in the structural model tested, to use two groups in a multi-group analysis, the required sample size doubles, with the use of three groups, triples, and so on.

A total of 890 observations were collected, being a relatively higher number than the minimum sample, and sufficient to perform a multigroup analysis with moderating variables of two groups. The post-hoc analysis was performed to verify the statistical power with this sample size. According to the Graphic 10, a sample between 500 and 600 would already be enough to obtain maximum statistical power.

Graphic 10: Sample obtained - post-hoc test



Source: Research data, via G*Power, 2018.

Although considering the minimum sample based on the complexity of the model, certain factors such as the minimum sample requirement for the use of a technique were also evaluated. The possibility of using multi-group analysis from moderating variables also impacts the sample calculation, since the sample should be doubled, triplicated and even quadrupled depending on the number of groups of each moderator variable used.

Regarding the use of structural equation modeling by covariance (CBSEM), Maroco (2014) and Hair et al. (2011) ensure that it is appropriate to have at least ten observations per manifest variable. With the use of thirty-one predictor variables and six dependent variables; and using in the calculation 10 observations per manifest variable, a sample of 370 observations are required for use of the CBSEM in this study. Using a moderator variable with two groups, the sample requirement doubles, reaching 740 observations. Table 15 reveals the different sample requirements.

Table 15: Meeting the requirements of the sample

Requirements	Reference values	Sample
Model complexity	31 predictor variables; statistical power = 0.95 and confidence level = 0.99	328
CBSEM	10 observations per manifest variable (10 x 37)	370
Moderators in model research	Multi-group analysis with 2 groups (x2)	740
Sample obtained		**890**

Source: Research data, 2018.

4.4 Analysis plan

For the analysis of the data collected, the software Statistical Package for the Social Sciences (SPSS v. 021) and Analysis of Moment Structures (AMOS v. 021) were used. In addition, the R language was used to generate statistical results not available in the software used. With R, the R-Studio graphical interface and the R-Markdown tool were used as support for the analyzes generated in R. Some packages of R were used, such as: dplyr, ggplot2, ltm and caret.

The database was exported from the Google Forms online tool in the Google Drive application suite, the platform on which the survey was hosted. From a text file with extension "comma delimited value", the data was encoded in UTF-8 and imported into SPSS. The use of the data in the R was given in text format, needing only to create the object with the read.table function.

The analysis process first went through the data treatment, diagnosing and treating missing data and univariate and multivariate outliers. Then, an exploratory analysis of the data was performed, based on descriptive statistical results, with measures of central tendency and variability; graphs such as box-plot, histograms and scatter; correlations and tests of means comparisons. At the end of the exploratory analysis, the research model is tested through exploratory and confirmatory multivariate analysis techniques, with its due assumptions and validation tests being considered.

The statistical procedures used in the CBSEM are described in chronological order, as shown in Table 16, where the statistical tests and procedures for adjusting the factorial model and the structural model are presented.

Table 16: Statistical procedures in CBSEM

Modification Indexes	**Procedures**
Exclusion of observations detected as outliers	Mahalanobis Square Distance (D^2)
Checking of correlations between errors	Need to estimate correlations between errors
Adjustment quality Indexes	**Reference values**
RMR (Root Mean Square Residual)	The smaller, the better
TLI (Tucker Lewis Index)	>0.80 – good adjustment
CFI (Comparative Adjustment Index)	> 0.80 – good adjustment
NFI (Parsimonious Adjustment Index)	> 0.80 – good adjustment
PCFI (Parsimonious CFI)	] 0.6 : 0.8 [- good adjustment
PNFI (Parsimonious NFI)	] 0.6: 0.8 [- good adjustment
RMSEA (Root Mean Square Approximation Error)	] 0.05: 0.10 [- good adjustment
ECVI (Expected Cross-Validation Index)	The smaller, the better...
MECVI (ECVI adjusted)	The smaller, the better...
Assessment of assumptions	**Procedures**
Multivariate normality	Calculation of the asymmetry (sk) and Kurtosis (ku)
Linearity	Confirms the linearity of the model
Non-zero sample covariances	Must be 0 in endogenous variables
Multiple indicators	Presence of three or more manifest variables for each endogenous variable
Absence of multicollinearity	Calculation of Variance Inflation Factor (VIF)

Strong measure	Use of metric scale above five points generating discrete or continuous variables
Inexistence of *outliers*	Exclusion of cases with higher D^2
Validation of the measurement model (CFA)	**Procedures and references**
Factorial validation	Factorial load (ANASTASI, 1997).
Convergent validation	Average Variance Extracted (AVE). (FORNELL; LARCKER, 1981)
Discriminant validation	Composite reliability (CC) and Average Variance Extracted (AVE). (MAROCO, 2010).
Structural model validation (SEM)	**References**
RNFI (Relative Normal Adjustment Index) RPR (Relative Index of Parsimony)	(MAROCO, 2010; MULAIK *et al.* 1989)

Source: Research data, 2018.

These statistical procedures were presented in chronological order and follow a logical chain that reflects the operationalization of the research. The statistical procedures are consequences resulting from the formation of the research model, as well as the process of making the research instrument and subsequent data collection. In Figure X, the steps of the research, from the first effort to obtain the initial model of the research, being this effort the bibliometric and systematic revision, until the validation of the final model found in the research.

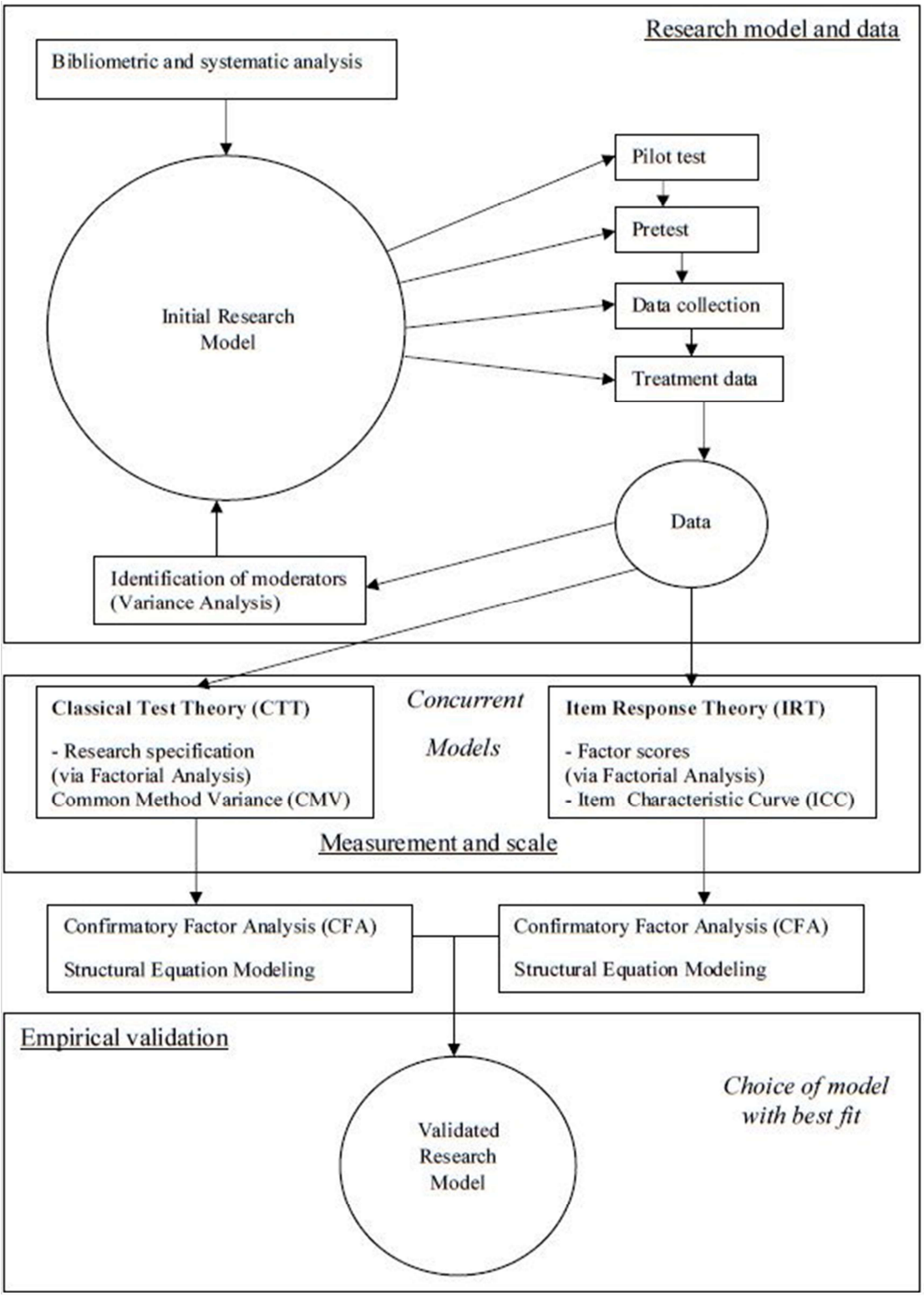
In order to obtain the data, the following were carried out: pilot study, pretest, data collection and treatment data. Then, through the family of variance analysis techniques, including ANOVA, ANCOVA, MANOVA and MANCOVA, the moderating variables were identified, adding more elements to the initial model.

In a second step, of measurement and scale, two competing models are tested in order to obtain the best exploratory factorial model and better preparation for the structural analysis. The Classical Test Theory (CTT) uses Exploratory Factor Analysis (EFA) as a specification research, and then the constructs are evaluated for Common Method Variance (CMV), according to the methodology of Podsakoff (2003). Concurrent to CTT, a factorial model is

generated from the statistical basis of the Item Response Theory (IRT), evaluating the IRT parameters and evaluating each model variable by the Item Characteristic Curve (ICC).

Finally, the last step is responsible for the empirical validation through CBSEM, starting with the Confirmatory Factor Analysis (CFA), to then include the second order constructs and the causal relations. The best model, that is, with better indices of quality adjustment is presented as the final model of the research. The flowchart of the study details the steps mentioned.

Figure 9: Study flowchart



Source: Elaborated by author, 2018.

5. Results

This chapter presents the main results obtained from the analysis of the data, starting with a descriptive analysis through summary measures, graphs and correlation index. Next, an exploratory analysis is conducted using the Item Response Theory (IRT) approach and Classical Test Theory (CTT) approach, with the common method variance (CMV), Exploratory Factor Analysis (EFA) and other strategies for exploring data. The structural model is then validated by two steps: the first by Confirmatory Factor Analysis (CFA) and by path analysis, after addition of causal relationships. The moderating variables are tested after the convergence of the structural model. For each generated output the software used, and the statistical procedures are explained briefly or in more detail if necessary to justify decisions about the theoretical model tested.

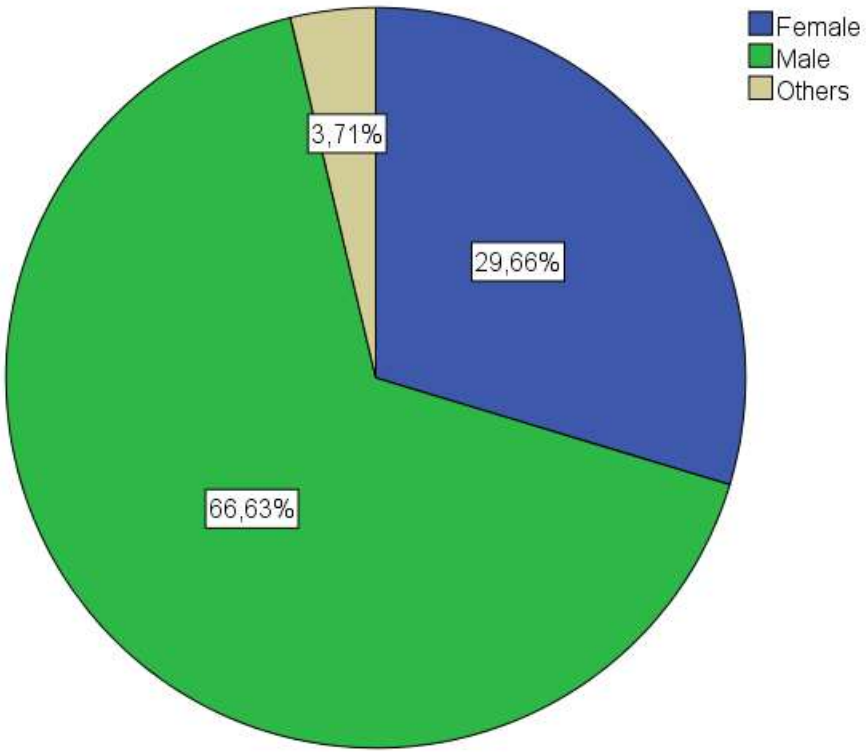
5.1 Descriptive analysis

In this first topic an investigation of the variables of the study is carried out, being summarized, crossed and correlated. The profile of the interviewee was traced, the classification of engagement of the individuals was performed, as well as the classification of other possible moderating variables and, finally, the manifested variables were summarized and correlated.

5.1.1 Profile

When analyzing the sample profile, composed of 890 observations, the objective is to synthesize the data and identify possible similarities and differences between the study variables. In addition, the variables of the profile may indicate the formation of distinct groups in the sample, being possible moderating variables. The first variable of the profile measured was the gender, with the categories: male, female and others. The results are presented in Graphic 11. It can be seen that the majority of the respondents (593; 66.6%) are male. The representation of the other gender, which includes the non-identification of known genders was marked by 33 respondents (3.7%).

Graphic 11: Gender

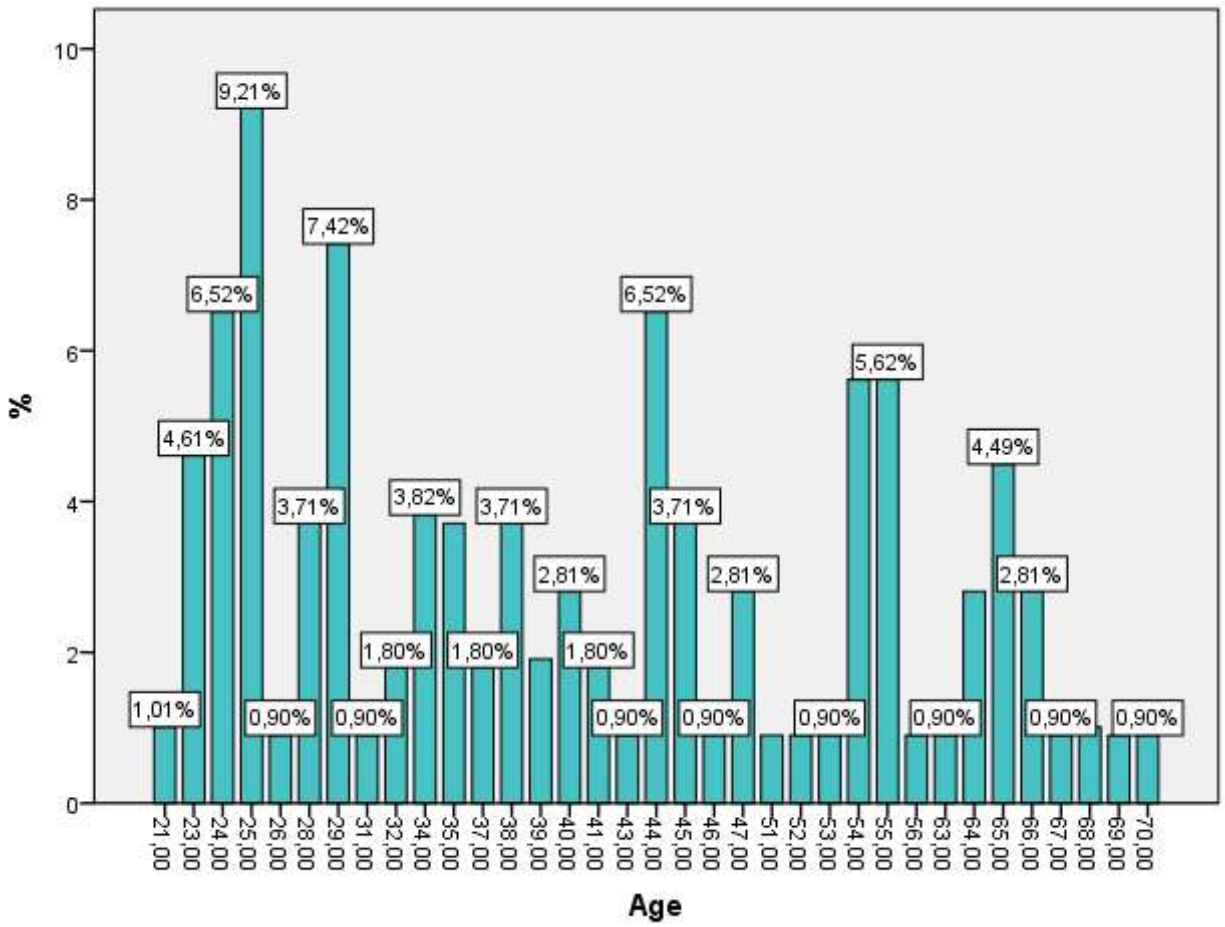


Source: Research data, 2018.

When they answered about the age, a discrete quantitative variable was created. In Graphic 10, the ages are arranged on the x-axis. It is a fact that among respondents, the youngest is 21 years old and the oldest is 70 years old. As for the minimum age, it is natural to be 21 years since most of the courses offered in MOOCs relate to contents of undergraduate and post-graduate courses. There are many respondents, represented by the age range up to 29 years. People aged 44 helps represent respondents in their 40-50 age range. After 50 years, the respondents with 54, 55 and 65 years increase the representation for this range. The mean age of the respondents was 41 years, with a deviation of 14 years, indicating a range between 27 and 55 years.

Thus, when looking at the Graphic 12, we can see three peaks in the data distribution, indicating the formation of three age groups. A more accurate analysis in the data exploration step may confirm this observation or indicate the formation of fewer or more groups.

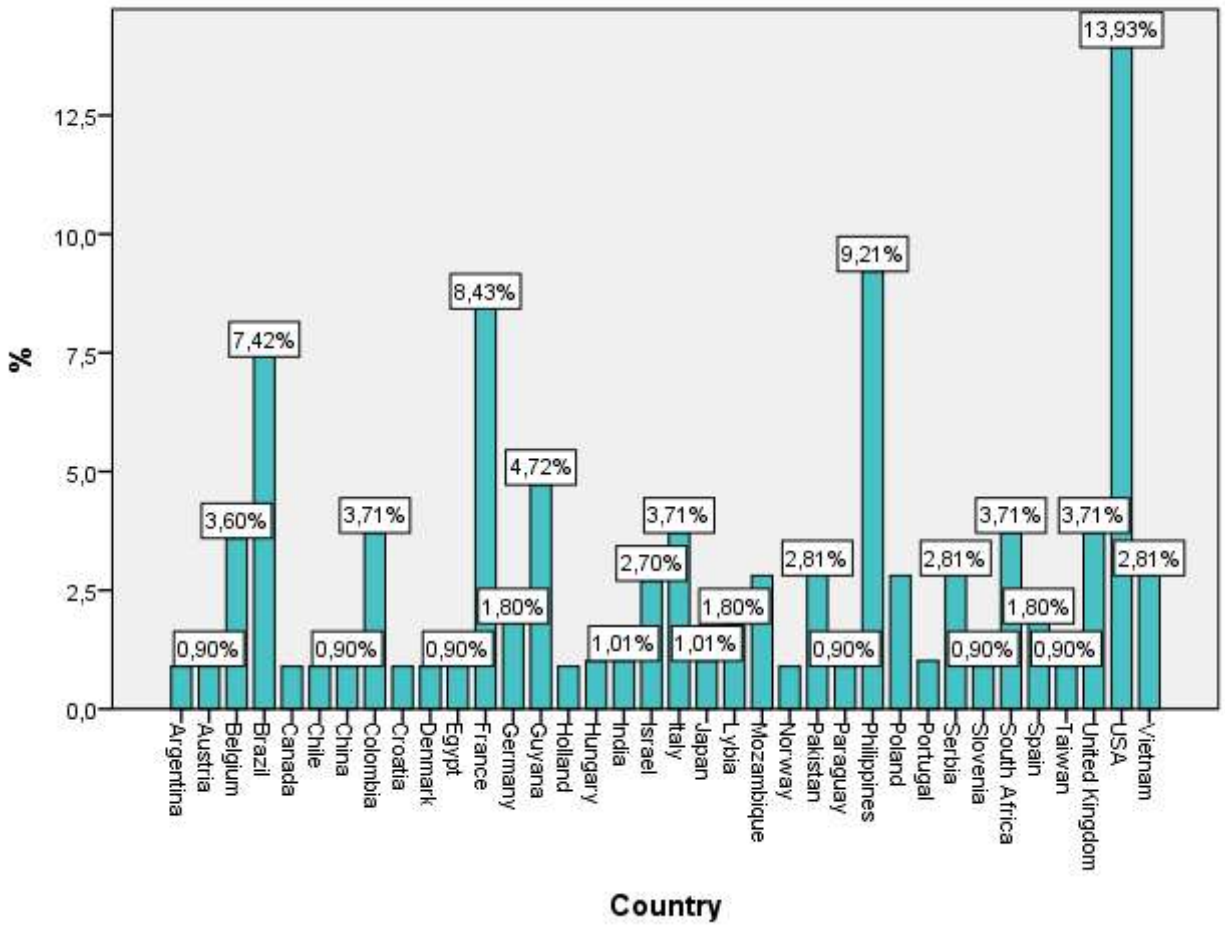
Graphic 12: Age



Source: Research data, 2018.

Next, the country of origin of the respondent was identified. For this question, the answer was opened. Thus, names of countries with the wrong spelling were changed. In all, respondents from thirty-six countries participate in the sample. The most representative country is the USA. The USA is the country with the highest number of MOOCs enrolled and has many MOOCs in its higher education institutions. Also, with outstanding representativeness are the Philippines, France and Brazil. By continents, the European, American and Asian continent are well represented, but the African and Oceanic continents have low representation in the sample.

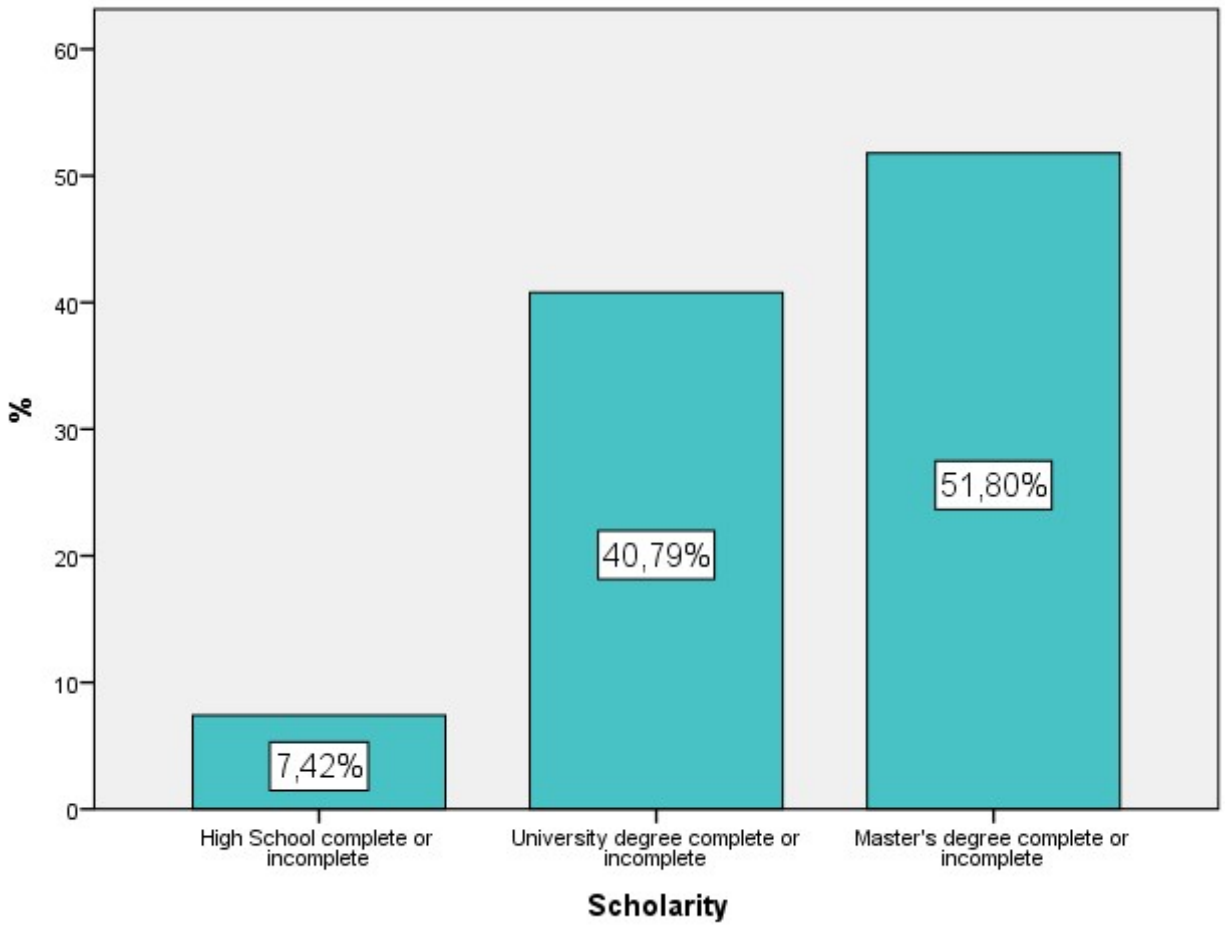
Graphic 13: Country of respondents



Source: Research data, 2018.

As the last demographic variable, the respondent's level of education was questioned, with three options without distinction between those who completed and those who did not. The results of Graphic 14 show the low representation of individuals who did not enter a higher education course. This result converges with the nature of the content addressed in the courses that are not geared towards high school generally. Nevertheless, the university public and especially those who have already entered the post-graduate studies represent the sample of this research.

Graphic 14: Scholarity



Source: Research data, 2018

Going beyond demographic variables, to profile was questioned for respondents if they have completed at least one MOOC course and if they have already paid to attend at least 1 course in MOOC platforms. Due to the collection process allied to the target audience of the research, it was expected that people already involved with MOOCs who completed more than one course would be more interested in participating in this research. Thus, Table 17 shows that most individuals have completed at least one MOOC course (758; 85.2).

Table 17: Summary of "Completed" variable

Have you completed at least one course MOOC?					
		Frequency	%	Valid %	Cumulative %
Categories	No	132	14.8	14.8	14.8
	Yes	758	85.2	85.2	100.0
	Total	890	100.0	100.0	

Source: Research data, 2018.

About the payment, just over half of the respondents said they have already paid at least one course (495; 55.6). This value agrees with the development of a business model in MOOCs that generates income for universities. However, a good part of the users still did not pay to use a MOOC platform (395; 44.4).

Table 18: Summary of "payment" variable

Have you paid for at least 1 course on MOOC platforms?					
		Frequency	%	Valid %	Cumulative %
Válido	No	495	55.6	55.6	55.6
	Yes	395	44.4	44.4	100.0
	Total	890	100.0	100.0	

Source: Research data, 2018.

Summarizing the characteristics of the interviewee's profile, this sample is predominantly male, with a different age and country of origin, who is at least attending a higher education course, who has already completed at least one MOOC course and who may or may not have paid to complete a MOOC course. Searching to verify possible group formations, cross tables were generated to identify if the payment of courses is more common for certain groups. It has been found that there are differences in gender groups and in the variable that asks if they have completed at least one course, as shown in Table 19.

Table 19: Cross table for gender vs payment and payment vs completed

Gender/Paid	**No**	**Yes**
Female	198	66
Male	297	296
Other	0	33
Paid/Completed	**No**	**Yes**
No	132	363
Yes	0	395

Source: Research data, 2018.

According to Table 19, most women did not pay for any course, while in the "other" group, all paid. In relation to men, there is a balance. In the cross-table between payment and completeness of the course, the majority of those who did not pay completed the course, but the most interesting result is that all those who paid also completed the course, evidencing that the payment can be an incentive to complete one or more courses. Next, the variables that measure the engagement are displayed, and eventually cross tabulations are presented.

5.1.2 Engagement

The engagement in this study was used as a construct formed by five variables that measure respectively: adaptability in distance learning, persistence, self-reported commitment, dedication and hours of study. Each variable was considered in the study in the same way as used in previous studies, which resulted in variables of different types, some nominal and ordinal categories and other continuous variables. Table 20 summarizes the results of each engagement variable, highlighting the scale used by each question.

Table 20: Engagement summary

ENG1 (Adaptability in distance learning): Taking a distance course is easy for me	
Yes = 890	No = 0
ENG2 (Persistence) : Check your actions that have already been done or what you will do	
10 categories: Access to 1st course reading; Access to part of the course readings; Access to all course readings; Access to the first evaluation instrument; Access to a part of the evaluation tools; Access to all assessment tools; Access to all contents of the first part (week) of the course; Access to all contents of some parts (weeks) of the course; Access to all content offered in the course.	

Mean = 5.70
ENG3 (Self Reported Commitment): Choose one or more options about your self reported commitment 5 categories: I am committed to the course; I intend to complete the course, or already conclude; I intend to get a good grade or have already; I intend to get a certificate or have already obtained; I intend to include in my curriculum the course certificate. **Mean = 3.18**
ENG4 (Dedication): Mark one option about your dedication (hours per week) 5 categories: There is no regularity (1); Fewer hours than suggested by the course (2); The hours suggested by the course (3); More hours than suggested by the course (4); More than twice the number of hours suggested by the course (5) **Mean = 2.89**
ENG5 (Assiduity): Mark one option about your access (times per week) 5 categories: Everyday (5); 4 to 6 times a week (4); 2 to 3 times a week (3); 1 time a week (2); A few times (1) **Mean = 3.15**

Source: Research data, 2018.

The first engagement variable (ENG1) is a nominal categorical variable, the ENG2 and ENG3 variables use categories that add up, can mark more than one option, generating discrete quantitative variables, the ENG4 and ENG5 variables use multiple choice, generating qualitative ordinal variables. In the first question (ENG1) the answer was unanimous in confirming that it is easy to conduct a distance course. This result is compatible with the nature of the MOOC courses, without obligation or need to study to obtain a compulsory certificate. Thus, it is inferred that people who have difficulty adapting to study at a distance are probably not part of the target audience of MOOCs.

In the other variables, the means were near the center point of each scale, indicating there is a balance in the sample of people with smaller and greater engagement. ENG2 and ENG3 are questions that respectively measure the actions of users on the platforms and their commitment to the course. The questions ENG4 and ENG5 respectively measure the regularity of access, standardized by the hours of dedication required in each course and the frequency of access on weekdays.

To use engagement in the structural model, some decisions had to be made. First, engagement provides theoretical support to be used as a moderating and mediating variable.

However, for the use of mediating variables in the modeling of structural equations, the mediator effect fails in the composition of the variable, since it is a statistical regression procedure, based on a latent reason scale. Thus, the relationship between Satisfaction -> | Engagement | -> Continuance of use cannot be measured for the purposes of this study. In other studies, with theoretical support for the use of variables with a ratio scale, the engagement can be tested as a mediating effect.

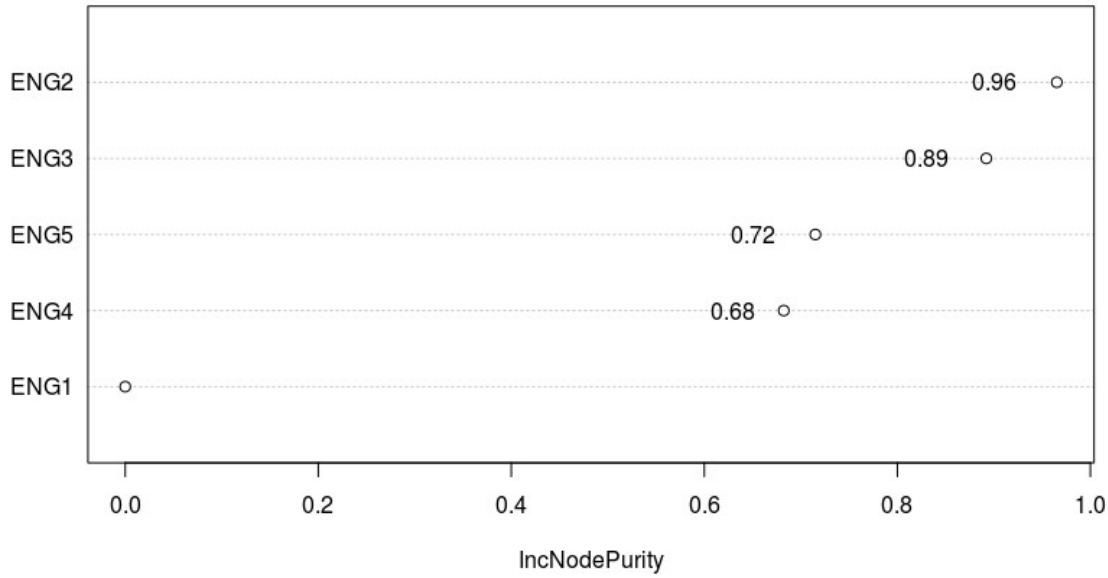
A second decision stems from the various types of variables composing the engagement. To classify the engagement groups, linear multivariate techniques require the use of discrete or continuous variables, as in cluster analysis. As an alternative, nonlinear classification techniques were considered.

In the R environment, the machine learning tools are in full development and aggregate several non-linear classification techniques. With the help of the "caret" package, which abbreviated the term classification and regression training, several models can be compared, indicating the one with the best fit (Kuhn, 2017). Of the several models that have different types of variables, two have obtained satisfactory adjustment, were the Support Vector Machine (SVM), which uses the "e1071" package by Meyer et al. (2015), and Random Forest (RF), which uses the "randomForest" package authored by Liaw and Wiener (2002).

Both techniques obtained convergence using two categories, low engagement and high engagement. The variable ENG1 in which all the answers were "yes", that is, that has facility in distance courses, did not contribute in the classification of the groups. The other variables in both models contributed significantly. However, random forest achieved a better ranking power, reaching 83%.

Since RF and SVM are machine learning techniques, a training sample must be used for the algorithm to classify the observations. To do so, the five engagement variables were used in the analysis together with a median based variable, calculated by the sum of the standardized scores of the engagement variables. Thus, the technique was conducted with two samples, the training, with 20% of the original sample and the test sample, with the other 80% of the original sample. 100 trees were generated as a standard measure, common to achieve results with less residuals. In addition to the classification power of the model, another result that is of interest is the ranking of the variables that best classify low and high engagement groups. Graphic 15 shows this ranking using the statistical test of Inc Node Purity, being the rate of variance of each variable in the model. The syntax for generating this graphic as well as the random forest model is given in Appendix A.

Graphic 15: Random Forest: variance level



Source: Research data, 2018.

It can be seen from the Graphic 15 that the ranking of the variables in descending order was: ENG2, ENG3, ENG5, ENG4 and ENG1. The most important variable to classify engagement is referent to actions inside the platform, also called persistence (ENG2). The actions involve the readings of the texts made available in the course, the accomplishment of the evaluation activities and the access to the contents generally available. Table 21 summarizes the results of the persistence by category.

Table 21: ENG2 - Persistence frequency

		Frequency	%	Valid %	Cumulative %
Valid	1	99	11.1	11.1	11.1
	2	66	7.4	7.4	18.5
	3	132	14.8	14.8	33.4
	4	132	14.8	14.8	48.2
	5	66	7.4	7.4	55.6
	6	33	3.7	3.7	59.3
	7	66	7.4	7.4	66.7
	9	33	3.7	3.7	70.4
	10	263	29.6	29.6	100.0
	Total	890	100.0	100.0	

Source: Research data, 2018.

According to Table 21, it is possible to identify two percentage peaks, the first one for individuals who scored 3 or 4 of the 10 available, collecting 29.6% of the sample. Those who scored the 10 options, reaching 29.6%. These categories support the results obtained in the random forest regression model.

The second variable with the highest classification power was the self-reported commitment, which measures the student's intention in five levels. In the same way as seen in persistence, compromise also presents representativeness at lower and higher levels, as can be seen in Table 22.

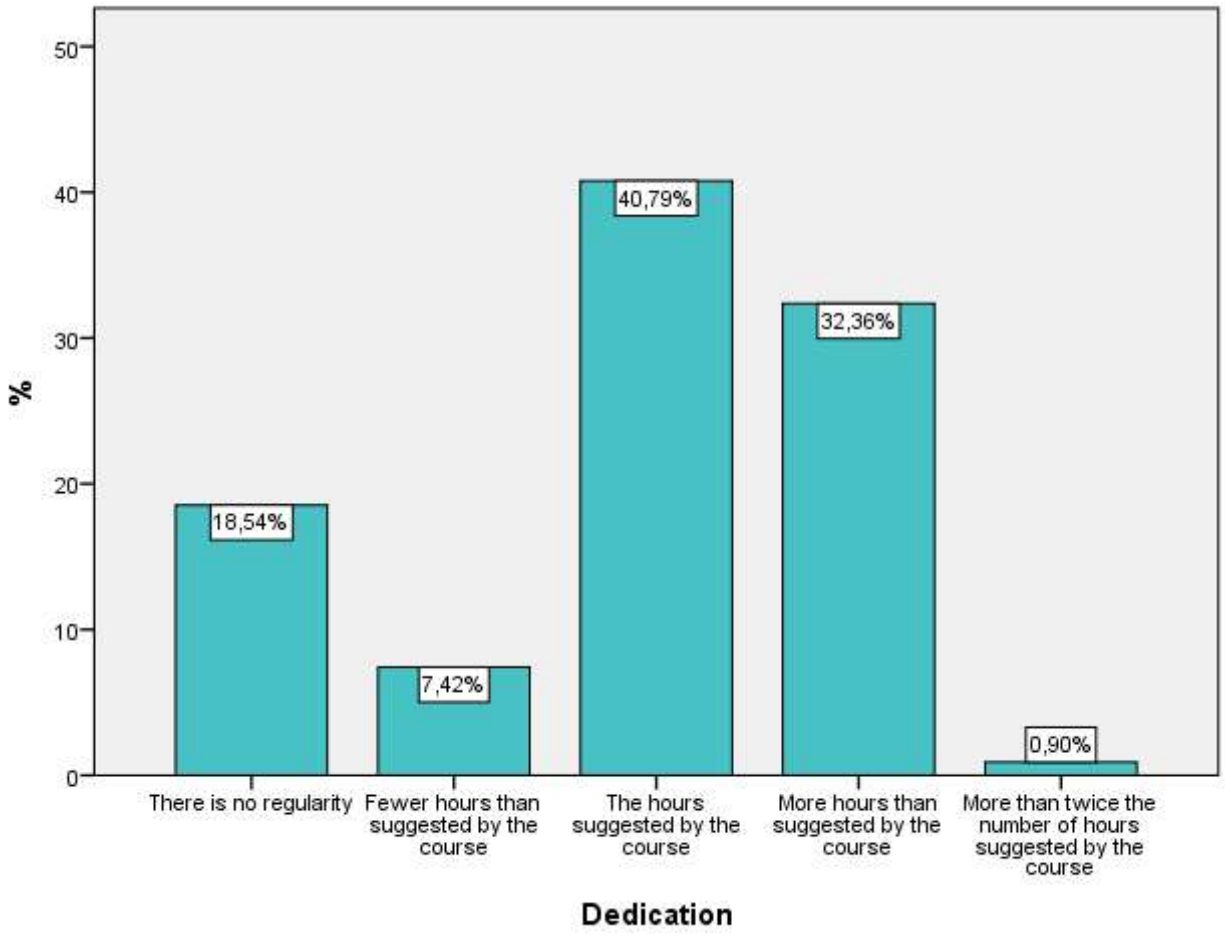
Table 22: ENG3 – Self-Reported Commitment

		Absolut	%	Valid %	Cumulative %
Valid	1	165	18.5	18.5	18.5
	2	198	22.2	22.2	40.8
	3	132	14.8	14.8	55.6
	4	98	11.0	11.0	66.6
	5	297	33.4	33.4	100.0
	Total	890	100.0	100.0	

Source: Research data, 2018.

As can be seen in Table 22, individuals who scored one or two of the five options together represent 40.8% of the sample, while 33.4% of the sample scored the five options available. The other variables: ENG5 and ENG4, measure attendance and dedication to the course, respectively. The following graphics detail the results of the variables by category.

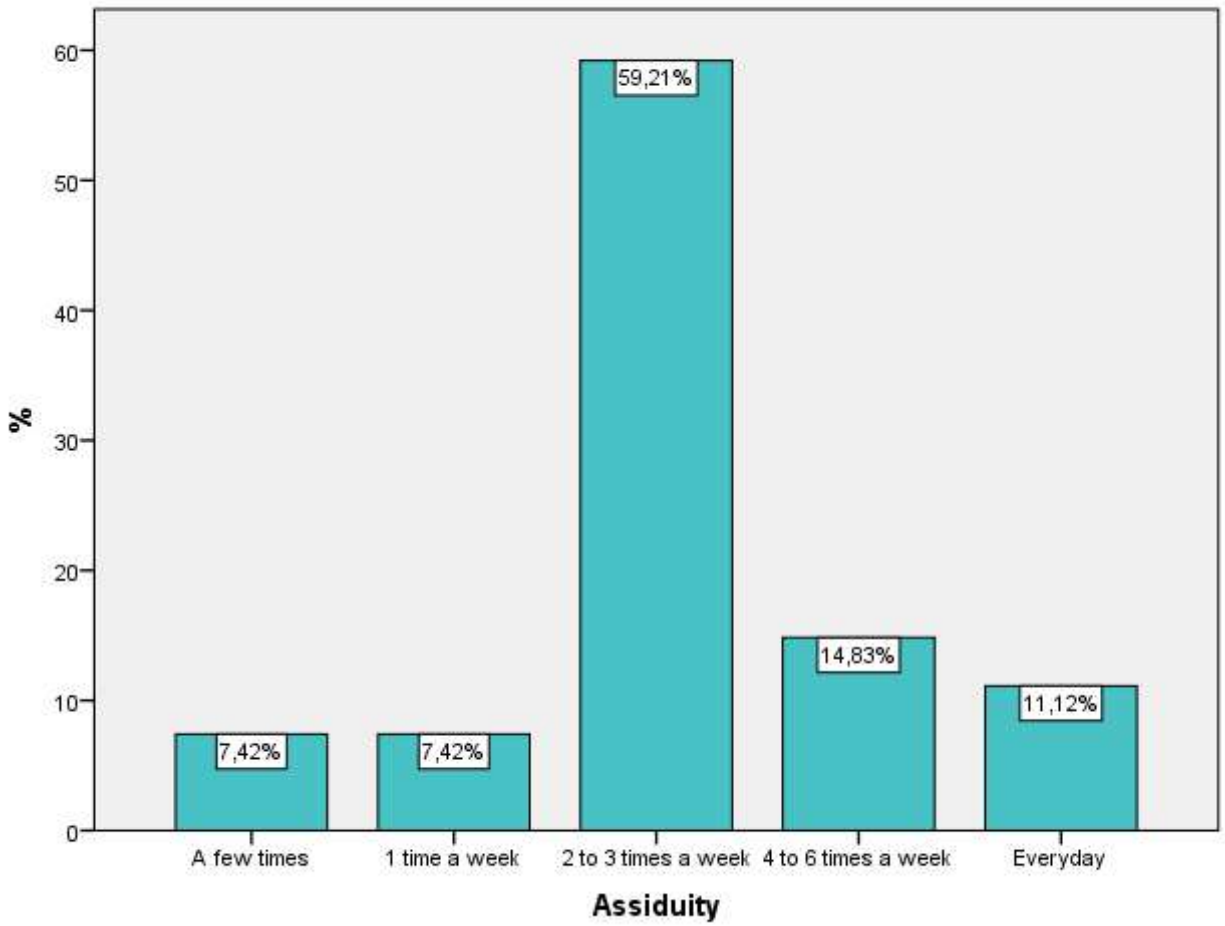
Graphic 16: ENG4 - Dedication



Source: Research data, 2018.

With a variance rate of 0.68 in the engagement ranking power, dedication to the course was measured in a multiple-choice question. It is observed that most of the respondents are concentrated in categories 3 and 4, indicating that they devote themselves in hours per week, the hours of the course being suggested by the teachers or more hours than suggested by the teachers. However, there is a significant portion of 18.54% that does not present regularity in the study in relation to the number of hours dedicated per week. The variable that measures attendance obtained a slightly higher variance of 0.72 and is presented as to follow.

Graphic 17: ENG5 - Assiduity



Source: Research data, 2018.

According to the Graphic 17, it is noticed that the assiduity is marked by more than half of the respondents in the range of 2 to 3 times in the week. This concentration in a range corroborates the results of the regression, since even with adequate classification power, no peaks were identified in different parts of the measurement scale. Thus, with the use of random forest, the low and high engagement classification variable was created and later used as a moderator variable of the structural model. In the next section, other variables next to the engagement were tested as possible moderating variables, using other analysis strategies.

5.1.3 Moderator variables

Other variables, such as those in the demographic profile, were tested to see if there were differences of means throughout the groups in the dependent variables. The variables tested

were engagement, gender, age, paid and scholarity. The selection procedures to include a variable as moderator of the structural model were given through the family of techniques of analysis of variance. At first, cross-tabulations of each respondent profile variable, such as gender and age, were evaluated together with the six dependent variables of the study, using t test or ANOVA, verifying possible differences in means of satisfaction and continuance variables throughout the groups. Obtaining p-value below 0.05, the variables were evaluated jointly through the Multivariate Analysis of Variance (MANOVA).

However, before beginning the comparison of means, the variable age was measured as an open question in the questionnaire, generating a discrete quantitative variable. In this way, it was necessary to make it categorical by means of age groups. With the management of a hierarchical cluster, as recommended by Maroco (2014), solutions with 2 or 3 clusters were plausible. The solution with three clusters was generated first. By means of a cross-table, it was verified that the categories of respondents with a cluster center equal to 49 years of age and the category with cluster center of 58 presented similar means.

In the analysis of the solution with 2 clusters, it was possible to differentiate non-hierarchical groups, the variable age was then classified into two age groups, with cluster centers at 32 years and 58 years respectively. From the categorization of the variable age, cross tabulations were generated to verify the means of the dependent variables by each group of the categorical variable. The first to be analyzed was age, as shown in Table 23.

Table 23: Cross table for dependent variables and age

Clusters by age / Dependent variables			Sat1: Course performance	Sat2: Choose to take a MOOC	Sat3: Be involved in a MOOC	Con1: I want	Con2: I will continue	Con3: I will do regularly
Young	N	Valid	502	502	502	502	502	502
		Missing	0	0	0	0	0	0
	Mean		7.30	7.19	7.16	7.15	7.32	6.68
	Standard deviation		2.2664	2.4033	2.4768	2.5123	2.5550	2.8070
More young	N	Valid	388	388	388	388	388	388
		Missing	0	0	0	0	0	0
	Mean		7.16	7.10	7.12	7.25	7.31	6.74
	Standard deviation		2.3500	2.4075	2.5969	2.5210	2.5519	2.7112

Source: Research data, 2018.

According to Table 23, the means in the age groups are very close, and the difference between the means is much smaller than the standard deviation. These results indicate that a

test and comparison of means will hardly identify significant differences. Thus, the means of satisfaction and continuance of use are the same for the groups with more and less age. Thereafter, cross-tabulations were conducted for the other demographic variables. The second one was for gender. The variable is formed by three groups; however, the other group has few representatives and would hardly fit into a variance analysis model. Thus, the male and female groups were compared for the mean, as shown in Table 24.

Table 24: Cross table for dependent variables and gender

Gender			Course performance	Choose to take a MOOC	Be involved in a MOOC	I want	I will continue	I will do regularly
Female	N	Valid	264	264	264	264	264	264
		Missing	0	0	0	0	0	0
	Mean		7.20	7.12	7.16	7.16	7.24	6.57
	Standard deviation		2.3738	2.4343	2.5065	2.4756	2.5913	2.7251
Male	N	Valid	593	593	593	593	593	593
		Missing	0	0	0	0	0	0
	Mean		7.24	7.16	7.13	7.20	7.35	6.77
	Standard deviation		2.2986	2.4055	2.5494	2.5291	2.5446	2.7934

Source: Research data, 2018.

According to Table 24, results similar to age are found in gender. Valuation of men and women is practically the same, with mean differences close to 0. Similarly, the means of the dependent variables were crossed with the educational variable, measured in three ranges. The results also show a low average difference, as can be observed in Table 25.

Table 25: Cross table for dependent variables and scholarity

Scholarity			Course performance	Choose to take a MOOC	Be involved in a MOOC	I want	I will continue	I will do regularly
High School complete or incomplete	N	Valid	66	66	66	66	66	66
		Missing	0	0	0	0	0	0
	Mean		7.07	7.31	7.16	7.11	7.29	6.28
	Standard deviation		2.3067	2.1534	2.4143	2.4168	2.4919	2.8936
University degree complete or incomplete	N	Valid	363	363	363	363	363	363
		Missing	0	0	0	0	0	0
	Mean		7.17	7.06	7.06	7.16	7.20	6.59
	Standard deviation		2.3593	2.4339	2.5792	2.5878	2.6302	2.7766
Master's degree complete or incomplete	N	Valid	461	461	461	461	461	461
		Missing	0	0	0	0	0	0
	Mean		7.32	7.20	7.21	7.22	7.42	6.86
	Standard deviation		2.2589	2.4167	2.5074	2.4754	2.4991	2.7310

Source: Research data, 2018.

Another variable analyzed was payment, which measures through of a nominal categorical variable if respondents have already paid to perform at least one MOOC course. The results, as shown in Table 26, showed the same behavior of the other cross tables, with averages between groups close to 0. Light differences of means can be noticed, but they do not approach the amplitude of 1 standard deviation.

Table 26: Cross table for dependent variables and payment

Have you paid for at least 1 course on MOOC platforms?			Course performance	Choose to take a MOOC	Be involved in a MOOC	I want	I will continue	I will do regularly
No	N	Valid	495	495	495	495	495	495
		Missing	0	0	0	0	0	0
	Mean		7.29	7.22	7.22	7.28	7.41	6.82
	Standard deviation		2.2904	2.3917	2.5281	2.4466	2.5490	2.7468
Yes	N	Valid	395	395	395	395	395	395
		Missing	0	0	0	0	0	0
	Mean		7.18	7.07	7.05	7.07	7.20	6.56
	Standard deviation		2.3202	2.4203	2.5293	2.5969	2.5545	2.7829

Source: Research data, 2018.

Finally, the categorical variable engagement, created from the regression analysis by random forest, was crossed with the dependent variables. The mean differences were notorious among the high and low engagement groups. In Table 27 it is possible to notice that the low engagement group obtained averages below 6 points on the scale of 0 to 10, while the high engagement group obtained averages much higher. These results suggest that a comparison test of means should be conducted for engagement.

Table 27: Cross table for dependent variables and engagement

Engagement			Course performance	Choose to take a MOOC	Be involved in a MOOC	I want	I will continue	I will do regularly
Low	N	Valid	402	402	402	402	402	402
		Missing	0	0	0	0	0	0
	Mean		5.45	5.23	5.17	5.33	5.39	4.71
	Standard deviation		2.0223	2.0420	2.1570	2.1979	2.2977	2.3480
High	N	Valid	488	488	488	488	488	488
		Missing	0	0	0	0	0	0
	Mean		8.72	8.73	8.77	8.71	8.90	8.35
	Standard deviation		1.2241	1.2607	1.4067	1.5536	1.4062	1.8521

Source: Research data, 2018.

Although the mean difference was low along the cross-tables with age, gender, scholarity and paid. The respective tests were conducted, and it was statistically verified that the null hypothesis that there was no difference of averages was not rejected. For the engagement the result was different. First, the test to be applied considers the number of groups of the categorical variable and the number of dependent variables. According to the classification by Hair et al. (2011), when a dependent variable is used, and two groups are evaluated in each independent variable, the most appropriate test is the t test for independent samples.

The hypotheses for the t test can be postulated as follows:

H0: There are no significant differences of means in the dependent variables of satisfaction and continuance of use of MOOC users among low and high engagement groups.

H1: There are significant differences of means in the dependent variables of satisfaction and continuity of use of MOOC users among low and high engagement groups.

For the application of the t test, are evaluated for the independent variables if these do not violate the assumption of normality and if the population variances are homogeneous or homoscedastic. For that, the asymmetry and kurtosis indexes were used to evaluate the

normality and the Levene's test to evaluate the homoscedasticity. Regarding normality, the asymmetry and kurtosis indexes are between -3 and 3, and these reference values are suggested by Maroco (2014) to affirm that there is no violation of the normality assumption.

The Levene's test in turn is responsible for verifying if the population variances are homogeneous along k population averages. In the event of violating the homoscedasticity assumption, this violation could be related to the imbalance of the size of the groups: 402 for low engagement and 488 for high engagement. In this case, there is clearly a violation of this assumption, but the student t test with Welch's correlation dispenses mathematical transformations and corrects the imbalance of the groups. To test whether the population averages are equal, the following statistical hypotheses are proposed:

$$H_0: \mu A = \mu B \qquad vs. \qquad H_0: \mu A \neq \mu B$$

It is noticeable that the hypotheses: null and alternative are bilateral, because we want to test if the population averages are the same or different. If the population variances are not homogeneous, such as the variables sat1, sat2, sat3, con1, con2 and con3, the test statistic is called the Welch t test, and can be obtained by the following formula:

$$t = \frac{\mu A - \mu B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} \qquad (5)$$

The population means are subtracted and divide the root of the sum of the deviation of each group over the size of each group. With the homogenous population variances, the standard deviation is set, with the test statistic being used:

$$t = \frac{\mu A - \mu B}{S\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \qquad (6)$$

Welch's t-test was conducted for each of the six variables, encompassing the Levene's test results, the test result, with its respective p-value, and the confidence interval of the mean difference between the evaluated groups. These results can be seen in Table 28.

Table 28: t test for satisfaction and continuance versus engagement

Variables	Levene's test		Test t for independent samples					
	F	Sig.	t	Sig.	Difference Mean	S.E	95% C.I for difference	
							Lower	Upper
Course performance	93.685	0.000	-28.352	0.000	-3.26	0.1151	-3.49	-3.03
Choose to take a MOOC	72.636	0.000	-29.949	0.000	-3.49	0.1168	-3.72	-3.27
Be involved in a MOOC	60.135	0.000	-28.799	0.000	-3.60	0.1251	-3.84	-3.35
I want	51.817	0.000	-25.919	0.000	-3.37	0.1303	-3.63	-3.12
I will continue	89.613	0.000	-26.765	0.000	-3.51	0.1312	-3.76	-3.25
I will do regularly	24.764	0.000	-25.252	0.000	-3.63	0.1441	-3.92	-3.35

Source: Research data, 2018.

The results of the Levene's test prove that the violation of the assumption of homoscedasticity, resulting in the use of the equation mentioned. With values of t well above the critical t, it is perceptible that there is evidence that the null hypothesis was rejected for the six dependent variables of the study. The confidence interval of the mean difference remained low, contributing to the reliability of the results. The negative values of the means indicate that the group with high engagement obtained higher means along the variables.

Thus, of the five analyzed variables, only the engagement obtained significant differences of averages and has the potential to be a moderating variable in the structural model, moderating the causal relationship between satisfaction and continuance of use. In the next topic, the manifest variables are synthesized by summary measures of central tendency and variability. In addition, the relationship between construct variables is measured using Pearson's correlation.

5.1.4 Manifest variables

For descriptive purposes, the thirty-seven manifest variables were summarized, seeking to investigate the behavior of the respondents along each variable measured and aggregated to the latent factor of which each variable is a part. Two measures of central tendency are calculated, the unweighted arithmetic mean and the median. The mean synthesizes the data by identifying the center of gravity of the distribution, while the median informs the point on the scale that divides the respondents into two halves.

Two other measures measured are those of variability, standard deviation and coefficient of variation. These measures allow to know the dispersion of data around the mean, being a

marker of the representativeness of measures of central tendency. Finally, the minimum and maximum scale were identified by completing the summary measures used.

Together with the summary measures, the variables of each construct were correlated through the Pearson correlation index, which is a measure that standardizes the covariance result, offering greater interpretability on the existence of a relationship between two variables and their magnitude. The results are presented for the variables of each construct and the scale used in each question is a ten points scale, through a metric scale of semantic differential, as reported in the chapter on methodological procedures.

The first construct analyzed was the interactivity (INT), containing six items, which in turn integrates the second-order factor MOOCs Adherence Index (MAI), next to the construct "Collaborative Learning" (COL). The results for the INT items are shown in Table 29.

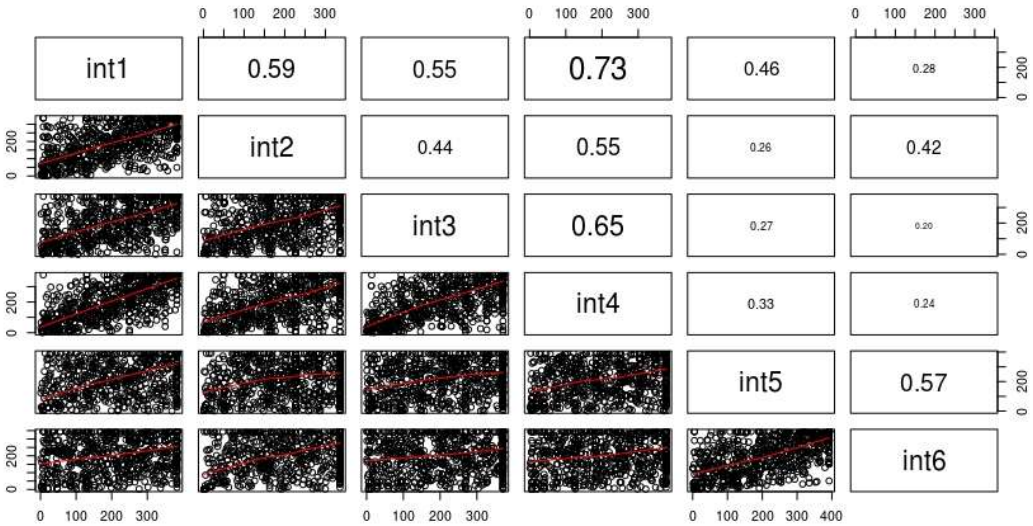
Table 29: Interactivity summary

Variable	min	median	max	mean	sd	CV
int1: Hypermedia	0	6.8	10	6.50	2.3975	0.3683
int2: Reading forums and chats	1	7.9	10	7.62	1.8402	0.2414
int3: Writing forums and chats	0	7.4	10	7.04	2.4444	0.3470
int4: Stimulate participation	0	6.8	10	6.55	2.4458	0.3730
int5: Doubts and deepening	0	6.6	10	6.40	2.3975	0.3742
int6: Interaction outside the platform	1	8.0	10	7.71	1.8192	0.2359

Source: Research data, 2018.

According to Table 29, the averages ranged from 6.40 to 7.71, indicating similar opinions regarding interactivity issues. The six questions present a controlled variability, as shown in the coefficient of variation (CV) indices, all being below 0.5. The question with the highest median and mean was int6, which measures the interaction outside the MOOC platform, that is, the respondents agreed more or judged more important to have interaction among course participants in virtual environments outside the teaching platform. Deepening the analysis of the items, the relationship between the variables measured, being presented in Graphic 18. The graphic contains below the diagonal the dispersion diagram of the variables in pairs, above the diagonal correlation indices are presented, with an increased source for indices with greater values in the scale of -1 to 1. The main diagonal identifies the variables by the code.

Graphic 18: Interactivity correlations



Source: Research data, 2018.

Graphic 18 reveals that the highest correlation indices were concentrated between the variables int1 to int4, with the highest index being the relation between int1 and int4 (0.73), both measures the use of multiple media and the stimulated participation, respectively. It is inferred that the use of multiple media favors the stimulation of participation. Similarly, int4 is the variable that nurtures relations of higher magnitude with int2 and int3, both measure actions in forums and chats. Despite the weak correlations of int5 and int6, among them the correlation is moderate (0.57). These variables measure the actions of questioning and delving into the content with off-platform interaction.

The second construct analyzed was collaborative learning (COL), also being part of the MAI. The means obtained in COL were slightly higher than the INT variables. As in INT, the variability of COL variables was low, represented by standard deviation and coefficient of variation. By sd, the variability magnitude is interpretable subjectively, but the CV, being below 0.5, standardizes the variability in a percentage measure. Table 30 presents these results.

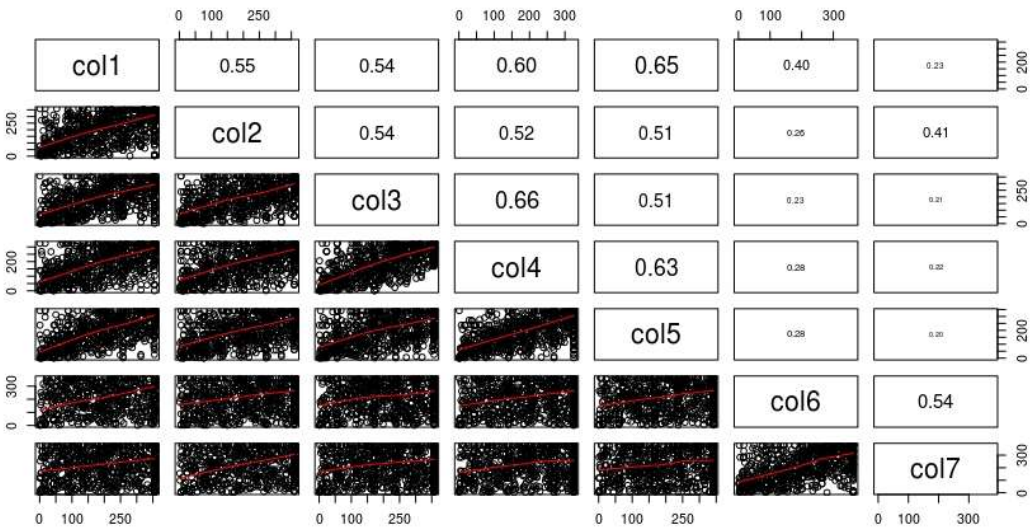
Table 30: Collaborative Learning summary

variable	min	median	max	mean	sd	CV
col1: correction of activities (Learning)	1	7.4	10	7.23	2.1016	0.29
col2: Correction of activities (Empathy)	0	7.3	10	7.16	2.1740	0.30
col3: Freedom to propose new forms to study	1	7.9	10	7.47	1.9289	0.26
col4: Freedom to indicate materials	1	8.1	10	7.88	1.9108	0.24
col5: Teamwork	1	7.9	10	7.53	1.9569	0.26
col6: Interest to share	0	7.6	10	7.28	2.1120	0.29
col7: Knowledge shared	0	7.1	10	7.00	2.3239	0.33

Source: Research data, 2018.

The variable with the highest mean was col4, which measures the freedom of the user to indicate study materials in the teaching platform (7.88). Shortly afterwards, with a median of 7.9, there are col3 and col5, which respectively measure the user's freedom to propose new forms of study and to contribute to teamwork. In fact, the averages oscillate between 7.0 and 7.88, maintaining a high correlation. In Graphic 19, the relations between the COL variables are set.

Graphic 19: Collaborative Learning correlations



Source: Research data, 2018.

According to Graphic 19, the trend line in the dispersion charts highlights linear and directly proportional relations between the variables. However, the variables col6 and col7 seem to be disconnected with the rest of the COL variables. The correlation between col6 and col7 is

moderate, but the correlation of these variables with the others is moderate to low. The variables from col1 to col5 on the other hand, correlate between 0.51 and 0.66. The first five variables measure the ways of correcting activities (col1 and col2), the freedom to propose and suggest (col3 and col4) and teamwork (col5). On the other hand, col6 and col7 measure aspects of content sharing.

After analyzing the variables of the MAI, with its two constructs, in this section are evaluated the variables that measure the factor of second order Performance, formed by three constructs: quality, usability and value. Each construct consists of six variables and all variables use the same ten-point scale. The quality variables are summarized in Table 31.

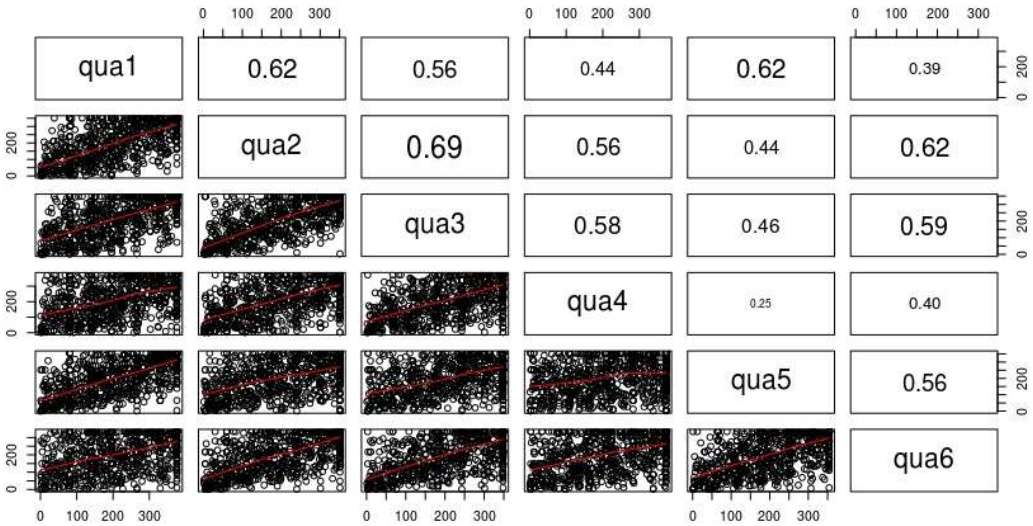
Table 31: Quality summary

variable	min	median	max	mean	sd	CV
qua1: Layout	1	7.1	10	7.07	1.9105	0.27
qua2: Navigation	0	7.5	10	7.34	2.0164	0.27
qua3: Comfort	0	7.9	10	7.47	2.0317	0.27
qua4: Completeness	0	7.0	10	6.78	2.2248	0.33
qua5: Content quality	1	7.4	10	7.29	1.8295	0.25
qua6: Support readiness	0	8.0	10	7.72	1.8593	0.24

Source: Research data, 2018.

According to Table 31, the averages remained high, above 7.07, with medians above 7.1. However, the highest mean reaches 7.72, qua6, which measures the readiness support. The variability remains low throughout the variables, reinforcing the representativeness of the mean. The relationships between the variables are shown in Graphic 20.

Graphic 20: Quality correlations



Source: Research data, 2018.

All correlations are presented as moderate, between 0.40 and 0.69. The highest correlation occurred between qua2 and qua3, which respectively measure navigation on the teaching platform and comfort related to the virtual teaching environment. Navigation also fosters a high relation with qua6, which measures support for readiness. Another relevant correlation is between qua1 and qua5, which measure the quality of the layout and the quality of the content offered. To compose a structural model, moderate correlations are adequate, since low correlations between variables indicate that one or more variables do not explain latent factors and, if they are very high, can cause a high multicollinearity effect.

Continuing the analysis of Performance constructs, the second analyzed is usability. This construct assesses the way in which users use the teaching platform, which consists of variables that evaluate tools and measure the impact of the use of usability tools. The variables are presented in Table 32.

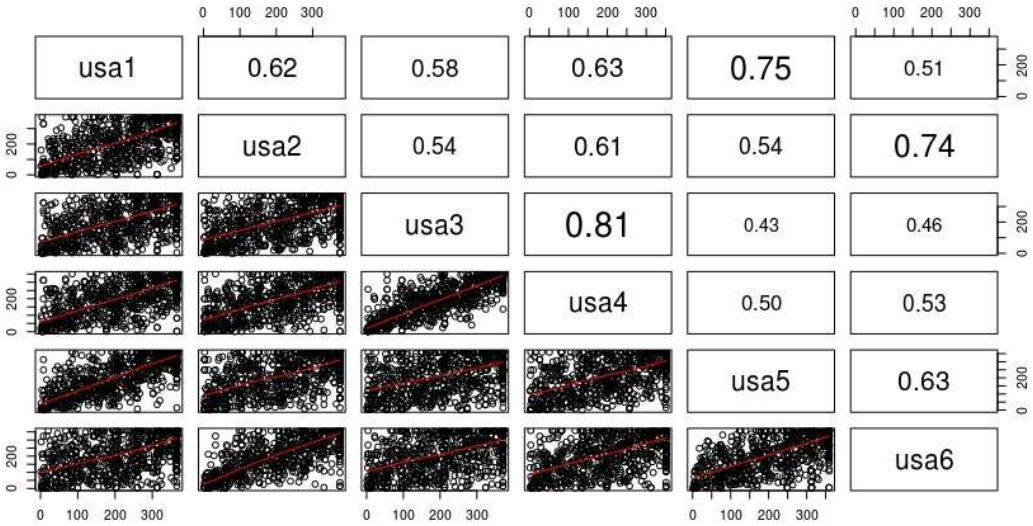
Table 32: Usability summary

variable	min	median	max	mean	sd	CV
usa1: Easiness	0	7.5	10	7.23	2.1449	0.29
usa2: Skill	0	7.2	10	6.96	2.3572	0.33
usa3: Features and activities	0	7.0	10	6.91	2.2237	0.32
usa4: Tools	0	7.2	10	7.07	2.1335	0.30
usa5: Content transition	0	7.9	10	7.40	2.0723	0.28
usa6: Support productivity	0	7.0	10	6.94	2.4186	0.34

Source: Research data, 2018.

Usability, as well as quality, has variables with averages close to 7, ranging from 6.91 to 7.40. The variability also remained low throughout the variables, as can be seen in the CV column. Despite the low amplitude of the mean, there are respondents who marked the 0 point in the scale, represented by the minimum column. The relationships between usability variables were also measured and are represented in Graph 21.

Graphic 21: Usability correlations



Source: Research data, 2018.

Different from the quality construct, there are usability variables with high correlations (> 0.7) in three relationships. The highest correlation occurs between usa3 and usa4, being 0.81. These variables measure activity options and the tools used in the education system. In the same way, there are high correlations in relations (usa1 <-> usa5; 0.75) and (usa2 <-> usa6; 0.74). These relationships make sense by the concept of the measured variables, usa1 and usa5

respectively measure ease of use and content transition. The relations between usa2 and usa6 refer to skills assessment and productivity support.

As the last construct of Performance, one has the value, composed of six variables and that measure the non-monetary value generation for the respondent. The variables with the exception of val4 present medians from 7 to 7.4. Val4 presented a median of 6.3 and a mean of 6.35. This variable measures the sense of intelligence, that is, if taking the course on the MOOC platform gives the feeling of greater intelligence. The results are shown in Table 33.

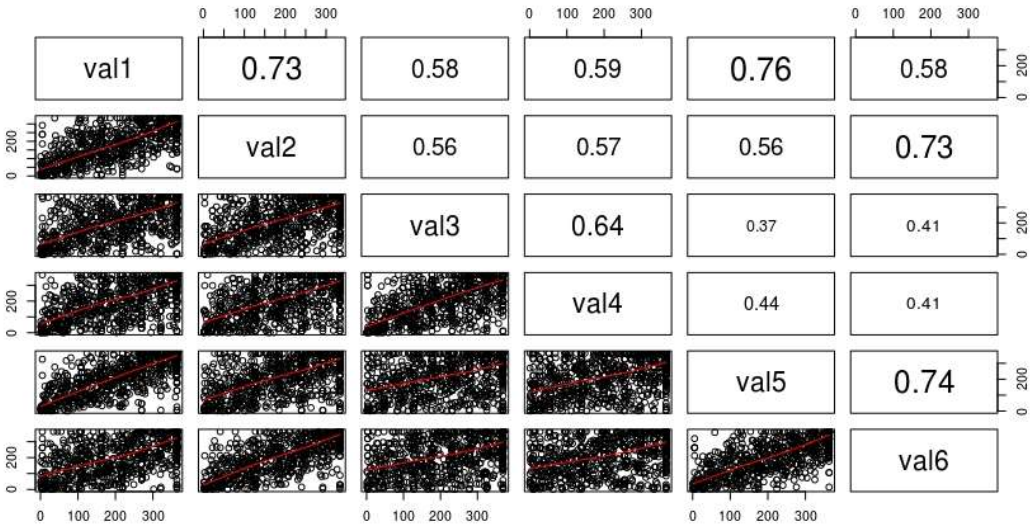
Table 33: Value summary

variable	min	median	max	mean	sd	CV
val1: Achievement	0	7.0	10	6.93	2.4813	0.36
val2: Needs	0	7.4	10	7.18	2.2275	0.31
val3: Follow trend	0	7.0	10	6.76	2.5235	0.37
val4: Intelligence sense	0	6.3	10	6.35	2.6953	0.42
val5: Independence	0	7.0	10	6.83	2.5643	0.37
val6: Entertainment	0	7.3	10	7.00	2.3955	0.34

Source: Research data, 2018.

The variable val2 that measures the attendance to the needs was with a higher average, at 7.18, followed by val6, which measures the entertainment. In general, the means are close to 7, indicating agreement with the statements measured in the metric scale. The assertions used in the research instrument are presented in Appendix A. The variability remains low, as in the other evaluated constructs. Correlations are evaluated in Graph 22.

Graphic 22: Value correlations



Source: Research data, 2018.

Although the correlations do not exceed 0.8, as happened in the usability construct, there are four correlations that exceed 0.7. Two of them involve val1, which measures the sense of achievement that the course provides. Val1 maintains high correlations with val2 - Needs and with val5, which measures the sense of independence, the latter being the highest correlation among all this construct. Another variable that has strong correlations with other two variables is val6 - Entertainment, with val2 - Needs and val5 - Independence. The variables val3 and val4 are the ones that have the lowest correlations with the others. Finally, the dependent variables, which form the latent factors: satisfaction and continuance of use, were summarized together, both for the summary measures and for the evaluation of the correlations. Thus, Table 34 presents the results for these two constructs.

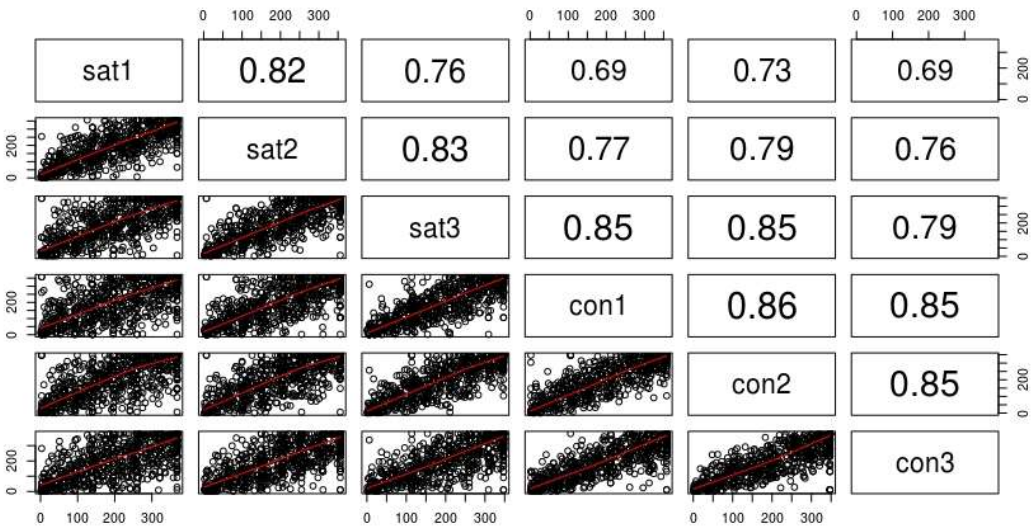
Table 34: Satisfaction and continuance summary

variable	min	median	max	mean	sd	CV
con1: I want do courses	0	7.8	10	7.19	2.5152	0.34
con2: I will continue do courses	1	8.0	10	7.32	2.5521	0.34
con3: I will do courses regularly	1	7.0	10	6.71	2.7643	0.41
sat1: Course performance	0	7.7	10	7.25	2.3030	0.31
sat2: Choose to take a MOOC	0	7.6	10	7.16	2.4039	0.33
sat3: Be involved in a MOOC	0	7.6	10	7.15	2.5286	0.35

Source: Research data, 2018.

According to the results, there were respondents in the sample who scored 0 on all questions, disagreeing with the affirmative. Nevertheless, the means follow the results already found in the previous constructs, with averages ranging from 6.71 to 7.32. The medians were slightly higher, ranging from 7.6 to 8.0. With the low variability, it is possible to infer that for the sample, a general agreement regarding the variables of satisfaction and continuance of use. The relationships among the variables complete this initial analysis of the constructs, being presented in Graphic 23.

Graphic 23: Satisfaction and continuance correlations



Source: Research data, 2018.

Among all the constructs evaluated, the dependent variables present the highest correlations among them, ranging from 0.69 to 0.86. These results may be of concern because of the imminent high impact of multicollinearity, which is an assumption that should not be violated in the conduct of CBSEM. Evaluating only the results of the correlations, it is plausible that the satisfaction variables nourish relations of great magnitude with the continuity of use variables.

In the next topic, the manifest variables are evaluated in exploratory ways, but with a deeper understanding of the use of variables in CBSEM. The use of variables is given in two approaches, the first by Item Response Theory (IRT), seeking to identify the items that best fit a factorial model and provide more information that explains the phenomenon. On the other

hand, the Classical Test Theory (CTT) comprises the Exploratory Factor Analysis EFA and the measurement and scale procedures.

5.2 Exploratory approaches on the measurement scale

According to the flowchart of the study, presented in the methodological procedures chapter. At this point of the analysis, two approaches are conflicted to select the variables of the study, both the variables that are present in the research instrument and those that will be analyzed in the CBSEM in later stage. Both approaches, however conflicting in the study, are used for the same common purpose: to develop a measurement scale. The methodological strategies reported in this topic and that rescues previous steps of the research, follow the steps for the construction of a measurement scale as suggested by Mackenzie, Podsakoff and Podsakoff (2011) and Maroco (2014).

Divided into seven steps, this research sought to define the constructs used (1); generated items that define the constructs (2); validated the content through the use of validated scales and indicator survey techniques (3), generated a measurement model through CFA (4), collected data to conduct a pre-test (5), refined the research instrument through of exploratory techniques by different approaches (6) and validated the measurement model and the structural model (7).

It is known that to measure the characteristics and behavior of individuals, we call a latent unobservable trait to represent them. This latent trait in turn reflects variables that can be measurable. The measurement scales aggregate these latent traits and can be formed by two approaches, CTT and IRT, as already mentioned in this chapter of analysis. This topic seeks to select the study variables from the two approaches. At first, the approaches are competing, but throughout the analysis the results of one approach may be similar or even complement the results of another approach. The contributions of these methods impact from obtaining the research instrument used, to the selection of variables that may present better fit in the structural model via CBSEM.

Nebojsa (2014) states that in empirical self-evaluation management studies it is almost exclusively based on a reflective model containing relationships between latent traits, analyzed through the CBSEM under the theoretical scope of CTT. This finding results in the scarce use of the IRT (Nunes & Goldszmidt, 2013). Kose and Demirtasli (2012) claim that the most important improvements of the latent century are IRT in psychological measurement. IRT is a

modern test theory which examines the ability level by using responses to test items with strong assumptions against CTT's weak assumptions with mathematical models.

CTT uses the score of a statistical test as a measurement reference. The IRT focuses on the item rather than the test. Both contemplate the analysis of items from the parameter estimates, but in TRI there is no dependence on the measure of proficiency in relation to the test and there is no dependence on the parameters of the items in relation to the set of respondents. For these two CTT limitations, Moghadamzadeh, Salehi and Khodaie (2011) respectively refer to symmetry and invariance. The factorial analysis of discrete or continuous items are conceived as latent variables. The model underlying the factorial analysis is equivalent to the TRI model with three parameters. So, what changes from one approach to another are parametrizations.

5.2.1 Item Response Theory approach

The database was imported into R-Studio, which uses the R programming language for statistical analysis purposes (R, 2018). The use of R to conduct the approach through IRT was due to the advancement of this approach in the R language and the scarcity of statistical tools that adopt this approach in the main statistical packages such as SPSS and SAS. The graphs were used as a visual resource to synthesize the information extracted from the data. The items and constructs were analyzed by the IRT approach, using factorial scores, item characteristic curves (ICC) and item information curves (IIC).

The use of the IRT approach is encouraged because it is a more flexible approach than the linear CFA, since it is not necessary to assume linearity between indicators and the required sample tends to be much smaller, guaranteeing a high reliability index with a small sample (Nunes & Goldszmidt, 2013).

5.2.1.1 Before data collection

Before even collecting the data that compose the sample of this research, in the bibliometric and systematic review stage, reported in topics of chapter two, pre-tests were performed to verify the behavior of the variables and possible adjustments that should be made. In addition to a descriptive analysis, observing mean, variability and asymmetry and kurtosis indexes, IRT was used in the same way that it was used in the complete sample.

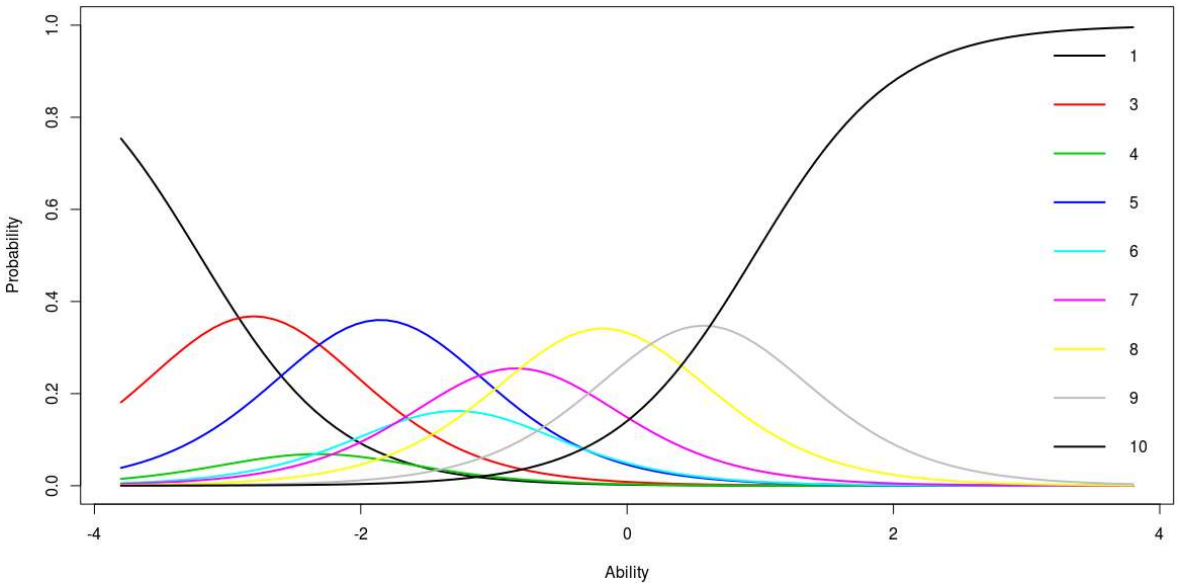
The objective of using the IRT still in the pre-test phase was given by the flexibility of the approach in working with small samples, and to verify the distribution of the points of the scale through the characteristic curve of the item (ICC). In this graphic, the points of the scale are

plotted by the ability level theta, being the ability one of the parameters measured in the IRT. Thus, is plotted a chart that measures along skill levels the probability hits for each point of the scale. In other words, the highest point on the scale is expected to be more likely to hit at higher skill levels. Thus, whoever scored this point has a high level of agreement on the measured variable. Proportionally, the other points are expected to have high probabilities in the skill range that gives that point.

This graphic allows us to detect anomalies and biases that would not normally be perceived by evaluating only the parameters of central tendency and variability. In fact, finding that a variable presents high dispersion does not inform the cause of the dispersion, nor whether this dispersion is caused by a bias. The ICC chart provides supplemental information that helps you calibrate the wording of the item. For example, if a question presents the majority of respondents marking the largest point of the scale, it is interesting to subjectively assess the wording of the question to see if it encourages or influences the respondent to score higher on the scale.

In the pretest test of this research, as reported in the chapter on methodological procedures, the IRT approach was used as complementary support through the ICC, giving clues and directing an analysis of the items more efficiently. The ICC was generated from the use of the latent trait model (ltm) in the R environment. The package designer defines the ICC as an indicator of the relative ability of an item, to discriminate among contiguous trait scores at various locations along the trait continuum (Rizopoulos, 2006). Exemplifying the graphic, each variable generated an ICC, such as the variable shown below, the con3, which evaluates the regularity in the use of MOOC courses.

Graphic 24: Item Characteristic Curve (ICC) for use regularity



Source: Research data, 2018.

Graphic 24 shows the points on the scale marked by the respondents, from 1 to 10, representing the lines. In the x axis is the parameter of ability, being a parameter of continuum, that is, with limits from - infinity to more infinity, but with greater proportion of skill concentrated in the range of -4 to 4. In the y axis is shown the probability of answer the point k of the scale on each skill level. There is also a third parameter not shown in the graphic, which measures the difficulty of the item, being a parameter similar to an index of variability.

The greater the difficulty of the item, the greater the dispersion. The coefficients of skill at each point of the scale are expected to follow the order of the points on the scale and that the points on the scale can be easily identified at the highest levels of probability. This information helped to adjust the wording of the questions, but besides the ICC, factorial scores per individual could be traced, identifying the individual's position in the ICC. This information can identify the pattern of responses of individuals who have a pattern of responses that deviate from the mean, and therefore, tend to have a lower discriminating power of the points of the scale.

In the ICC analysis it was also possible to identify the points of the scale where the information has low probability. Variables with high information rate may be being driven by a few points on the scale. Thus, variables that contain many points on the scale with a probability close to 0 indicate that the information provided by the variable is being inflated by a few points on the scale. In the analysis of the study variables, some variables presented this behavior, with 5 or more of the scale with probability close to 0, they were: col2, col3, qua1,

qua6, usa6, val3 and val4. These variables can cause measurement errors if they remain in the model.

In the next topic, the statistical procedures using the IRT approach are expanded, generating information underlying the EFA and CFA, as well as appropriating some CTT measures, such as the Bisserial Correlation Index and the Cronbach Alpha.

5.2.1.2 After data collection

After data collection, the treatment of missing data and univariate outliers was performed, the most appropriate IRT model was chosen to compose the data. The most suitable models for polytomic ordinal variables are the Graded Response Model (grm) and the Generalized Partial Credit Model (gpcm). With the help of the code of the caret package (Kuhn, 2017), the models were compared, and the one that got the best fit for the data was the grm. It is necessary to point out that a polytomic ordinal variable has the same syntactic effect of manifest ordinal variables. In the case of the data of this research, they are used as ratio variables in CTT, however, in fact they are ordinal qualitative variables.

While in the gpcm model, it is possible to have 1 to 3 parameters, using dichotomous variables, the grm model was idealized specifically for manifest ordinal variables. The first appearance of the model in the literature is authored by Samejima (1969) and has been used since then for variables that attend to the presupposition of having ordinal levels of answers in the scale of the items. The model is defined by the following equation:

$$\log\left(\frac{\gamma_{ik}}{1-\gamma_{ik}}\right) = \beta_i z - \beta_{ik} \qquad (7)$$

Where β_{ik} is the ratio of β_{ik*} to β_i. The term γ_{ik} denotes the cumulated probability of a response in a category k of item i, given a latent ability z. If constrained = True, then it is assumed that β_i is equal to β. The constrained refers to the equalization of the distance of the points of the scale. When true, arbitrarily the distance between the points of the scale is the same.

The ltm package allows to apply a variance analysis to see which model best fits the data, which arbitrates the same beta value or leaves it free. The function used is anova.grm. The fit of the model is based on the marginal approximation by maximum likelihood, using the Gauss-

Hermite quadrature rule, to calculate the approximation of the required integrals. In the case of grm, each point of the scale has an integral, which forms the area of the curve represented in the ICC.

As each item is independent in parameter definition, the general model was divided into models for each construct, evaluating for each construct whether the model with the free beta would be the most appropriate. For the variables of all constructs, the betas were not symmetrical, indicating that the model with free beta, that is, with false constrained is the most appropriate. For example, the result for the Collaborative Learning construct is shown in Table 35.

Table 35: Anova for grm model

Likelihood table					
Variables	AIC	BIC	log.lik	LRT	p.value
Int2 -> constrained = T	2326.20	2464.27	-1110.10		
Int -> constrained = F	2323.27	2474.37	-1103.64	12.93	0.024

Source: Research data, 2018.

According to Table 35, the comparison of grm models with free or equal beta for the points of the scale is performed from the AIC and BIC measurements. The AIC is the Akaike Information Criterion, which uses the maximum likelihood function of the model and the number of variables, through equation AIC = -2log (Lp) +2[(p+1) +1]. The BIC is the Bayesian Information Criterion, defined to BIC = -2log (Lp) +[(p+1) +1]log(n). These measurements are marked by the sum of the squares of the residuals (SQE), that is, AIC and BIC increase when the number of residues increases. The model with less residues and lower AIC and BIC values are preferable. The model presented is the one of interactivity, and this presented better fit in the model with free beta, being interpreted by the value of the likelihood ratio test (LRT) and its p-value smaller than 0.05. The other constructs presented similar results.

Continuing with the analysis, for each construct, Cronbach's Alpha was calculated, since each construct generated a grm model, and each model must have an adequate internal consistency. In addition, the amount of test or item information for the IRT model in the specific range of -4 to 4 was also measured. This interval is commonly used, concentrating most of the item information (Rizopoulos, 2006). The Table 36 shows the results of the Cronbach Alpha for each construct, making possible the conduction of grm models for each construct. In addition, the accumulated information values of each construct are presented.

Table 36: Internal consistency and amount of information of each grm model

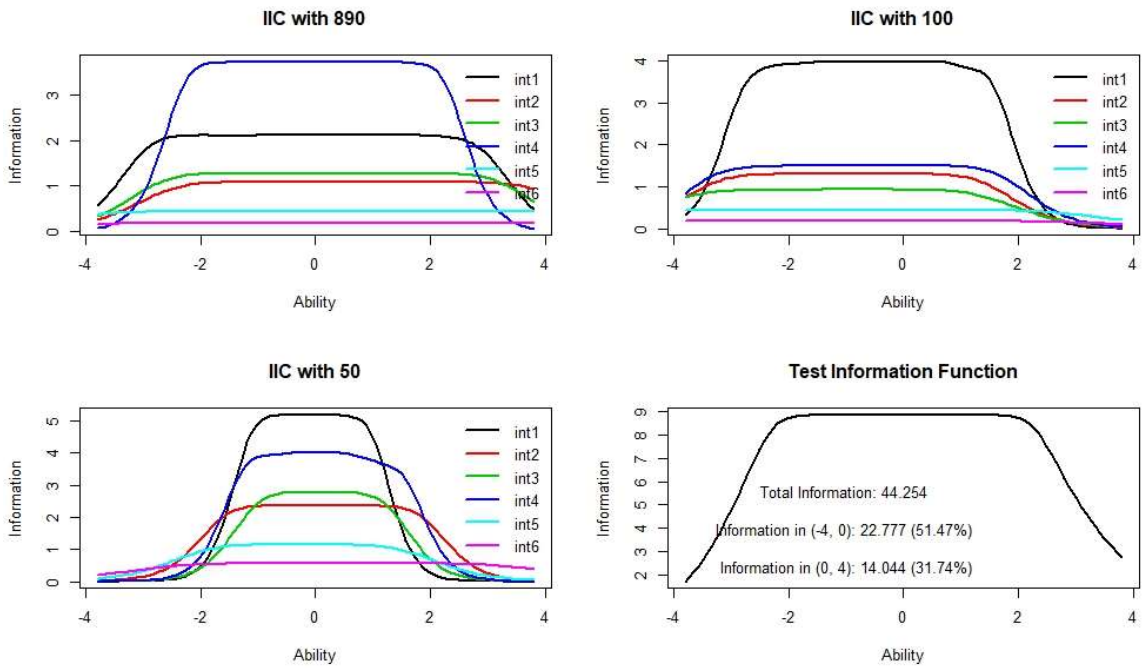
Reliability via Alpha de Cronbach					
Interactivity	Collaborative learning	Quality	Usability	Value	Satisfaction and continuance
0.822	0.837	0.866	0.897	0.890	0.898
Amount of information of each grm model					
Int_grm	Col_grm	Qua_grm	Usa_grm	Val_grm	Satcon_grm
36.81 (83,22%)	52.68 (87,70%)	56.85 (88.08%)	62.81 (88.39%)	78.24 (99.05%)	85.62 (99.45%)

Source: Research data, 2018.

According to Table 36, the Cronbach Alpha for all constructs was above 0.8, indicating adequate internal consistency (Hair et al., 2011; Maroco, 2014).Given that the models are feasible, we have the amount of information resulting from the sum of information of the variables of each construct. It is noticed that the grm model of the six dependent variables contains more information if compared with the other constructs. Although the ranking is an important result, it is notorious for the IIC graphics that all the models manage to gather a great amount of information.

As each model obtained was generated with the complete sample, alternative models were generated using sub-samples of 100 and 50 observations respectively. The objective of generating models with sub-samples is to verify if the obtained results could be generated with a smaller sample size. The results presented are the graphs of Item Information Curves (IIC). In addition to the IIC graphics for the original sample and sub-sample values, the Kernel Density Estimation (KDE) is also presented for the model's ability parameter, which is responsible for indicating the amount of information accumulated by the construct. The KDE curve uses the skill values and factorial scores to estimate the probability density function. The results for the interactivity model are presented in Graphic 25.

Graphic 25: Item Information Curve for grm interactivity model

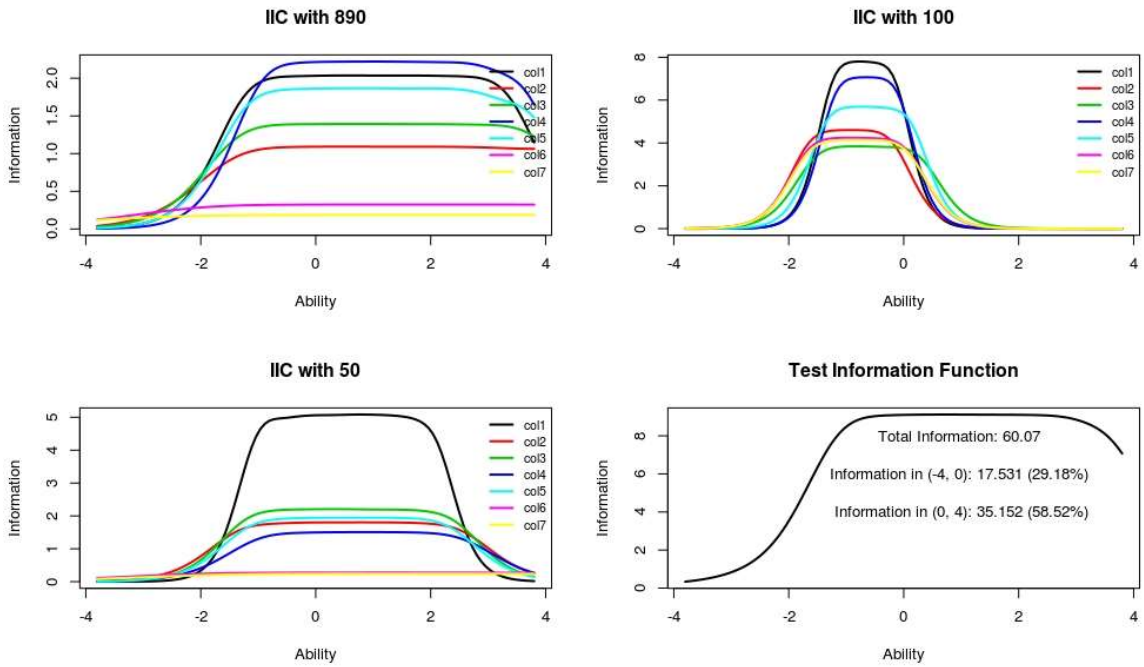


Source: Research data, 2018.

According to Graphic 25, the variable that contributes with more information to the model is int4. The variables with less information power are int5 and int6, precisely the same variables that have low correlations with the other variables of the model. The results for the sub-sample of 100 respondents revealed a different interpretation on the variable that most contributes with information, choosing int1. It was identical to the interpretation of the variables with less power of information, being int5 and int6. With a smaller sample number, the decision to remove variables from the model would be the same as with the original sample.

In fact, the variables mentioned above throughout the skill interval have low information and are strongly selected for exclusion. The KDE graphic shows the test information curve and is considered high along the analyzed skill interval. A final important result is verified by the amount of information in the negative range and the positive range of ability. It is noticed that the lower skill levels provide more information for the interactivity model. Then the results for the collaborative learning construct are presented.

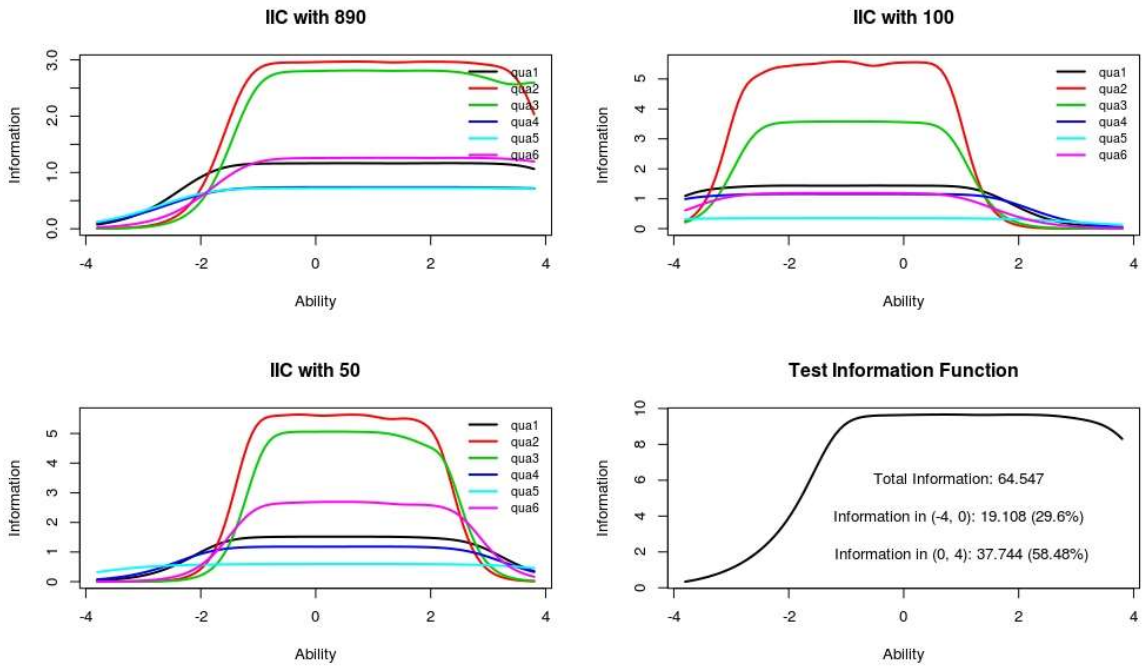
Graphic 26: Item Information Curve for grm collaborative learning model



Source: Research data, 2018.

The results for Graphic 26 show that the variables col1 and col4 provide more information for the IRT model. Nevertheless, col6 and col7 contribute less to the model, these being the variables that presented low correlations with the other variables of the construct. The models obtained with the sub-samples presented some divergent results. As a form of correction, the model was run with bootstrap, using 328 resamples, which is the minimum sample required to obtain a high statistical power, as presented in the chapter on methodological procedures. With bootstrap, the divergences of the sub-samples were corrected, obtaining an IIC similar to the original sample. Regarding the test information function, the highest skill levels are those that provide more information for the model. The following is the summary of the quality grm template.

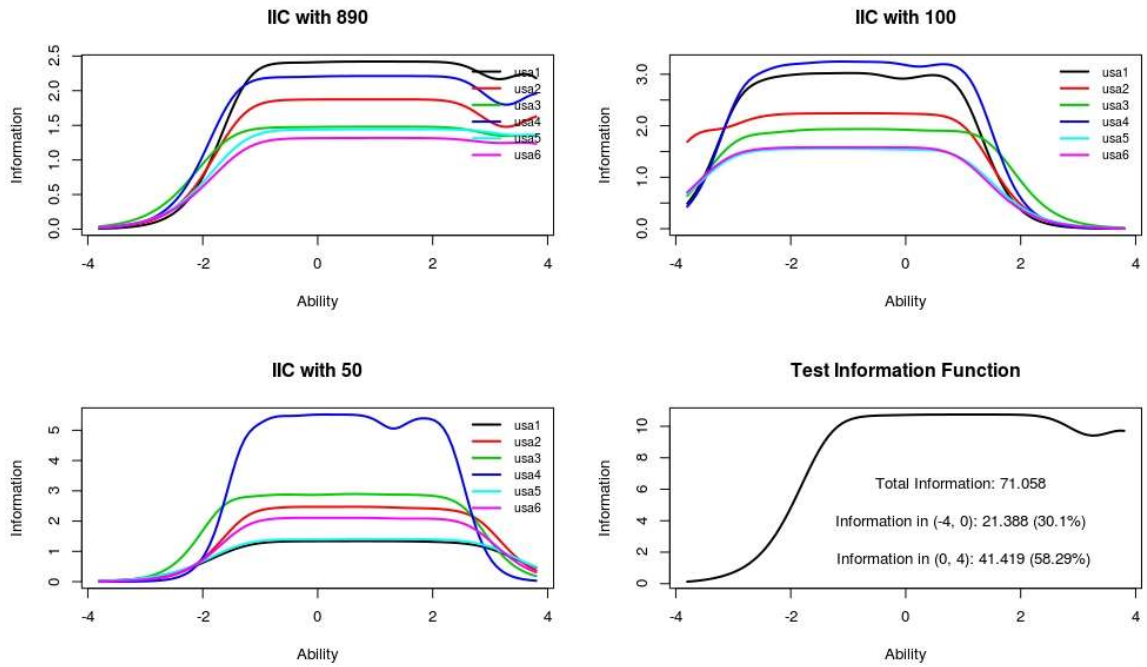
Graphic 27: Item Information Curve for grm quality model



Source: Research data, 2018.

The results of Graphic 27 show that the variable col5 is the one that least contributes with the model. This result is constant for the sub-samples. Without the use of the bootstrap, the same interpretation is also obtained for the items that provide this information, being qua2 and qua3. However, despite the high information levels of qua2 and qua3, these are the variables with the highest correlation with each other. The highest skill points are those that provide more information for the model. The usability model is presented to follow.

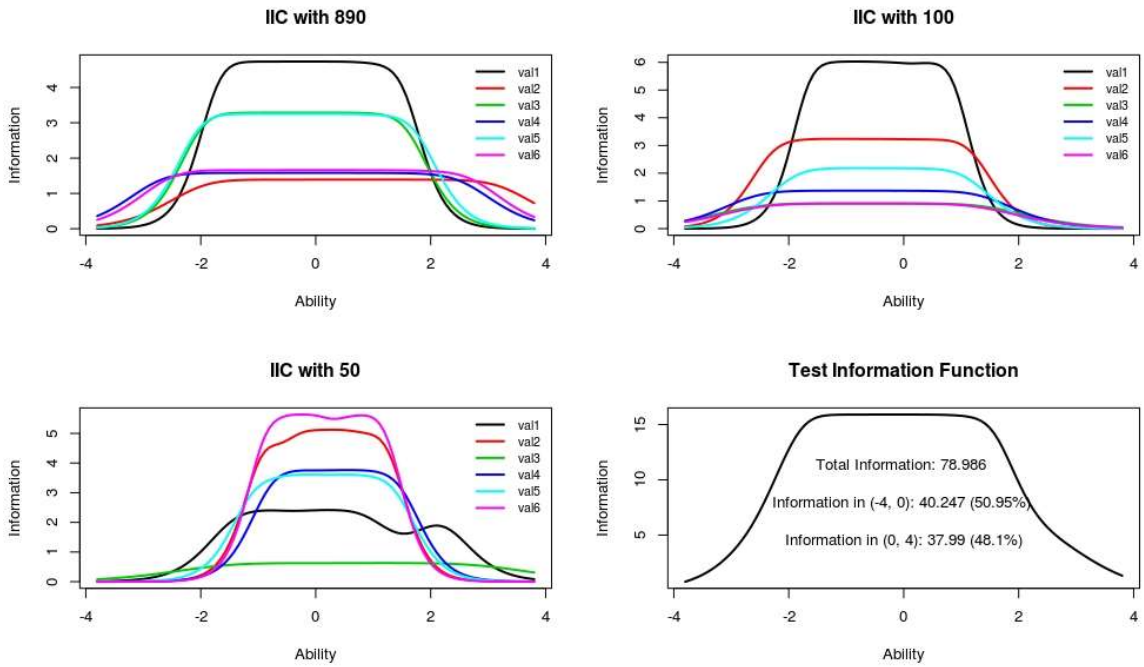
Graphic 28: Item Information Curve for grm usability model



Source: Research data, 2018.

According to Graphic 28, the usability variables showed high levels of information for all items of the model. The models of the sub-samples reversed the positions, but they kept presenting an amount of information far from 0 for each item. As with the quality model, high skill levels are those that provide more information for the model. The value grm model is the next to be summarized.

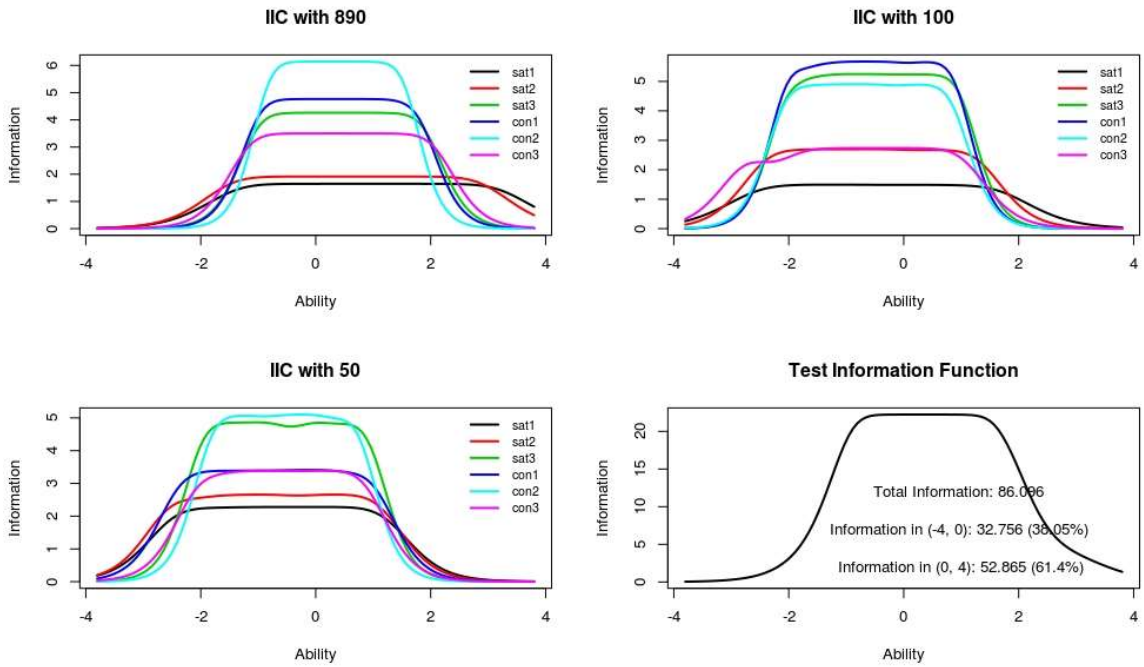
Graphic 29: Item Information Curve for grm value model



Source: Research data, 2018.

In the results of the value model, all the variables presented an amount of information relevant to the IRT model. For the sub-samples, it was necessary to adopt the bootstrap to generate the models, since the interpretation in the sub-samples differed from the results of the original sample. With the use of 328 resamples, the results of the sub-samples converged to the same of the original samples. Different from the models of quality and usability, there is an equilibrium in the contribution of information between levels of ability. In fact, the KDE curve indicates that the information is well concentrated in the range of -2 to 2. Finally, the results for the dependent variables model are presented in Graphic 30.

Graphic 30: Item Information Curve for grm satisfaction and continuance model



Source: Research data, 2018.

The results for the satisfaction and continuance model demonstrate high levels of information for all variables. According to the KDE curve, levels above 0 of skill contribute more information to the model. In general, variables int5, int6, col6, col7 and qua5 would be considered candidates for model exclusion. In spite of these results, the high information rate of some variables may have been enhanced by the high correlations and consequent presence of multicollinearity. In addition, biases can be statistically identified using CTT techniques. Evaluating the assumptions of CBSEM, the structural model with these variables removed was conducted in a later stage.

To validate the grm models, a model residual matrix is used. If there are no values below the main diagonal, it means that the model has items with good adjustment. When calculating the chi-square residues in each model, all presented good adjustment. This result is similar to the comparison of an observed and expected correction matrix. As an example, the fit for the interactivity model is shown in Table 37.

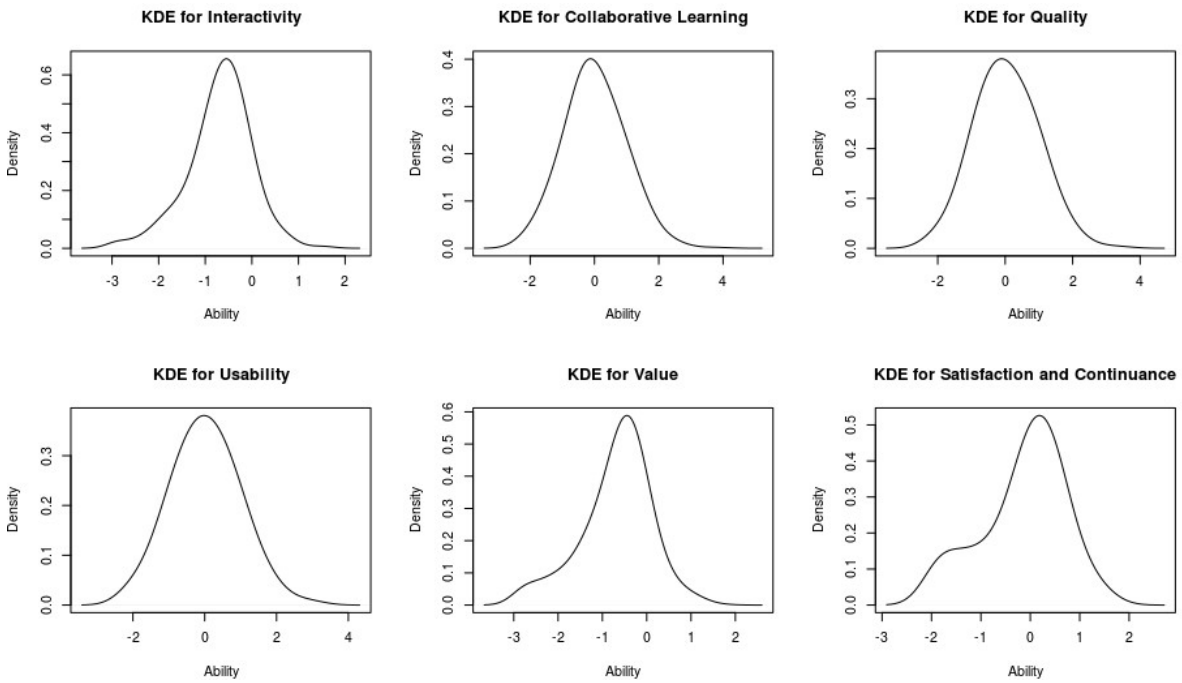
Table 37: Fit grm model

	Int1	Int2	Int3	Int4	Int5	Int6
Int1	-	86.37	79.42	73.18	100.63	97.55
Int2		-	79.84	66.55	67.47	56.62
Int3			-	66.15	114.62	75.03
Int4				-	133.68	80.72
Int5					-	130.29
Int6						-

Source: Research data, 2018.

As a way of summarizing the factorial scores of the validated models, Graph 31 presents the KDE obtained for each model. Those that match high density levels with the largest area of the curve were those that obtained higher factor loads per respondent.

Graphic 31: KDE for grm models



Source: Research data, 2018.

According to the results found in the KDE graphics, the quality of information per model already presented is confirmed by the areas of the curves in each graph. It is noticeable to notice that there is variation in the density curve along the levels of the skill parameter. Nunes and Goldszmidt (2013) explain that in order to obtain high levels of reliability throughout the latent trait scale, it is necessary to create items with varying difficulty parameters that discriminate both individuals with low levels of latent trait (analogous to " easy "or non-demanding in a test, with which subjects with a low level of the latent trait tend to agree or strongly agree) and high

levels of latent trait (analogous to" difficult "or very demanding questions in which only high level individuals of the latent trait tend to agree or strongly agree).

Additional analyzes of factorial scores per person were omitted because of the infeasibility of presenting such results to 890 participants. However, these measures are used in later moments, being compared to similar outputs of CTT. In the next topic, the exploration of the data is performed using CTT, through the EFA. In addition, the procedure for obtaining a scale, adopting the JAD strategy, is reported. Finally, the possible influence of common variance to the method (CMV) on the evaluated items is verified. The results of CTT tend to generate more obstacles for variables to be accepted in the model, which is not the same as being more robust, but rather to have statistical procedures with more premises and assumptions.

5.2.2 Classical Test Theory approach

The empirical cycle captures the process of coming up with hypotheses about how stuff works and testing these hypotheses against empirical data in a systematic and rigorous way. It characterizes the hypothetic-deductive approach to science. (Scholten, 2017b). Recalling this important approach to doing science, it is natural to ask the question: what is truly the question to ask or what hypotheses to be tested by the study? To assist in this process of obtaining hypotheses and operationalizing them through indicators, some metrics and indicator survey techniques are used. Among the techniques, there are those used before obtaining a research instrument and others that are used to calibrate or refine the indicators already used (Assis, 2015). This topic addresses the measurement scale formation techniques used in this study under the CTT approach.

5.2.2.1 JAD/NGT integration

Some steps prior to conducting a CBSEM analysis, such as literature review and the design of the research tool, are of extreme importance to avoid adjustment problems in structural models. In fact, validation problems in empirical models can be minimized in earlier steps. One of the ways of minimization is to conduct an unbiased and standardized literature review as possible. Another way is to obtain a search instrument based on previous research and previously validated models (Scholten, 2017b).

As reported in the chapter on bibliometric and systematic analysis, for the composition of this theoretical model, it was used research on e-learning and MOOCs. E-learning research is more advanced empirically, with more maturation time, many case studies, and validated

models. The constructs related to the second order factor "Performance", being quality, usability and value, as well as the variables dependent on satisfaction and continuance of use, have theoretical support of models already validated empirically. Thus, the items of each construct have already been tested in other studies, generating measures that allow selecting the ones that best fit in a model of this nature.

This consolidation of e-learning studies is not similar to studies on MOOCs. The interactivity and collaborative learning constructs, which measure the second-order fact "Moocs Adherence Index" (MAI), do not support literature with validated models, and usually have their items measured by unvalidated scales or by validated scales, but are one-dimensional.

As a strategy to obtain measurable indicators that could explain a latent factor, the Joint Application Development (JAD) data collection technique was used, being a facilitated group technique, which can be used to systematize information requirements, through the synergy of participants in the theme in question. JAD is used in management as a qualitative technique of data collection, similar to the focus group in qualitative research. The JAD is superior to the interview technique, for promoting interaction and for reducing the time of data collection. However, there is a low standardization in conducting a facilitated session. Together with JAD, the Nominal Group Technique (NGT) is incorporated into the steps of JAD. NGT in a nutshell consists of a series of standardized procedures used in group sessions (Duggan, & Tachenkary, 2004).

In a workshop or qualitative strategy outside the axis of the interviews, the objective is different. The main difference is the number of participants, usually higher, but can also be differentiated by the guidelines and products expected in the data collection. The main divergent characteristic between focus groups and interviews is the final product and the achievement of a consensus that represents the group of interviewees.

The facilitated sessions have the common goal of identifying information requirements or indicators for a particular issue or area of interest (Assis, 2015). JAD can reduce communication barriers and identify needs that improve the quality of a system or process (Duggan, & Tachenkary, 2004). However, JAD groups face problems with free interaction, this lack of control is also one of the defects of a brainstorm. NGT is used in conjunction with JAD to minimize these problems, with a well-defined structure and rules, seeking equal participation and resolution of common conflicts in JAD sessions.

The focus group, even though it is a well-defined collection strategy in the literature (Creswell, 2012), seeks to obtain a qualitative analysis product based on tacit knowledge, so the analyzed data are not quantifiable in the traditional sense of the word, or measurable. The

use of a qualitative strategy with the focal group collection method using NGT and JAD facilitator is similar to the difference between qualitative and quantitative research strategies. Creswell (2012) states that a qualitative research seeks to detect experiences and perceptions related to the research topic and the way in which they make sense of these experiences. Qualitative research also focuses on the process that is occurring and, on the product, or result of that process. In quantitative research, the objective is to synthesize analytically an already perceived and credible reality, but with unknown standards.

In this research, the theme presents the standards gap, the absence of studies with quantitative approaches or even the absence of attempts to quantify, relate variables or even identify them, aiming at the standards to be controlled and managed. Thus, this collection strategy seeks to take the first step in the search for a group of measurable indicators that represent the latent factors that explain the rate of adherence to the MOOCs by the users.

NGT is a result of the techniques applied by the New England Regional Leadership Program, concluding that people working alone but in a group, environment express a greater number of ideas. Thus, the NGT is a collective process with well-defined stages and that seeks to maintain the autonomy of the interviewee, with the minimum influence of the other participants in certain stages, which were popularized by Duggan and Tachenkary (2004). The suggested steps for NGT are in accordance with Assis (2015, p.65) suggested and were conducted in this research following the following steps.

• Opening of the session - The workshop participants, being six professional participants and researchers, three from the area of education and three from the management area, were divided into trios, of which two will act as leaders, being a facilitator (researcher) and another documentary (assistant researcher). These leaders were chosen by the members of each group in a subjective evaluation of the member's qualifications in the theme. Each participant presented in a few words and the facilitator described the group's process of action, besides presenting the objective of generating items that measure the users' adherence to the MOOCs;

• Initial registration of requirements - Each participant registered the information requirements that he / she deemed appropriate to the process under analysis; the orientation given by the facilitator was to think of measurable indicators, so the information requirements were thought as indicators of user adherence to the MOOCs;

• Presentation of requirements or initial questions - Each member of the group presented, without discussing, one of their indicators, which was recorded by the documenter; then the facilitator requested that each person submit a new requirement or question, and so on, until all

items were recorded. In each group, each member suggested five items, obtaining altogether thirty items at the end of this stage;

• Review of requirements or questions - the facilitator read to the group each requirement or question recorded and questioned the existence of doubts, interpretations or explanations on the part of the participants, eliminating possible overlaps; during this process each requirement or question was numbered and the list of them was printed by the documenter and distributed to the participants of the group;

• List of relevant requirements or issues - the facilitator requested that each participant record the numbers of indicators judged relevant in the list of requirements or issues reviewed. The ranking of each participant should have ten items numbered from 1 to 10 in order of importance; the documenter tabulated the results of each participant;

• Assignment of importance - with a single list of indicators, the facilitator requested that participants assign each indicator a level of importance, from zero to ten, which was recorded by the documenter;

• Ranking of requirements or critical questions - the degree of importance assigned by the group to each indicator is calculated as the average of the grades assigned to it, weighted by the number of indications received by them. The final outcome of the process was a list of thirteen indicators. These indicators were compared with current theory, seeking theoretical support and allocation of indicators in latent factors worked on the studies on the subject.

5.2.2.2 Exploratory Factor Analysis (EFA)

The EFA was used in this research as one of the exploratory procedures with the objective of refining the research model, knowing a priori the number of latent factors to consider. Maroco (2014) explains that the use of EFA in confirmatory models is usually used as a specification search, which seeks to find a subset of variables that present an optimal combination of simplicity (parsimony) and quality of adjustment, but without looking for latent factors. One of the main advantages of using EFA as an exploratory step is that the EFA can be driven by different methods of factor extraction, while the AFC commonly uses maximum likelihood estimation.

The EFA is part of the family of multivariate techniques by the CTT approach. Hair et al. (2011) conceptualizes multivariate analysis as an analysis of multiple relationships in a single set or model. The EFA is conceptualized as an interdependence and exploratory technique. According to Fávero et al. (2009) the EFA allows the study of interrelationships between many

variables, with the aim of synthesizing or condensing all the information into groups of variables, without any considerable loss of information.

In the data treatment step, the incidence of missing values per Missing Value Analysis (MVA) was verified in SPSS v. 023. As the survey was conducted online, with all the mandatory issues, no missing data was identified. In addition, extreme values were verified using the interquartile deviation method (IQR = Q3 - Q1), with the difference between quartile 3 and quartile 1. The extreme extremes were determined by Q1 - 1.5 * IQR, while values high extremes calculated by Q3 + 1.5 * IQR. They were identified in 14 manifest variables, totalizing 1.2% of the sample.

As the outliers do not exceed 3% of the sample, Maroco (2014) suggests as a remedy to arbitrate the standardized value of 3 deviations in a module, minimizing the effects of extreme values and thus avoiding measurement errors. Thus, the data were standardized by the normal distribution and then recoded in the same variables, leading to extreme cases of -3 or 3. At the end of the corrective measures, the assumptions of the EFA were verified, such as: sample size (1), homoscedasticity (2), linearity (3), multivariate normality (4) and multicollinearity (5).

Regarding the sample size, Hair et al. (2011) suggests that there are five observations per manifest variable. 31 manifest variables were used in the analysis with 890 observations, requiring 155 observations. Regarding homoscedasticity, the Levene's test was conducted to verify the homogeneity of the variances. Homoscedasticity for all variables was violated, indicating that the data are heteroscedastic.

As the Levene's test is quite robust, an alternative strategy for verifying homoscedasticity is by looking at the residue graphs with the predicted values. The graphs for each variable indicated that they did not have any behavior or tendency, generating indications that the variances of the residues were not unequal. To confirm the subjective analysis of the graphics, the Breusch-Pagan test was calculated, being a similar test to the Levene's test, but more suitable for large samples. With the conduction of the test, the homogeneity of variances was confirmed.

The third assumption is that of linearity, which is verified through the indexes of correlations between the manifested variables. The correlation matrix formed among the 31 manifest variables showed linear relationships in most pairs of variables. Some variables had weak relationships, with correlation indexes below 0.4. However, were included in the model so that the measures generated by the EFA were evaluated.

The multivariate normality was tested through the asymmetry and kurtosis indexes, as was used in the t test for the dependent variables in the previous topic. The variables presented

values between -3 and 3 in both indexes, confirming the attendance to this assumption. Finally, the presence of multicollinearity indicates measurement errors related to indexes of inflated correlations. To verify the absence of multicollinearity, the Variance Inflation Factor (VIF) test was used. It was identified that the variable used6 of the usability construct exceeded the maximum value of 5 of the VIF thesis, being then a candidate to leave the factorial model. The other variables obtained values well below 5, indicating the absence of multicollinearity.

In the EFA, the global adjustment indices were verified at each round of analysis, being the sphericity test of Bartlett and the Keiser Meyer Olkin (KMO). The first tests the overall significance of all correlations, verifying if the correlations between the manifest variables are significantly different from 0. The second index informs if the sample size is adequate against the model's complexity. The ideal values suggested by Hair et al. (2011) and Maroco (2014) are: significance close to 0 for the Bartlett test and values close to 1 for the KMO.

The EFA conducted used the most common extraction method, that is principal components from the correlation matrix. Varimax orthogonal rotation was used in order to obtain a factorial structure in which a variable is strongly associated with only one factor or dimension. At first, a factorial without rotation was generated, only to investigate the scree-plot or sedimentation graphic, which identifies the points between the numbers of components or dimensions with eigenvalues. It was noticed that six dimensions emerge with eigenvalue above 1. As the objective is not to find an ideal number of latent factors, the factorial solution is arbitrated in five factors, as it is predicted in the theoretical model of the research. This process to fix the number of factors configure the specification search. The factorial solution obtained a variance rate of 69%, with significance in the Bartlett test = 0 and KMO = 0.911.

The same factorial model was tested using the maximum likelihood method by Varimax orthogonal rotation. The maximum likelihood method is considered more robust than the main components method. The total explained variance of the model increased to 70%, maintained the sig. Bartlett = 0 and KMO = 0.921, being these overall adjustment values considered very good. As correlations above 0.4 between the latent factors were verified, the orthogonal solution is not the most adequate, since it arbitrates correlations equal to 0 between the factors, to generate better interpretable factorial solutions. Thus, the maximum likelihood model was rotated obliquely, controlled by the Kappa index = 4. The Kappa index, when it is equal to 1, is analogous to orthogonal rotation, but when it is equal to 4, it leaves the correlations free. The new model obtained a rate of variance = 71%, with significance of Bartlett = 0 and KMO = 0.923. This was the best fit model.

To refine the model, the EFA generates measures of conformity of the variable to the factorial model. The variables that best fit the model are those that do not violate the minimum limit of measures: anti-image, commonality, factor loads and cross-load. Hair et al. (2011) defines the reference values and conceptualizes each measure:

- Anti-image: is the degree to which the factors explain each other in the results, generated from the matrix of partial correlations (must be greater than 0.5).

- Commonality: is the total amount of variance that one variable shares with all others included in the analysis. It is the most important measure and its minimum limit must be followed without exception. If a variable exhibits commonality below 0.5, it means that the shared variance it nourishes is less than the chance variance shared.

- Factorial load: it is the correlation between the original variables and the factors. The factorial score is the measure used to characterize a latent factor. (Must be greater than 0.4).

- Cross-load: it is a phenomenon that occurs when a variable presents high factorial loads in more of a latent factor, preventing the delimitation of the variable to a single factor chance.

After evaluating these measures, four of them failed on one or more premises. Each variable was excluded in a round, since the withdrawal of a variable implies changes in the measures evaluated (Hair et al. 2011). It should be noted that the orthogonal models, by principal components and by maximum likelihood did not indicate the output of any variable. Only with the third model, with oblique rotation, by maximum likelihood, some variables disapproved in the measures analyzed. The results that reprove the variables are presented in Table 38.

Table 38: Variables excluded of factorial model

Qt.	Cod.	Variable	Criteria	Values
1	Col6	Interest of share	Commonality	0.374
2	Col7	Knowledge shared	Commonality	0.400
3	Int5	Doubts and deepening	Commonality	0.449
4	Int6	Interaction outside the platform	Commonality	0.450
5	Qua5	Content quality	Cross-load	High loads in four factors

Source: Research data, 2018.

It can be seen from Table 38 that the candidate variables leaving the model, converge with the factorial and IIC scores of the IRT approach. The variables col6, col7, int5 and int6 failed

in the commonality criterion, being less than 0.5. The analogous measure in IRT is the amount of information provided by the variable for the construct, visualized in the IIC. For the variable qua5, even obtaining adequate commonality, presented high factorial loads in four factors. This result indicates the low retention of the variable to the corresponding latent factor, in this case, the quality. After finding the desired factorial model, excluding the five variables that failed in the criteria, the anti-image, commonality and factor load measurements are presented in Table 39.

Table 39: Measures of EFA

Cod.	Variable	Anti-image	Factorial load	Commonality
Int1	Hypermedia	0.935	0.698	0.698
Int2	Reading forums and chats	0.961	0.705	0.635
Int3	Writing forums and chats	0.954	0.680	0.584
Int4	Stimulate participation	0.933	0.702	0.698
Col1	Correction of activities (Learning)	0.941	0.797	0.665
Col2	Correction of activities (Empathy)	0.941	0.688	0.573
Col3	Freedom to propose new forms to study	0.928	0.689	0.654
Col4	Freedom to indicate materials	0.930	0.739	0.679
Col5	Teamwork	0.948	0.746	0.629
Qua1	Layout	0.913	0.756	0.687
Qua2	Navigation	0.924	0.788	0.726
Qua3	Comfort	0.966	0.805	0.786
Qua4	Completeness	0.958	0.719	0.592
Qua6	Support readiness	0.916	0.658	0.705
Usa1	Easiness	0.935	0.658	0.823
Usa2	Skill	0.923	0.879	0.736
Usa3	Features and activities	0.951	0.840	0.753
Usa4	Tools	0.946	0.833	0.821
Usa5	Content transition	0.937	0.845	0.813
Usa6	Support productivity	0.909	0.727	0.754
Val1	Achievement	0.910	0.847	0.852
Val2	Needs	0.903	0.810	0.822
Val3	Follow trend	0.908	0.697	0.636
Val4	Intelligence sense	0.920	0.731	0.613
Val5	Indepedence	0.855	0.699	0.867
Val6	Entertainment	0.894	0.818	0.801

Source: Research data, 2018.

In the next topic, the manifest variables are tested as they were predicted in the theoretical model, in order to verify the effect of the method and thus to identify another element that can cause measurement error.

5.2.2.3 Common Method Variance

It is widely accepted among theorists and researchers that relationships among variables measured with the same method can be distorted and inflated because of the Common Method Variance (CMV). Many theorists, in fact, have pointed out that monomethod biases are responsible for a certain amount of variance in measurement. This study discusses the relevance of CMV in applied research and describes the method used to understand the effect of common method biases at the scale of the study.

The problems with self-assessment questionnaires, in the case of this study, appear due to the questioned theme, which can be understood in different ways and can change over time and in different environments. In this study, there are different environments, since the sample can be composed of people from anywhere in the world who have access to the Internet. It is also true that the research presents a cross-section with respect to time, that is, it is measuring a picture of reality at that moment, even though the measurement scale contains disconfirmatory constructs. Faced with these limitations, it is natural that parameter estimates may be contaminated, causing errors in measurement and reporting of results (Nebojsa, 2014). Thus, the errors of measurement of the manifest items of this research were evaluated and remedied.

Method Variance refers to variance that is attributable to the measurement method rather than the construct of interest (Bagozzi & Yi, 1991). By method, Spector (2006) conceptualizes as the different ways of measuring something, in different levels of abstraction, either by the formats of the items, the writing of the items, the procedures of data collection, the different formats of the research instruments, the different sources of data, among others. Spector (2006) points out that some methods can be controlled, but not all, causing measurement errors.

The CMV is explained by Podsakoff et al (2003) that in the same paper offer solutions for method biases. At first, the authors argue that method biases vary from area to area. In behavioral sciences, there are different kinds of biases than evidenced in deterministic areas.

Another relevant point is the possible influence of method variance on the relationships between variables, including correlations and consequent factor loads. In other words, two constructs when correlated and measuring the same latent factor, must have a correlation equal to 1. When the actual correlations are compared the observed correlations, the difference

between them generates a measurement error. This error can inflate or deflate correlation indices.

Some common errors in behavioral self-evaluation research are portrayed and identified by various authors, such as social desirability (Thomas, & Kilmann, 1975), implicit theories (Phillips, & Lord, 1982) and negative affectivity (Chen, & Spector, 1991). In relation to negative affectivity, it refers to the tendency to experience a wide range of negative emotional states and represents the potential source of bias that produces CMV. Individuals with a negative view of the self tend to suffer from poor self-esteem, report stress and physical symptoms, and experience strains and dissatisfaction across time and situations (Chen, & Spector, 1991). Despite the authors' approach, they did not find negative affectivity effects in the variables investigated. Later, Spector (2006) confirmed the influence of negative affectivity on some behavioral variables.

Regarding social desirability, he refers to the need for social approval and acceptance and the belief that it can be attained by means of culturally acceptable and appropriated behavior (Thomas, & Kilmann, 1975). According to Podsakoff et al. (2003) individuals tend to hide their personal effects on the topic they are asked to answer in order to fulfill a social desire. It is well established that social desirability represents the potential source of bias in measurement process. The study of Thomas and Kilmann (1975) has the aim of highlighting the relevance of the social desirability variable in organizational research by focusing on two potential biasing effects in measuring conflict-handling modes: elevation of means and misleading correlations.

On the implicit theories, Phillips and Lord (1982) argue that raters classify others grounded on their most noticeable traits or behaviors and interpret these classifications into behavioral expectations or the behavioral scores requested by questionnaires; and that these classifications may introduce systematic bias into behavioral rating. After identifying three of the most recurrent CMV sources that lead to measurement errors, in the articles analyzed, the authors emphasize the influence of negative affectivity, social desirability, and implicit theories. Other method effects by common source can be consistency motif, leniency biases, positive affectivity, acquiescence and transient mood state. These studies converge on three points on the common method bias.

• The influence of such sources in fact generates biases that can influence the estimates and interpretation of the constructs worked.

• The biases, including CMV, should be investigated in order to have knowledge of the presence of noises and to facilitate understanding of the constructs studied.

• It is necessary when working with complex constructs and with a high degree of abstraction, the identification, understanding and actions to minimize biases found.

In a recent research, Lai, Li and Leung (2013), CMV is considered a serious threat to the validity of the findings in a survey and that there are not enough studies about its possible influences. On the application in organizational studies, the authors claim that the confounding influence of CMV has attracted a great deal of attention in organizational research (Lai, Li & Leung, 2013).

Lai. Li and Leung (2013) present six remedies to control CMV, and more a remedy is proposed using Monte Carlo simulation. In the study by Podsakoff et al (2003), other sources for common method biases are discussed, as the characteristics of the item and the item context. Podsakoff (2003) concluded that from the presence of multiple sources, there are more effective remedies for each source.

It is possible to conclude that CMV is a real problem, present in studies that using psychometric and behavioral constructs in general. To generate and interpret the parameter estimates, the researcher should investigate the impact of potential biases and take actions to reduce this impact, with the awareness that the biases cannot be fully controlled (Spector, 2006). From the investigation of the common method biases, is attributed to CMV a more complex design, however, enhancing the identification and use of measures to minimize their effects, advancing in the literature to address issues that contribute to the validity of the estimates in empirical researches.

After these considerations about CMV, the measurement errors were identified, and a statistical remedy was chosen to minimize the effect. Podsakoff et al. (2003) points out that there are two types of remedies: procedural and statistical. Procedures refer to good practices in data collection. From the methodology of Podsakoff et al. (2003), the best statistical method of identification of method variance was chosen.

At first, the Harman's single factor test was chosen to verify the magnitude of the effect of CMV on the measurement scale. The test identifies whether there is a factor that emerges in the EFA with more than 50% of the variance of the model. If this is confirmed, it is reasonable to conclude that CMV is a major problem. However, it is a method that does not control the effects of the method, but only allows evaluation of the extent of CMV. With the conduction of the test, the construct with the highest variance rate did not approach 50%, reaching 18%. A more efficient method is to isolate a general scoring factor. Although ignoring the measurement error, this method can accurately reflect CMV and identify items that undergo CMV influence.

From a set of remedies, four decision rules help you choose the best statistical medicine. The rules of decision were as follows;

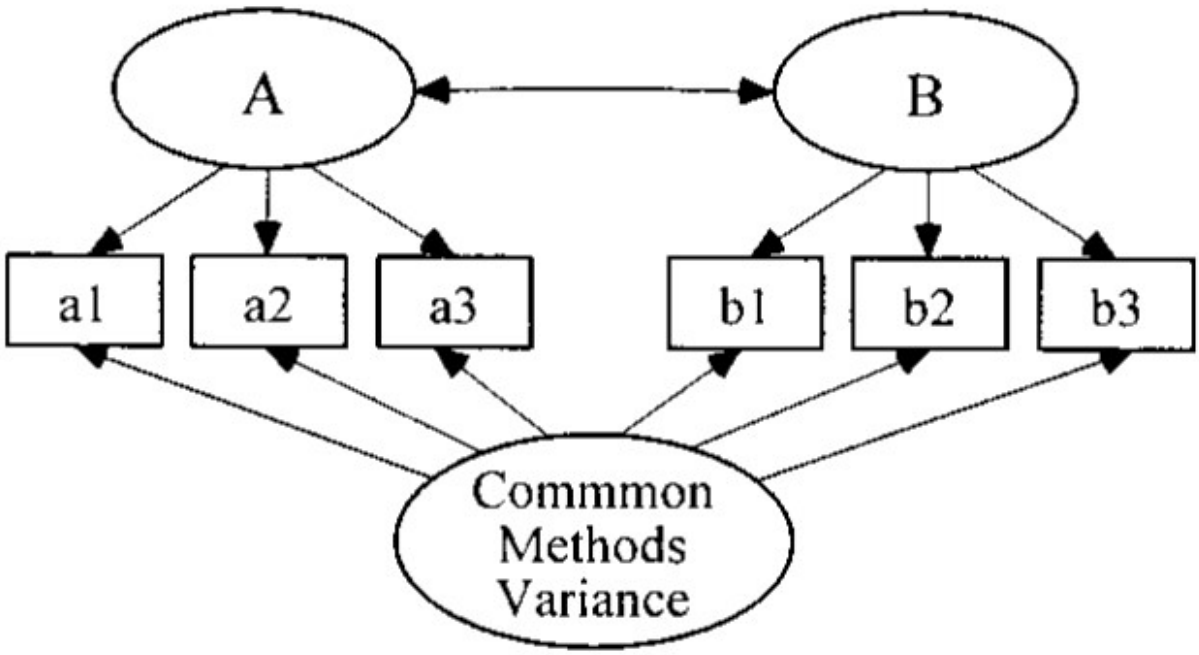
Can the predictor and criterion variables be obtained from different sources? R. No

Can the predictor and criterion variable be measured in different contexts? R. Yes

Can the source (s) of the method bias be identified? No

In this case, the fourth decision rule does not apply. In this research, a single collection route was used, but aggregating different contexts by the scope of the study. Regarding the source of the biases, it cannot be identified. To be able to identify, it would be necessary to have prior knowledge of the sample or to include a scale of the desired method effect, such as the scale of social desirability inserted through research instruments to identify the effect of social desirability. The statistical remedy is represented by the following path diagram.

Figure 10: Common Latent Factor approach



Source: Podsakoff et al. (2003).

Thus, Podsakoff et al. (2003) from these answers suggests to include a single factor going against manifest variables. This single factor is also called the Common Latent Factor (CLF), and this statistical remedy is a category of the single-method-factor approaches. The decision rule to identify high influence of CMV is given by the difference of the standardized weight of the variable to the latent factor with the same measure with the CLF included in the model.

Being greater than 0.4 in module indicates that there is a high influence of CMV, being deflated if the difference is negative and inflated if the difference is positive. The covariance values between the error and the item also reveal the direction and magnitude of the measurement error. The use of the common factor occurred for each group of constructs: performance and MAI. The results for the performance are presented to follow.

Table 40: Common latent effect on performance variables

Standardized Regression Weights: with CLF				Standardized Regression Weights: without CLF				
Relations			Estimate	Relations			Estimate	Delta
qua6	<---	QUA	0,201	qua6	<---	QUA	0,693	0,492
qua5	<---	QUA	0,754	qua5	<---	QUA	0,802	0,048
qua4	<---	QUA	0,544	qua4	<---	QUA	0,679	0,135
qua3	<---	QUA	0,705	qua3	<---	QUA	0,871	0,166
qua2	<---	QUA	0,708	qua2	<---	QUA	0,788	0,080
qua1	<---	QUA	0,236	qua1	<---	QUA	0,698	0,462
usa6	<---	USA	0,402	usa6	<---	USA	0,677	0,275
usa5	<---	USA	0,752	usa5	<---	USA	0,778	0,026
usa4	<---	USA	0,734	usa4	<---	USA	0,826	0,092
usa3	<---	USA	0,669	usa3	<---	USA	0,757	0,088
usa2	<---	USA	0,591	usa2	<---	USA	0,73	0,139
usa1	<---	USA	0,706	usa1	<---	USA	0,845	0,139
val6	<---	VAL	0,754	val6	<---	VAL	0,762	0,008
val5	<---	VAL	0,741	val5	<---	VAL	0,798	0,057
val4	<---	VAL	0,219	val4	<---	VAL	0,65	0,431
val3	<---	VAL	0,214	val3	<---	VAL	0,63	0,416
val2	<---	VAL	0,625	val2	<---	VAL	0,846	0,221
val1	<---	VAL	0,614	val1	<---	VAL	0,883	0,269
qua1	<---	CLF	0,712					
qua2	<---	CLF	0,304					
qua3	<---	CLF	0,201					
qua4	<---	CLF	0,304					
qua5	<---	CLF	0,431					
qua6	<---	CLF	0,757					
usa1	<---	CLF	0,474					
usa2	<---	CLF	0,421					
usa3	<---	CLF	0,321					
usa4	<---	CLF	0,351					
usa5	<---	CLF	0,181					
usa6	<---	CLF	0,423					
val1	<---	CLF	0,441					
val2	<---	CLF	0,238					
val3	<---	CLF	0,625					
val4	<---	CLF	0,637					

val5	<---	CLF	0,311
val6	<---	CLF	0,327

Source: Research data, 2018.

According to the results, in the left columns, the model was generated using the CLF, as suggested by Podsakoff et al. (2003). In the right columns, the model was generated without the CLF. The delta difference between the standardized weights of the variables without CLF and CLF were calculated. For the variables qua1, qua6, val2 and val3, the values exceeded the reference limit. Thus, the effect of the method on these variables is evidenced, causing measurement errors, even if the source of the error is unknown. These results are in agreement with the results found in the IRT approach in the ICC analysis. In the IRT analysis, the high information of the four problem variables was generated by few points of the scale, while several points of the scale generated information close to 0. The same analysis was performed for the variables of the MAI. The results can be seen in Table 41.

Table 41: Common latent effects on MAI variables

Standardized Regression Weights: with CLF				Standardized Regression Weights: without CLF				
Relations			Estimate	Relations			Estimate	Delta
INT6	<---	INT	0.578	INT6	<---	INT	0.635	0.057
INT5	<---	INT	0.452	INT5	<---	INT	0.476	0.024
INT4	<---	INT	0.511	INT4	<---	INT	0.818	0.307
INT3	<---	INT	0.467	INT3	<---	INT	0.721	0.254
INT2	<---	INT	0.458	INT2	<---	INT	0.735	0.277
INT1	<---	INT	0.495	INT1	<---	INT	0.656	0.161
COL7	<---	COL	0.358	COL7	<---	COL	0.788	0.430
COL6	<---	COL	0.298	COL6	<---	COL	0.727	0.429
COL5	<---	COL	0.547	COL5	<---	COL	0.757	0.210
COL4	<---	COL	0.558	COL4	<---	COL	0.797	0.239
COL3	<---	COL	0.341	COL3	<---	COL	0.752	0.411
COL2	<---	COL	0.226	COL2	<---	COL	0.696	0.470
COL1	<---	COL	0.599	COL1	<---	COL	0.725	0.126
INT1	<---	CLF	0.577					
INT2	<---	CLF	0.647					
INT3	<---	CLF	0.587					
INT4	<---	CLF	0.598					
INT5	<---	CLF	0.202					
INT6	<---	CLF	0.231					
COL1	<---	CLF	0.669					
COL2	<---	CLF	0.625					
COL3	<---	CLF	0.641					

COL4	<---	CLF	0.679
COL5	<---	CLF	0.647
COL6	<---	CLF	0.298

Source: Research data, 2018.

The results for the interactivity and collaborative learning variables, INT and COL respectively presented only problematic deltas in the variables, col2, col3, col6 and col7. These variables had already presented low information in the IIC, led by the IRT approach. The method effects for these variables reinforce the problematic character of these variables. In the next topic, the CFA is conducted as the initial stage of the CBSEM.

5.3 Measurement model

The present topic addresses the first step of a modeling of structural equations by covariance, the CBSEM, generating as product the measurement model or confirmatory factorial model. In this step, a confirmatory factorial analysis, the CFA, is conducted. The modeling follows its second and final step using the validated factorial model, adding the causal relations and testing the complete theoretical model.

SEM is a multivariate technique that considers multiple linear relationships between exogenous and endogenous variables simultaneously. SEM is undoubtedly the most used method to validate theoretical models (Kline, 2005). SEM contributes to bring theories to new levels of understanding and understanding of human perceptions (Babin & Svensson, 2013). In order to have a structural equation it is necessary that the causal relations be represented by structural relations and that these structural relations can be equated and modeled under the theory of study (Nebojsa, 2014). It is in this way that SEM is understood to start in the process of reviewing the literature, making a theoretical model, making a research instrument, testing and retesting, validating the measurement scale, conducting the SEM and validating the theoretical model (Babin & Svensson, 2013).

In the academic management community, SEM is widely used, under the covariance-based method, the CBSEM and the method based on partial least squares, the PLS-SEM. The choice of CBSEM is based on prior knowledge of the theory, just as MOOCs have inherited the literature on e-learning. PLS-SEM is best suited for exploratory approaches or replications of validated theoretical models (Nebojsa, 2014).

In relation to the measurement model, this can be reflexive, formative or mixed. The model

of this study is reflexive, establishing relations of the construct to the indicator. The reflexive model is more suitable for this research for several reasons. In the first place, the theoretical framework of this research is based on psychometric constructs, that is, that measures attitude, personality and behavior in general; the latent trait is empirically defined, that is, it is subject to common variance based on measurement error or construct definition error; requires internal consistency to validate latent traits, and may remove a variable from the model, without this influencing the use of the construct.

Thus, the CFA is conducted under the aegis of hypotheses that explain relations between constructs, the constructs being a property, an abstract term that is operationalized by variables, being in turn, the variable a concrete expression of the construct, being measurable and manipulable (Scholten, 2017b).

In previous topics, the data were synthesized and explored. The measurement scale of the theoretical model was able to refine the model from two approaches: the IRT and the CTT. At the end of the exploratory phase, some variables were excluded from the model because they presented threats, specifically for the potential to cause type I and type II errors. Table 42 organizes the various stages of analysis by approach, indicating at which stage of the analysis a determined variable was indicated to exit the model.

Table 42: Excluded variables of theoretical model

Variable	Stage per approach					
	IRT		CTT			
	ICC	IIC	EFA	VIF	CMV	CFA
Int5		x	x			
Int6		x	x			
Col2	x				x	
Col3	x				x	
Col6		x	x			
Col7		x	x			
Qua1	x				x	
Qua5		x				x
Qua6	x				x	
Usa6	x			x		
Val3	x				x	
Val4	x				x	
Diagnosis	The variables excluded in each approach (IRT and CTT) converge, and integrate the same structural model					

Source: Research data, 2018.

According to Table 42, in the IRT approach, the characteristic curves of the item and the information curve of the item allowed to select several variables and identify problem variables, from the parameters of discrimination and factorial scores. With the same objective, procedures

that used CTT also indicated the exit of variables and selected those with greater potential in the theoretical model. Initially, the two approaches would generate competing models, but the refinement decisions for each approach converged, leaving only one factorial model to be conducted, without the mentioned variables.

The variable qua5 (content quality) was the only variable to leave the model, which had already removed the variables that had failed in the previous stages. Each variable was evaluated for factorial load and commonality, and qua5 presented adequate indexes. In the IRT phase, qua5 was the variable that contributed least to the quality grm model. In spite of obtaining a satisfactory factorial load in the EFA, and not having failed in the other tests, in the last stage of refinement it was disapproved summarily. Next, the assumptions of CBSEM are evaluated, and are considered in both steps: CFA and full structural model.

5.3.1Assumptions of CBSEM

Several assumptions evaluated in the use of CBSEM have already been presented in this research. However, the list of assumptions is wide and should be carefully scrutinized to avoid biased results in parameter estimates.

1. Multivariate normality: this assumption was verified in the EFA stage, through the skewness and kurtosis indexes. According to Maroco (2014) and Téo (2010). These third and fourth order measures respectively have the ability to determine deviations from multivariate normality. Table 43 presents these measures for the study variables, indicating that the normality assumption was attended.

Table 43: Estimates of skewness and kurtosis

Variable	skew	c.r.	Kurtosis	c.r.
val6	-0.678	-8.260	-0.182	-1.111
val5	-0.596	-7.255	-0.422	-2.572
val2	-0.686	-8.349	0.058	0.354
val1	-0.612	-7.455	-0.339	-2.066
usa5	-0.763	-9.290	0.304	1.853
usa4	-0.643	-7.828	0.064	0.387
usa3	-0.585	-7.123	-0.115	-0.702
usa2	-0.570	-6.945	-0.402	-2.446
usa1	-0.621	-7.565	-0.118	-0.718
qua4	-0.450	-5.485	-0.384	-2.337
qua3	-0.662	-8.057	-0.056	-0.338
qua2	-0.653	-7.951	0.138	0.841
col5	-0.654	-7.960	-0.048	-0.294
col4	-0.838	-10.203	0.175	1.067
col1	-0.516	-6.283	-0.422	-2.571

Variable	skew	c.r.	Kurtosis	c.r.
int4	-0.373	-4.549	-0.589	-3.589
int3	-0.691	-8.416	-0.190	-1.156
int2	-0.710	-8.644	0.293	1.783
int1	-0.422	-5.135	-0.538	-3.279
Multivariate			95.015	50.171

Source: Research data, 2018.

2. Linearity: The use of CBSEM requires linearity in the model, so that all parameters can be estimated. The linearity of the model was confirmed in the stage of EFA and doubly confirmed in the conduction of CBSEM in AMOS graphics v. 21, which when rotating the model, automatically confirms the linearity of the model. In the same way, the model was tested in the R, with the use of the lavaan package (Rosseel, 2012) and with the model being generated, the linearity is also confirmed.

3. Non-zero sample covariates: this assumption is met in the modeling phase of the factorial model. At this stage, latent factors are operationalized by manifest variables, and it is required that these variables present some type of association, that is, that their covariance is not zero. It is an analogous verification to the Bartlett test in the EFA, with the difference that it does not use correlations, but rather covariance (Maroco, 2014).

4. Multiple indicators: A consensus in the literature is the estimation of latent variables by at least three manifest variables (HAIR et al. 2011). Combined the use of three indicators per construct, the use of large sample (greater than 300) avoids the production of Heywood cases. The Heywood cases refer to a factorial solution that produces an estimate of error variance below zero (negative). A result like this is impossible, but it can be obtained if there are small samples (below 300) or constructs measured by less than three indicators (Hair et al. 2011).

5. Multicollinearity: this assumption was also verified in the EFA stage. Multicollinearity occurs when variables are highly correlated, so certain statistical functions may not work correctly (Teo, 2010). The VIF test was conducted and the results are presented in Table 44, indicating that only the variable usa6 violated the assumption, being removed from the model still in the exploratory stage.

Table 44: VIF test for multicollinearity diagnosis

Model		Statistics	
		Tolerance	VIF
	Hypermedia	0.289	3.457
	Reading forums and chats	0.349	2.863
	Writing forums and chats	0.417	2.395
	Stimulate participation	0.299	3.341
	Doubts and deepening	0.510	1.962
	Interaction outside the platform	0.514	1.947
	Correction of activities (Learning)	0.322	3.105
	Correction of activities (Empathy)	0.412	2.426
	Freedom to propose new forms to study	0.359	2.785
	Freedom to indicate materials	0.339	2.946
	Teamwork	0.403	2.483
	Interest of share	0.580	1.725
	Knowledge shared	0.582	1.719
	Layout	0.321	3.114
	Navigation	0.289	3.460
	Comfort	0.241	4.147
	Completeness	0.408	2.450
	Content quality	0.393	2.542
	Support readiness	0.310	3.221
	Easiness	0.194	3.157
	Skill	0.277	3.605
	Features and activities	0.264	3.783
	Tools	0.189	4.295
	Content transition	0.199	3.014
	Support productivity	0.262	5.824
	Achievement	0.165	3.043
	Needs	0.187	3.345
	Follow trend	0.366	2.730
	Intelligence sense	0.401	2.495
	Indepedence	0.146	3.852
	Entertainment	0.211	3.731

Source: Research data, 2018.

6. Strong measure: this assumption is met still in the development of the research instrument, requiring the measurement of the manifest variables by a scale above five points, which generate discrete or continuous variables. The scale used in this research was a metric scale of semantic differential and amplitude from 0 to 10, thus meeting this assumption.

7. Inexistence of outliers: to meet this assumption, univariate outliers were still identified

in the data treatment phase. However, this assumption also encompasses multivariate outliers. In general, the square distance of Mahalanobis (D^2) is used as a distance measure that can identify non-standard observations. As the outliers may inflate or deflate the covariance between the variables, this measure should be used to exclude the observations with higher D^2 (Hair et al., 2011). There is no reference value for D^2, but exclusions cannot leave the sample smaller than the minimum required. Thus, all the assumptions were evaluated and met. In the next step, the adjustment quality indexes of the model are presented.

5.3.2 Model adjustment

When conducting the factorial model for the first time, the model can be adjusted at each step using D^2 of Mahalanobis to exclude extreme observations and add a correlation between variable errors. The first method of model re-specification is due to the need to remove multivariate outliers from the sample. The second, by the possibility of strong correlations between manifest variables that were not contemplated in the theoretical model.

In relation to D^2, 100 observations were excluded, being 20 in each round, in a total of 6 rounds. In each round, a re-specification was performed adding correlations between variable errors, being in total 5 re-specifications. Such re-specification is identified by the modification index. Silva (2006) explains that for each parameter specified, there is a value that represents the expected fall in the value of the general chi-square if the parameter is no longer fixed. Thus, each re-specification made in the path diagram neutralizes the effect of high correlation.

The correlations between errors occurred between items of the disconfirmatory constructs of the performance and the dependent variables. These high correlations are justified in the literature, mainly for the relations between value and satisfaction constructs (Chiu et al. 2005; Pereira et al. 2015).

Once adjustments were made in the factorial model, the adjustment quality indexes improved with each round of re-specification. These indexes measure the overall fit quality of the model. The indexes are divided by family, and each family has several indexes that can represent it. The choice of one index over the other is due to the recurrent use in the literature. Maroco (2014), Kline (2005) and suggest those indexes that are more common. The indexes chosen by family were:

1. Absolute indexes: Evaluate the quality of the model by itself, without comparison with any other model. Normally the X^2 / gl indexes are used, the chi-square being divided by the

degrees of freedom, or the RMR, being the root mean square residual, is the calculation of the root of the residues on the degrees of freedom. Maroco (2014) affirms that these indexes are not important, since nothing is good or bad, but yes, comparable or not. Anyway, the measure chosen was the RMR. The smaller and closer to 0 is the RMR, the better.

2. Relative indexes: Evaluate the quality of the model by testing the model with the worst possible adjustment and / or with the best possible adjustment (CFI - Comparative Adjustment Index, TLI - Tucker Lewis Index, NFI - Normal Adjustment Index). CFI is the most used of relative indexes. Increasing variables in small samples tends to reduce both CFI and TLI. The NFI is the higher in the same proportion of the number of model variables and the sample size (Maroco, 2014). The adjustment is good if it is above 0.8.

3. Indexes of parsimony: Parsimony indices are obtained by correcting the relative indices with a penalty factor associated with the complexity of the model (Maroco, 2014). (PCFI - CFI parsimonious; PNFI - NFI parsimonious). Mulaik et al. (1989) sets values lower than or equal to 0.6 as a bad fit, between 0.6 and 0.8 are good and above 0.8 has optimal adjustment.

4. Population discrepancy indexes: Compare the adjustment of the obtained model with means and variances of the sample to evaluate if the model is approximately correct (RMSEA - Root of the Average Approximate Quadratic Error). It has a good fit between 0.05 and 0.1 and a very good adjustment at an index lower than 0.05 calculated in a 90% confidence interval.

5. Indexes based on information theory: Based on the X^2 statistic, it penalizes the model in function of its complexity. (ECVI - Expected Cross-Validation Index; MECVI - Adjusted ECVI Index). They have no reference values. The ECVI reflects the theoretical adjustment of the model in other samples similar to the one in which the model was adjusted, from a single sample (Kline, 2005).

Thus, the indexes were calculated for the adjusted measurement model and are presented in Table 45. The measures calculated with their respective reference values demonstrate an adequate overall adjustment quality. The next topic presents the validation of the measurement model.

Table 45: Global adjustment for measurement model

Index	Results	Diagnosis
RMR	0.258	Smaller, the better
TLI	0.899	>0.80 – Very good fit
CFI	0.922	> 0.80 – Very good fit
NFI	0.914	> 0.80 – Very good fit
PCFI	0.712	> 0.80 – Very good fit
PNFI	0.705	] 0.6; 0.8 [- Good fit
RMSEA	0.096	] 0.05; 0.10 [- Good fit
ECVI	1.485	Smaller, the better
MECVI	1.488	Smaller, the better

Source: Research data, 2018.

5.3.3 Validation of measurement model

After estimating the model by the maximum likelihood method, it was showed the factorial scores of the measurement model, represented by the non-standardized weights of the parameter estimates. Using non-standardized results, it is possible to rank the variables by the level of contribution and the importance of the variable in the model. There is also the same of standard error (S.E), which measures the estimated error, reflecting the precision with which each parameter was estimated.

The C.R is the critical ratio, consisting of the ratio between the parameter estimates by the errors. This test follows the normal distribution, at a 99% two-tailed confidence level. The symbols *** indicate that the p-value of C.R was very close to 0, and consequently rejects the null hypothesis, showing that the estimates are different from 0. The results of the measurement model are shown in Table 46.

Table 46: Estimates of measurement model

			Estimate	S.E.	C.R.	P	Standardized
int1	<---	INT	1.000				0.820
int2	<---	INT	0.678	0.029	23.208	***	0.714
int3	<---	INT	0.919	0.038	24.112	***	0.739
int4	<---	INT	1.064	0.036	29.246	***	0.855
col1	<---	COL	1.000				0.821
col4	<---	COL	0.846	0.036	23.604	***	0.764
col5	<---	COL	0.897	0.037	24.561	***	0.793
qua2	<---	QUA	1.000				0.733

			Estimate	S.E.	C.R.	P	Standardized
qua3	<---	QUA	1.235	0.047	26.281	***	0.882
qua4	<---	QUA	1.090	0.052	21.002	***	0.712
usa1	<---	USA	1.000				0.839
usa2	<---	USA	0.899	0.038	23.464	***	0.687
usa3	<---	USA	0.889	0.035	25.244	***	0.721
usa4	<---	USA	0.955	0.032	29.655	***	0.804
usa5	<---	USA	0.876	0.025	34.992	***	0.745
val1	<---	VAL	1.000				0.876
val2	<---	VAL	0.885	0.025	35.389	***	0.864
val5	<---	VAL	1.032	0.028	36.350	***	0.874
val6	<---	VAL	0.934	0.034	27.775	***	0.846

Source: Research data, 2018.

The validity of the adjusted model passes through three components: the factorial, convergent and discriminant validity. Together they form the validity of the construct. The first component is the factorial validity, which is verified through the standardized weights of the variables. A considerable and acceptable weight according to Kline (2005) and Teo (2010) is above 0.5. From the results of Table 46, the 19 variables present in the model have a factorial load above this value.

With the final model adjusted, the Composite Reliability (CR) and Average Variance Extracted (AVE) indexes were calculated, in order to demonstrate the compliance with convergent and discriminant validity. The convergent validity is met when the constructs under study present positive correlations between theoretically parallel constructs and depend on the proof of the factorial validity (HAIR et al. 2011).

Fornell and Larcker (1981) proposed to evaluate the convergent validity by calculating the CR and AVE indexes. AMOS does not have these indexes in the output file. Therefore, the AVE can be calculated as follows. Where $\Sigma\lambda$ is the factorial weight and $\Sigma\varepsilon\lambda$ is the error associated with the factorial weight. Calculations were made for the 4 antecedent confidence constructs. $\sum\lambda$ is the factorial load and $\sum\varepsilon\lambda$ is the error associated with the factorial weight. Calculations were made for the five latent factors.

$$AVE = (\sum \lambda^2) / (\sum \lambda^2) + (\sum \varepsilon \lambda^2) \qquad (8)$$

The CR is also not present among the results generated by AMOS v. 21, so it is calculated manually, just like the AVE. The reliability of a construct refers to the consistency and reproducibility property of the measure (MAROCO, 2010). Hair et al. (2011) explains that the use of CR is analogous to the calculation of Crombach's Alpha in EFA. However, for the AFC,

the CR is more appropriate to contribute to the discriminant validation of a measurement model. The CR can be calculated as follows.

$$CR = (\sum \lambda)^2 / (\sum \lambda)^2 + (\sum \varepsilon \lambda) \qquad (9)$$

Anderson and Gerbin (1988) and Fornell and Larker (1981) propose to guarantee convergent validity that we must obtain values of AVE> 0.5 and CR> 0.7. To guarantee discriminant validity, Fornell and Larcker (1981) define the AVEs of the dimensions as greater than or equal to the square of the correlations between the factors. If AVE> (corr C1xC2) ² then the discriminant validity is confirmed. If confirmed, together with the CR results, the discriminant validity meets the required conditions. The following table compares the required values.

Table 47: Comparison of AVE with square of correlations

	AVE	INT	COL	QUA	USA	VAL	**CR**
INT	**0,615**	**0,615**					**0.83**
COL	**0,629**	0,5184	**0,629**				**0.84**
QUA	**0,607**	0,4624	0,4624	**0,607**			**0.87**
USA	**0,609**	0,5329	0,4225	0,6084	**0,609**		**0.90**
VAL	**0,748**	0,6084	0,3136	0,5041	0,6084	**0,748**	**0.89**

Source: Research data, 2018.

It is observed in Table 47 that convergent and discriminant validities by the method of Fornell and Larcker (1981) are met. With the three components of validity met, it is possible to affirm that the factorial model is reliable, consistent and reproducible.

5.4 Full structural model

After the validation of the measurement model, the second stage of the CBSEM is started, where the causal relations of the theoretical model are included. With the adjustments made until then, a structural model is conducted, with two second order factors, being performance and MAI, using 790 observations and 5 re-specifications, which correspond to correlations between errors.

The relationship between MAI and satisfaction was tested in two ways, having a direct causal relationship and a performance-mediated relationship. The mediated relationship did not present significant differences in the standardized weight of the relationship between MAI and

satisfaction, evidencing that there is no performance measurement for this causal relationship. In this step, the adjustment quality indices are calculated again, since the path diagram was changed by the causal relations. The results of the overall fit of the structural model can be seen in Table 48.

Table 48: Global adjustment for full structural model

Global adjustment	**Results**	**Diagnosis**
RMR	0.255	Smaller, the better
TLI	0.908	>0.80 – Very good fit
CFI	0.924	> 0.80 – Very good fit
NFI	0.914	> 0.80 – Very good fit
PCFI	0.758	> 0.80 – Very good fit
PNFI	0.750	] 0.6; 0.8 [- Good fit
RMSEA	0.087	] 0.05; 0.10 [- Good fit
ECVI	2.321	Smaller, the better
MECVI	2.326	Smaller, the better
RNFI	0.921	Very good fit
RPR	0.902	Very good fit

Source: Research data, 2018.

The global adjustment indexes are once again adequate, and therefore, the relationships that undermine the model are discarded. With the adjusted model, the estimates are validated by means of Z test, in the same way as done in the CFA, using the CR and the p-value, considering significant the relations between the parameters < 0.05, and the significance of the relevant trajectories when > 0.05. The standardized weights are also shown in Table 49.

Table 49: Results of structural model

			Estimate	S.E.	C.R.	P	Standardized
SAT	<---	P	1.303	0.148	8.799	***	0.542
SAT	<---	MAI	0.893	0.147	6.072	***	0.372
VAL	<---	P	2.118	0.068	30.966	***	0.969
USA	<---	P	1.620	0.061	26.717	***	0.904
QUA	<---	P	1.098	0.057	19.291	***	0.764
INT	<---	MAI	1.837	0.073	25.156	***	0.935
COL	<---	MAI	1.226	0.063	19.419	***	0.742
CON	<---	SAT	0.926	0.020	46.810	***	0.955
INT1	<---	INT	1.000				0.819
INT2	<---	INT	0.673	0.029	23.049	***	0.709
INT3	<---	INT	0.905	0.038	23.887	***	0.731
INT4	<---	INT	1.067	0.036	29.315	***	0.857

			Estimate	S.E.	C.R.	P	Standardized
COL1	<---	COL	1.000				0.800
COL4	<---	COL	0.878	0.038	23.113	***	0.767
COL5	<---	COL	0.929	0.039	23.973	***	0.794
QUA2	<---	QUA	1.000				0.732
QUA3	<---	QUA	1.226	0.047	25.926	***	0.875
QUA4	<---	QUA	1.095	0.052	20.981	***	0.712
USA1	<---	USA	1.000				0.837
USA2	<---	USA	0.912	0.038	24.290	***	0.694
USA3	<---	USA	0.899	0.035	25.371	***	0.725
USA4	<---	USA	0.955	0.032	29.560	***	0.802
USA5	<---	USA	0.882	0.026	34.310	***	0.756
VAL1	<---	VAL	1.000				0.882
VAL2	<---	VAL	0.873	0.025	35.591	***	0.858
VAL5	<---	VAL	1.015	0.027	37.296	***	0.878
VAL6	<---	VAL	0.928	0.034	27.586	***	0.847
SAT3	<---	SAT	1.000				0.948
SAT2	<---	SAT	0.881	0.019	45.607	***	0.883
SAT1	<---	SAT	0.765	0.022	35.139	***	0.800
CON1	<---	CON	1.000				0.927
CON2	<---	CON	1.038	0.019	53.914	***	0.940
CON3	<---	CON	1.060	0.023	45.657	***	0.896

Source: Research data, 2018.

According to Table 49, the C.R indicates that all measures are significant, with the possibility to rank them from the non-standardized weights (estimates). With the adjusted structural model, it is possible to insert the moderator engagement variable, and compare the solutions of the high and low engagement groups. The procedure for the inclusion of moderating variables is multi-group analysis followed by confirmation of the moderating effect in each causal relationship.

According to Maroco (2014), the multi-group analysis aims to assess whether the structure of the structural model is equivalent or invariant in different groups. With the support of the means comparison test, it is already possible to know in advance that the averages for high and low engagement are significantly different. It remains to be seen whether there are differences in the parameter estimates throughout the groups. It is also known previously that the other profile variables are invariant throughout the groups. The standardized results of the two groups are presented in Table 50.

Table 50: Standardized values by group of engagement

Standardized Regression Weights: (Low)				Standardized Regression Weights: (High)				
			Estimate				Estimate	Delta
SAT	<---	P	0.582	SAT	<---	P	0.499	-0.083
SAT	<---	MAI	0.236	SAT	<---	MAI	0.285	0.049
VAL	<---	P	1.006	VAL	<---	P	0.955	-0.051
USA	<---	P	0.793	USA	<---	P	0.8	0.007
QUA	<---	P	0.531	QUA	<---	P	0.647	0.116
INT	<---	MAI	0.979	INT	<---	MAI	0.891	-0.088
COL	<---	MAI	0.448	COL	<---	MAI	0.371	-0.077
CON	<---	SAT	0.903	CON	<---	SAT	0.895	-0.008
INT1	<---	INT	0.748	INT1	<---	INT	0.747	-0.001
INT2	<---	INT	0.508	INT2	<---	INT	0.577	0.069
INT3	<---	INT	0.66	INT3	<---	INT	0.518	-0.142
INT4	<---	INT	0.834	INT4	<---	INT	0.779	-0.055
COL1	<---	COL	0.689	COL1	<---	COL	0.616	-0.073
COL4	<---	COL	0.689	COL4	<---	COL	0.589	-0.1
COL5	<---	COL	0.732	COL5	<---	COL	0.736	0.004
QUA2	<---	QUA	0.753	QUA2	<---	QUA	0.654	-0.099
QUA3	<---	QUA	0.854	QUA3	<---	QUA	0.736	-0.118
QUA4	<---	QUA	0.694	QUA4	<---	QUA	0.684	-0.01
USA1	<---	USA	0.809	USA1	<---	USA	0.732	-0.077
USA2	<---	USA	0.615	USA2	<---	USA	0.574	-0.041
USA3	<---	USA	0.662	USA3	<---	USA	0.541	-0.121
USA4	<---	USA	0.741	USA4	<---	USA	0.693	-0.048
USA5	<---	USA	0.646	USA5	<---	USA	0.608	-0.038
VAL1	<---	VAL	0.821	VAL1	<---	VAL	0.796	-0.025
VAL2	<---	VAL	0.816	VAL2	<---	VAL	0.791	-0.025
VAL5	<---	VAL	0.782	VAL5	<---	VAL	0.781	-0.001
VAL6	<---	VAL	0.755	VAL6	<---	VAL	0.829	0.074
SAT3	<---	SAT	0.91	SAT3	<---	SAT	0.88	-0.03
SAT2	<---	SAT	0.773	SAT2	<---	SAT	0.672	-0.101
SAT1	<---	SAT	0.576	SAT1	<---	SAT	0.557	-0.019
CON1	<---	CON	0.882	CON1	<---	CON	0.837	-0.045
CON2	<---	CON	0.909	CON2	<---	CON	0.84	-0.069
CON3	<---	CON	0.824	CON3	<---	CON	0.803	-0.021

Source: Research data, 2018.

According to Table 50, the differences between the standardized weights vary from 0.01 to 0.142. The negative values indicate that the largest weight is that of the Low group. The groups tend to give importance to each variable in a similar way, even though the means are different. Both models showed good overall adjustment.

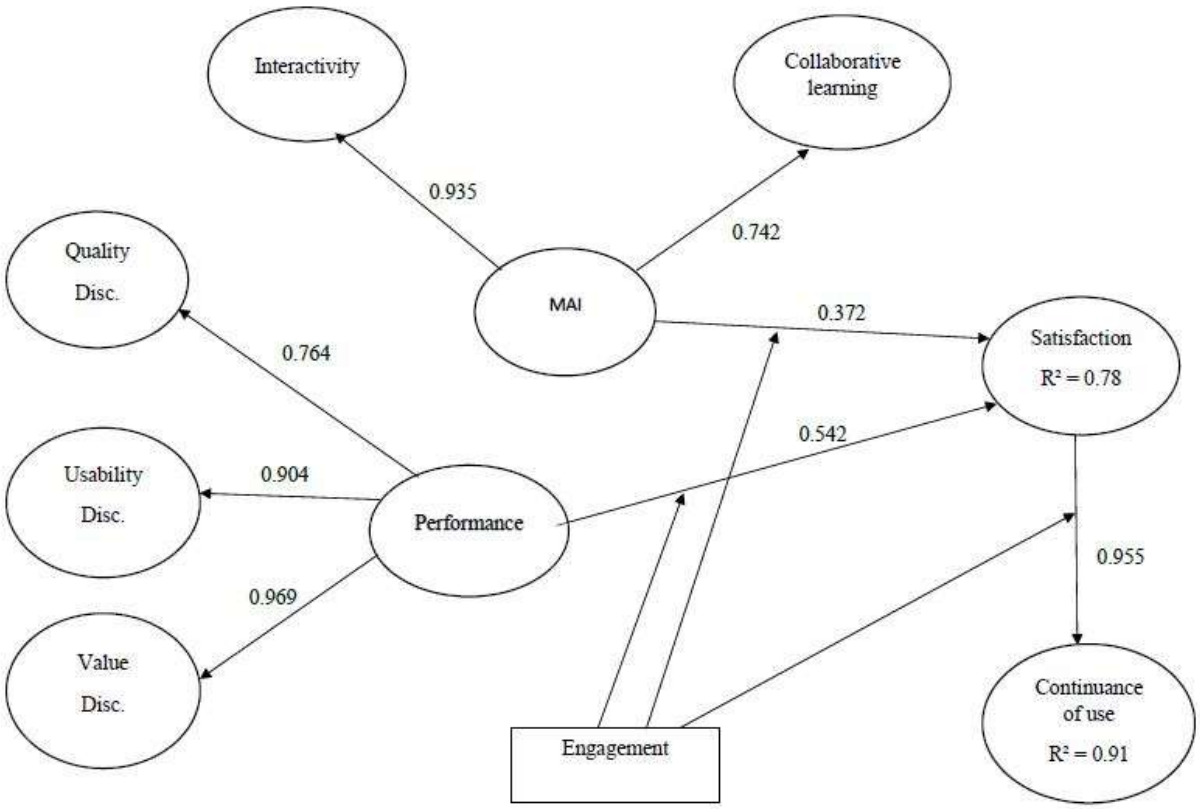
However, to determine if these differences exhibit invariance in the models, the multi-group analysis continues to compare the models. After this first step, with the adjustment of the

model for each group (Low and high), in the second step of the analysis, the invariance of the model was evaluated in the two groups by comparison of the unconstrained model (with factorial weights and variances / covariance of free factors) with a constrained model where the factorial weights and the variances / covariance of the two groups were fixed.

Finally, the invariance of the structural model was evaluated by comparing the model with free structural coefficients vs the model with fixed and equal structural coefficients in the two groups. The statistical significance of the difference of the two methods was made with the chi-square test, as described by Maroco (2014). With the difference of chi-squares between the model with fixed and free coefficients is 52.253, and p-value = 0.000, can be conclude that, for a 95% confidence level, that the model with fixed structural coefficients has a worse adjustment than the model with free structural coefficients, and because of this, the causal model is not invariant in the engagement groups.

Thus, the engagement variable integrates the final model as a moderator variable. The final model presented R^2 for high satisfaction and continuity of use. The final model found in this research is shown in Figure 11, containing the standardized weights.

Figure 11: Final model found by research



Source: Elaborated by author, 2018.

6. Discussion and conclusion

The study began with a review of the literature, investigating at an early stage the most appropriate terms to investigate MOOCs. In the first readings, the most relevant ones, such as Nicholson (2007), Spector (2014), Phelan (2015) and Steffens (2015) revealed that MOOCs are a more advanced stage of consolidated e-learning courses. Faced with this, seeking what has already been consolidated about e-learning is sensible to understand the MOOCs. Thus, the literature review was divided into two blocks, using search terms for e-learning and other search terms for MOOCs.

The main studies on e-learning and MOOCs were researched and then a bibliometric and systematic review with meta-analysis was conducted. This strategy of literature review was more adequate to the objective of the review, to seek studies that would provide a theoretical foundation for a research model focused on xMOOCs, considering the current business model of the platforms. In all, 89 studies were used.

Of the studies on e-learning, the majority are quantitative studies, with validation of theoretical models or replications of already validated models. While in the MOOCs, there is a clear division between studies with a qualitative approach and quantitative exploratory studies. These characteristics allow us to take advantage of what has already been consolidated with e-learning to study MOOCs. When evaluating data from the Web of Science, it was perceptible that the quantitative studies belonged to a cluster of publications on the subject, being in this cluster of publications that this research finds theoretical support.

The engagement variable in this research was formed from five variables. Some studies, such as Sawang, Newton and Jamieson (2012), Perna et al. (2014) and Greene, Oswald, and Pomerantz (2015) used the engagement to evaluate them in a descriptive way that would help to understand certain profiles of users of MOOCs. In this research, the intention was the same, with the particularity that it was necessary to use the engagement inserted in a theoretical model to be tested empirically.

Due to the theoretical support given by these studies, the engagement variables were measured as they were measured in previous studies. In order to form a single measure of engagement, the random forest procedure was used to classify individuals as to their engagement. Two distinct groups were identified, those of low and high engagement, corroborating the approach and findings of the studies that investigated the influence of engagement on satisfaction with MOOCs.

The variables of the demographic profile were crossed with the variables of satisfaction and continuance of use, in order to verify the disparity between the means by group. None of the demographic variables presented different means of satisfaction and continuance across their categories. These results contradict the findings on demographic variables in e-learning studies, such as that of Mohammadyari and Singh (2015), Chen (2011), Pynoo (2011), Raaij and Schepers (2008) and Teo (2010). In these studies, differences by gender and age range are present. In studies on MOOCs, however, studies that address differences by demographic variables are losing ground to other aspects, such as the level of online collaboration (Shen, & Ku (2015), social networking usage profile (Zhuhadar, Kruk and Daday, 2015) and the user's learning profile (Chang, Hung, & Lin, 2015).

When crossing variables, some results reveal different behaviors in certain groups. First, women, despite evaluating dependent variables with similar means as men, pay much less for courses. This result indicates that gender, although not a moderating variable of the theoretical model, influences variables in the user profile and consequently in the framework of the business model in MOOCs.

Regarding the payment, it is possible to identify that the individuals with higher scholarity are the ones who pay the most. However, the means of the dependent variables for those who pay are slightly lower than for those who do not pay, and there is no difference in the engagement among those who pay.

The first impression of these results is that payment is an option more considered by men and that deciding to pay generates higher expectation and, therefore, generates a slight decrease in satisfaction and continuance intention. However, when evaluating cross tabulation of payment with the completeness of the courses, all those who paid completed the course. Given this result, it can be inferred that the payment is not an action of engagement, nor does it increase the levels of satisfaction and continuance intention. On the other hand, it is the payment action that ensures that the student will not evade the course. Even those users with low engagement, if they pay to take the course, they conclude. Thus, payment is a crucial factor in decreasing the high dropout rate commonly measured in MOOC courses.

Going deeper, when evaluating engagement, the variable that best differentiates levels of engagement is persistence. Persistence is a variable measured by the actions of the users within the courses, actions ranging from the obligatory tasks to the access to all the content of the course. By crossing data from persistence to payment, it was possible to identify that there is an equal distribution among those who pay along persistence levels. That is, a considerable portion of those who pay are not very persistent, accessing only the contents and obligatory

tasks to finish the course. The low persistence is an indication that precedes the evasion, however, the payment keeps the user until the end of the course.

When evaluating the summary measures for the manifested variables, the means in all the constructs oscillated little, around two points in the scale of zero to ten points, being between six and eight in the concordance scale. With affirmative items, without inversion of perspective, it is evidenced the general agreement of the sample regarding the evaluated items. For Performance items, users generally agree that the quality of the teaching platform, usability, and value-added are critical to achieving performance on the courses. Likewise, online collaboration during learning and the use of tools for interaction are also decisive to suit the learning style of MOOCs.

The results of the performance constructs were already expected, since they came from measures already validated in the studies of Chiu et al. (2005), Lambropoulos, Faulkner and Culwin (2012), Pereira et al. (2015), among others. The studies of adhesion to the MOOCs are supported in the literature and, although they have not been previously validated, have frequencies similar to those of previous studies, such as Ventura, Bárcena and Monje (2014) and Gillani and Eynon (2014). In addition, the items measured are also supported by the joint procedure of JAD and NGT for selection of variables.

As a measure of refinement and adjustment of the measurement scale, several approaches and techniques were used. When using the IRT, the items were evaluated in terms of the amount of information per point in the scale and the contribution of the item to each construct. Similarly, under the CTT approach, the items were evaluated for the adherence of the items to the factorial analysis and the internal consistency of the constructs. Although they presented different approaches, both converged on the refinement of the model.

In some questions, such as col5 (Interest to share) and col6 (knowledge shared), the correlations between these variables and the others of the collaborative learning factor are low, which may justify the low contribution of the items to the factor. Likewise, int5 (doubts and deepening) and int6 (interaction outside the platform) exhibited this behavior. In the selection of variables in the NGT / JAD, these are the last variables to be considered, that is, they were the least important ones highlighted by the specialists. Another factor corroborating the results of the correlations is the use of interaction and collaborative learning in exploratory studies, as in Lambropoulos, Faulkner and Culwin (2012) or in Zhuhadar, Kruk and Daday (2015). With the absence of the formation of theoretical constructs, the difficulties to conceptualize and define the scope of the concept of a construct are exacerbated. Thus, these items can be studied

in greater depth to find out whether they form another latent factor or whether they should be measured in a way other than by self-evaluation.

The other excluded items violated one or more EFA or CFA assumptions, or even presented a high level of common method bias. Violating some assumption of statistical techniques or even presenting an unknown bias may have causes such as: influence of data collection, wording of items and possible biases characteristic of the population. Being an infinite population, the sources of bias become more difficult to identify. However, the variable usa6 (Support productivity) violated the multicollinearity assumption. This assumption in other studies such as Chiu et al. (2005), Cheung and Lee (2011) and Pereira et al. (2015) suggest that a confirmatory item has high correlations with its disconfirmatory analogues, which may have allowed different items, even with diagnosed multicollinearity, to be admitted into the usability construct.

The model converged and became one in the AFC phase, when the last variable was excluded in CTT, qua5 (Content quality). In the IRT approach, qua5 was disapproved in the IIC, since it is the variable that contributes least to the quality model. This item was specifically used in less heterogeneous samples, often with equal content being evaluated by the majority of respondents. It is not uncommon that it is variable to offer little information in this study, since the diversity of courses resulted in evaluation of different contents by each respondent.

In the evaluation of mediation and moderation, the mediator effect presents in the initial model did not have the null hypothesis rejected, indicating that the performance is not mediator of the MAI relation and satisfaction. For the moderating variables, of the variables tested, only the engagement showed moderation power in the low and high engagement groups. Engagement moderates the relations of second order factors to satisfaction and moderates the relationship between satisfaction and continuance of use.

In conducting the structural model, the 2nd order factors: performance and MAI, were relevant for satisfaction. The performance constructs presented high standardized weights, as in other studies (Yu & Yu, 2010; Hu, Kuo & Lin, 2010). The non-standardized weights allow us to draw a ranking of the most relevant constructs, indicating that the value stands out from the quality and usability, being the most relevant construct. This result was the same as that found by the studies of Chiu et al. (2005), Lin (2012) and Pereira et al. (2015). Thus, the expectation with the added value that the MOOC can bring is decisive for achieving high performance.

In relation to MAI, interactivity and collaborative learning define the 2nd order construct, corroborating the findings in the NGT / JAD and confirming the exploratory studies that had

been investigating these latent factors (Liaw, 2008; Paechter, Maier & Macher, 2010; Chang, Hung & Lin, 2011; Lambropoulos, Faulkner & Culwin, 2012; Ventura, Bárcena & Monje, 2014). By the non-standardized weight, it is evidenced that interactivity is more relevant than collaborative learning.

Among the second order factors, both presented a high correlation (0.86), indicating that better MAI index was, better will be user's performance. The indirect effects of the 2nd order constructs for continuance of use were also identified and were relevant. In the direct effects on satisfaction, performance and MAI are responsible for 78% of satisfaction variance, with performance contributing slightly more than MAI.

For continuance of use, the variance rate was 91%. These results are widely corroborated by the literature (Sawang, Newton Jamieson, 2012, Chien, 2012, Cheng, 2012, Alraimi, Zo & Ciganek, 2015; Lindsey, Rhoads & Lozano, 2015). It is important to point out that continuance of use is commonly referred to by other names, such as success (Lin & Chen, 2012), effectiveness (Chien, 2012) or continuance intention (Alraimi, Zo & Ciganek, 2015). In the use of all these expressions, the concept of the final construct is the user's desire to continue using the service.

The hypotheses of the study were evaluated in a bi-variate way, using the null and alternative hypotheses of each pathway performed by the constructs to the 2nd order factors and the causal pathway recommended in the initial model. All null hypotheses were rejected according to the standardized and non-standardized regression weights presented, obtaining a final model of five latent factors of 1st order and 2nd of 2nd order, with satisfaction and consequent continuance of use as dependents.

It is emphasized that performance driven by disconfirmatory constructs finds theoretical basis in EDT (Oliver, 1980). With the inclusion of MAI, based only on exploratory and non-consolidated bases, it is necessary to compare with similar studies and to use the same dependent variables in similar contexts. According to Table 51, it is well known that the results are adequate and resemble the results of other studies.

Table 51: Comparative variances

Authors	Year	Satisfaction R^2	Continuance R^2
Alraimi, Zo, Ciganek	2015	0.33	0.64
Chiu et al	2011	0.64	0.67
Lin, Chen	2012	0.82	0.87
Lee	2010	0.65	0.80

Liaw	2008	0.50	0.59
Pereira et al.	2015	0.74	0.84
Present study	2018	0.78	0.91

Source: Research data and information obtained in mentioned papers, 2018.

6.1 Implications

The motivation of this study was to validate a theoretical model that could understand the current context of the MOOCs and that contemplates their peculiarities in the scope of distance learning. The theoretical model was empirically presented two theoretical aspects at different moments. The first, coming from e-learning, with already consolidated studies, were integrated as historical base of MOOCs. The second, coming from studies still exploratory, seeking answers of a phenomenon that has not yet reached maturity.

The objectives of this study were reached, since the theoretical model was validated from an extensive literature review and rigorously applied statistical procedures. The model can be replicated in MOOC teaching environments and adapted for e-learning studies. The model also has the capacity to deal with a heterogeneous public, demonstrating adaptability in different economic and cultural contexts that can influence the evaluation of the users.

The main product of this study is an evaluation tool for satisfaction and continuance of use in MOOC courses, which can be used for management purposes and also to evaluate the evolution of a MOOC course, a group of users or even a platform education. The model of effectiveness of MOOCs also contributes to the understanding of the evolution of MOOCs as a teaching modality and as a business model, involving large universities, companies that invest in teaching and society in general, being potential users.

In the context of statistical and methodological procedures, the study sought to apply CBSEM in a complete way, as suggested by the main authors of the technique (Kline, 2005, Hair et al., 2011, Maroco, 2014), and aggregating impartial literature review procedures and standardized, and robust measurement and scale techniques, which guarantee reliability and parsimony of the studied phenomenon.

The IRT was introduced as a concurrent approach. However, the IRT findings are similar to and complement CTT findings. In the near future, the IRT with more development and insertion in more popular statistical packages, tends to be the most reliable approach in the conduction of theoretical behavioral models.

6.2 Limitations and future studies

The research presents some limitations, which do not make the study unfeasible, but should be considered to be minimized in later studies. First, there is a heterogeneity of the sample, since respondents could be from anywhere with Internet access. This heterogeneity can cause biases, which should be identified and treated as much as possible. In this study, biases were identified from CMV and items with problems were removed. In future studies, homogeneity of the sample can be achieved with collection strategies in a smaller range. In addition, some biases can be identified from a bias scale inserted in the search appliance.

Second, the IRT was used in this research, but not to its full potential. Factor scores were little explored and the results of the ICC and IIC graphics could be evaluated by means of summary measures and comparisons of mean values of the amount of information per point of the scale and per variable. As the IRT is still in development in 2018, even in R, there is little support material and much need for development of IRT applied to behavioral studies and that generate outputs analogous to CFA and CBSEM. The IRT is theoretically the best option to evaluate behavioral variables, in addition to requiring a smaller sample and less assumptions.

Future studies can develop and deepen model constructs that still deserve calibration, such as MAI. With qualitative and less embedded approaches to information retrieval, other variables and latent factors not considered in this study can emerge and increase the MAI. Aspects that are indirectly determinant must also be considered, such as the increasing number of courses that charge for the service, the imminent maturity of the MOOCs business model, and the role of universities in this business model.

References

Abeer, W., & Miri, B. (2014). Students' preferences about learning in a MOOC. *Procedia Social and Behavioral Sciences, v. 152*, pp. 318-323. doi:10.1016/j.sbspro.2014.09.203

Anderson, J., & Gerbin, D. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 3*(3).

APA. (2001). *Publication Manual of the American Psychological Association* (5 ed.). Washington: APA. Retrieved from American Psychological Association.

Assis, S. (2015). *Business Intelligence.* São Paulo: s/.

Babin, B., & Svensson, G. (2012). Structural equation modeling in social science research: Issues of validity and reliability in the research process. *European Business Review, 12*(4).

Bagozzi, R., & Yi, P. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly, 36*(3).

Bernhard, W., Bittel, N., Vlies, S., Bettoni, M., & Roth, N. (2013). The MOOCs business model. *Procedia Social and Behavioral Sciences, v. 106*, pp. 2931-2937. doi:10.1016/j.sbspro.2013.12.339

Bond, P., & Leibowitz, F. (2013). MOOCs and Serials. *Serials Review, v. 39*, 258-260. doi:10.1016/j.serrev.2013.10.007

Chen, J. (2011). The effects of education compatibility and technological expectancy on e-learning acceptance. *Computers & Education, v. 57*, pp. 1501-1511. doi:10.1016/j.compedu.2011.02.009

Chen, P., & Spector, P. (1991). Negative affectivity as the underlying cause of correlations between stressors and strains. *Journal of Applied Psychology, 52*.

Chiu, C., Hsu, M., Sun, S., Lin, T., & Sun, P. (2005). Usability, quality, value and e-learning continuance decisions. *Computers & Education, v. 45*, pp. 399-416. doi:10.1016/j.compedu.2004.06.001

Chow, W., & Shi, S. (2014). Investiganting students' satisfaction and continuance intention toward e-learning: An extension of the expectation confirmation model. *Procedia Social and Behavioral Sciences, 141*, 1145-1149. doi:10.1016/j.sbspro.2014.05.193

Creswell, J. (2012). *Qualitative inquiry & research design: Choosing among five approaches.* Thousand Oaks: Sage.

Darab, B., & Montazer, G. (2011). An ecletic model for assessing e-learning readiness in the Iranian universities. *Computers & Education, v. 56*, pp. 900-910. doi:10.1016/j.compedu.2010.11.002

Davis, M., & Behara, R. (2015). Navigating disruptive innovation in Undergraduate Business Education. *Decision Sciences Journal of Innovative Education, v. 13*(3), pp. 305-327.

DeLone, W., & McLean, E. (2003). The DeLone and McLean Model of information system success: a ten year update. *Journal of Management Information, 19*(4), 9-30.

Duggan, E., & Thachenkary, C. (2004). Integrating nominal group technique and joint application development for improved systems. *Information & Management, 41*.

Erdogmus, N., & Esen, M. (2011). An investigation of the effects of technology readiness on technology acceptance in e-HRM. *Procedia Social and Behavioral Sciences, v. 24*, pp. 487-495. doi:10.1016/j.sbspro.2011.09.131

Faul, F. (2017). G*Power 3.1 Manual. Dusseldorf, Germany. Retrieved from http://www.gpower.hhu.de

Faul, F., Erdfelder, E., Buchner, A., & A., L. (2009). Statistical power analyses using G*Power 3.1: test for correlation and regression analyses. *Behavior Research Methods, 41*(4).

Fávero, L., Belfiore, P., Silva, F., & Chan, B. (2009). *Análise de dados: modelagem multivariada para tomada de decisões.* Rio de Janeiro: Campus.

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1).

Gómez, F., Guardiola, J., Rodríguez, O., & Alonso, M. (2012). Gender differences in e-learning satisfaction. *Computers & Education, v. 58*, pp. 283-290. doi:10.1016/j.compedu.2011.08.017

Guanilo, M., Takahashi, R., & Bertolozzi, M. (2011). Revisão sistemática: noções gerais. *Revista da Escola de Enfermagem da USP*.

Hair, J., William, B., Babin, B., Anderson, R., & Tatham, R. (2011). *Análise multivariada de dados.* Porto Alegre: Bookman.

Hauge, O., Conradi, R., & Ayala, C. (2010). Adoption of open source software in software-intensive organizations - A systematic literature review. *Information and Software Technology, 52*(11).

Hew, K., & Cheung, W. (2014). Students' and instructors' use of massive open online courses (MOOCs): motivations and challenges. *Educational Research Review, v. 12*, pp. 45-58. doi:10.1016/j.edurev.2014.05.001

Kitchenham. (2007). *Guidelines for performing systematic literature reviews in software engineering.* Durham: Keele University.

Kline, R. (2005). *Análise multivariada de dados.* Porto Alegre: Bookman.

Koller, S., Couto, M., & Hohendorff, J. (2014). *Manual de produção científica.* Porto Algre: Penso.

Kose, I., & Demirtasli, N. (2012). Comparison of unidimensional and multidimensional models based on item response theory in terms of both variables of test length and sample size. *Procedia - Social and Behavioral Sciences, 46*.

Kuhn, M. (2017). caret: Classification and regression training. R-package version 6.0-78. Retrieved from https://CRAN.R-project.org/package=caret

Lai, X., Li, F., & Leung, K. (2013). A monte carlo study the effects of common method variances on significance testing and parameter bias in hierarchical linear modeling. *Organizationl Research Methods, 16*.

Lee, B., Yoon, J., & Lee, I. (2009). Learners' acceptance of e-learning in South Korea: theories and results. *Computers & Education, v. 53*, pp. 1320-1329. doi:10.1016/j.compedu.2009.06.014

Lee, M. (2010). Explaining and predicting user's continuance intention toward e-learning: An extension of the expectation-confirmation model. *Computers & Education, v. 56*, pp. 506-516. doi:10.1016/j.compedu.2009.09.002

Leonidou, C., & Leonidou, L. (2011). Research into environmental marketing/management: a bibliographic analysis. *European Journal of Marketing, 45*, pp. 68-103.

Liaw, M., & Wiener, M. (2002). Classification and regression by randomForest. *R-News, 2*(3).

Lin, Y., Lin, H., & Hung, T. (2015). Value hierarchy for massive open online courses. *Computers in Human Behavior, v. 53*, pp. 408-418. doi:10.1016/j.chb.2015.07.006

Littell, J., Corcoran, J., & Pillai, V. (2008). *Systematic review and meta-analysis.* New York: Oxford Universisty Press.

Little, G. (2013). Massively Open? *The Journal of Academic Librarianship, v. 39*, 308-309. doi:10.1016/j.acalib.2013.03.004

Loya, A., Gopal, A., Shukla, I., Jermann, P., & Tormey, R. (2015). Conscientious behaviour, flexibility and learning in massive open on-line courses. *Procedia Social and Behavioral Sciences, v. 191*, pp. 519-525. doi:10.1016/j.sbspro.2015.04.686

Luaces, O., Díez, J., Betanzos, A., Troncoso, A., & Bahamonde, A. (2015). A factorization approach to evaluate open-response assignments in MOOCs using preference learning on peer assessmets. *Knowledge-Based Systems, v. 85*, pp. 322-328. doi:10.1016/j.knosys.2015.05.019

Mackenzie, S., Podsakoff, P., & Podsakoff, N. (2011). Construct measurement and validation procedures in MIS and behavioral research: integrating new and existing techniques. *MIS Quarterly, 35*(2).

Maroco, J. (2014). *Análise de equações estruturais: Fundamentos teóricos, software e aplicações.* Perô Pinheiro: Report Number.

McAndrew, P., & Scanlon, E. (2013). Open learning at a distance: lessons for struggling MOOCs. *Science, v. 342*, pp. 1450-1451.

Merino, P., Valiente, J., Hoyos, C., Sanagustín, M., & Kloos, C. (2015). Precise effectiveness strategy for analyzing the effectiveness of students with educational resources and activities in MOOCs. *Computers in Human Behavior, v. 47*, pp. 108-118. doi:10.1016/j.chb.2014.10.003

Meyer, D., Dimitriadov, E., Hornik, C., Weingessel, A., & Leisch, F. (2015). e1071: Misc functions of the department of statistics, probability theory group. R package version 1.6-7. Retrieved from https://CRAN.R-project.org/package=e1071

Moghadamzadeh, A., Salehi, K., & Khodaie, E. (2011). A comparison method of equating classic and Item Response Theory (IRT): A case of iranian study in the University Entrance Exam. *Procedia - Social and Behavioral Sciences, 29*.

Muenchen, R. (2012). *The popularity of Data Analysis Software.*

Mulaik, S. (1989). Evaluation of goodness of fit indices for structural equation models. *Psychological Bulletin, 105*(3).

Nebojsa, S. (2014). The use and misuse of structural equation modeling in management research: a review and critique. *Journal of Advances in Management Research, 11*(1).

Nicholson, P. (2007). A history of e-learning: Echoes of the pioneers. In R. Manjón, *E-learning: from theory to practice* (p. 241). Dordrecht: Springer.

Nunes, M., & Goldszmidt, R. (2013). Uma análise da escala de liderança autêntica utilizando análise fatorial confirmatória linear e teoria da resposta ao item. *XXXVII Encontro da ANPAD.* Rio de Janeiro.

Oliver, J. (1980). Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services. *Journal of Marketing Research, 17*(4), pp. 460-469.

Pai, M. e. (2004). Clinical Research Methods - Systematic reviews and meta-analyses: an ilustrated, step-by-step guide. *The National Medical Journal of India*(2).

Parasuraman, A., Zeithaml, V., & Berry, L. (1988). Servqual: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12-37.

Pendlebury, D. (2008). *Using bibliometrics in evaluating research.* Thomson Reuters.

Pereira, F. (2013). *A satisfação e a intenção de continuidade de uso em serviços de e-elearning: validação empírica de um modelo aplicado no serviço público.* Natal: UFRN.

Phillips, J., & Lord, R. (n.d.). Schematic information processing and perceptions of leadership in problem-solving groups. *Journal of Applied Psychology, 67*.

Pinho, J., & Macedo, I. (2008). Examining the antecedents and consequences of online satifaction within the public sector: The case of taxation services. *Transforming Government People, Process and Policy*.

Pocinho, M. (2008). *Lições de metanálise.* Retrieved 12 15, 2017, from ISMT: http://docentes.ismt.pt/~m_pocinho/Licoes_de_revisao_sistematica_e_metanalise

Podsakoff, P., Mackenzie, S., Lee, J., & Podsakoff, N. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5).

Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation, 25*, 348-349.

Pynoo, B., Devolder, P., Tondeur, J., Braak, J., Duyck, W., & Duyck, P. (2011). Predicting secondary school teachers' acceptance and use of a digital learning environment: A cross-

sectional study. *Computers in Human Behavior, v. 27*, pp. 568-575. doi:10.1016/j.chb.2010.10.005

R, C. T. (2018). *R: A language and environment for statistical computing. R Foundation for statistical computing.* Vienna, Austria. Retrieved from https://www.R-project.org/

Raaij, E., & Schepers, J. (2008). The acceptance and use of a virtual learning environment. *Computers & Education, v. 50*, pp. 838-852. doi:10.1016/j.compedu.2006.09.001

Reich, J. (2015). Rebooting MOOC research. *Science, v. 347*.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*(5). Retrieved from http://www.jstasoft.org/v17/i05/

Roca, J., Chiu, C., & Martínez, F. (2006). Understanding e-learning continuance intention: An extension of the technology acceptance model. *International Journal of Human-Computers Studies, v. 64*, pp. 683-696. doi:10.1016/j.ijhcs.2006.01.003

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2).

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika, 34*.

Schaltegger, S., Gibassier, D., & Zvezdov, D. (2013). Is environmental management accounting a discipline? A bibliometric literature review. *Meditari Accountancy Research, 1*(1), pp. 4-31.

Scholten, A. (2017a). Quantitative methods - Module 1: origins. In *Quantitative methods.* University of Amsterdam.

Scholten, A. (2017b). Quantitative methods - Module 2: Scientific method. In *Quantitative methods.* University of Amsterdam.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *elearnspace: everything learning*, pp. 1-18.

Siritongthaworn, S., & Krairit, D. (2006). Satisfaction in e-learning: the context of supplementary instruction. *Campus-Wide Information Systems, v. 23*(2), pp. 76-91. doi:10.1108/10650740610654465

Sjostedt, E., Aldberg, H., & Jacobsson, C. (2015). *Guidelines for using bibliometrics at the Swedish Research Council.* Vetenskapsradet.

Spector, P. (2006). Method variance in organizational research: truth or urban legend? *Organizational Research Methods, 9*(2).

Spinak, E. (1988). Indicadores cienciométricos. *Ciência da Informação, 27*(2), 141-148.

Summak, M., Baglibel, M., & Samancioglu, M. (2010). Technology readiness of primary school teachers: A case study in Turkey. *Procedia Social and Behavioral Sciences, v. 2*, pp. 2671-2675. doi:10.1016/j.sbspro.2010.03.393

Teo, T. (2010). Using structural equation modelling (SEM) in educational technology research: Issues and guidelines. *British Journal of Educational Technology, 41*(6).

Thomas, K., & Kilmann, R. (1975). The social desirability variable in organizational research: An alternative explanation for reported findings. *Academy of Management Journal, 18*.

Tjahono, B. e. (2010). Six sigma: a literature review. *International Journal of Lean Six Sigma, 1*(3), pp. 216-233.

Udo, G., Bagchi, K., & Kirs, P. (2011). Using SERVQUAL to assess the quality of e-learning experience. *Computers in Human Behavior*. doi:10.1016/j.chb.2011.01.009

Ulrich, C., & Nedelcu, A. (2015). MOOCs in our university: Hopes and worries. *Procedia Social and Behavioral Sciences, v. 180*, pp. 1541-1547. doi:10.1016/j.sbspro.2015.02.304

Walpole, R., Myers, R., Myers, S., & Ye, K. (2009). *Probabilidade e estatística para engenharia e ciências* (8 ed.). São Paulo: Prentice Hall.

Wu, J., Tennyson, R., & Hsia, T. (2010). A study of student satisfaction in a blended e-learning system environment. *Computers & Education, v. 55*, pp. 155-164. doi:10.1016/j.compedu.2009.12.012

Zhang, L., Wen, H., Li, D., Fu, Z., & Cui, S. (2010). E-learning adoption intention and its key influence factors based on innovation adoption theory. *Mathematical and Computer Modelling, v. 51*, pp. 1428-1432. doi:10.1016/j.mcm.2009.11.013

For bibliometric and systematic analysis

Abeer, W., & Miri, B. (2014). Students' preferences about learning in a MOOC. *Procedia Social and Behavioral Sciences, v. 152*, pp. 318-323. doi:10.1016/j.sbspro.2014.09.203

Ahmed, H. (2010). Hybrid e-learning acceptance model: Learner perceptions. *Decision Sciences Journal of Innovative Education, v. 8*(2), pp. 313-346.

Alraimi, K., Zo, H., & Ciganek, A. (2015). Understanding the MOOCs continuance: The role of openness and reputation. *Computers & Education, v. 80*, pp. 28-38. doi:10.1016/j.compedu.2014.08.006

Baturay, M. (2015). An overview of the world of MOOCs. *Procedia Social and Behavioral Sciences, v. 174*, pp. 427-433. doi:10.1016/j.sbspro.2015.01.685

Behara, R., & Davis, M. (2015). Navigating disruptive innovation in undergraduate business education. *Decision Sciences Journal of Innovative Education, v. 13*(3), pp. 305-327.

Belleflamme, P., & Jacqmin, J. (2015). Business models and impacts on higher education. *CESifo Economic Studies*, pp. 1-22. doi:10.1093/cesifo/ifv016

Bernhard, W., Bittel, N., Vlies, S., Bettoni, M., & Roth, N. (2013). The MOOCs business model. *Procedia Social and Behavioral Sciences, v. 106*, pp. 2931-2937. doi:10.1016/j.sbspro.2013.12.339

Bharuthram, S., & Kies, C. (2013). Introducing e-learning in a South African higher education institution: Challenges arising from an intervention and possible responses. *British Journal of Educational Technology, v. 44*(3), pp. 410-420. doi:10.1111/j.1467-8535.2012.01307.x

Biasutti, M. (2011). The student experience of a collaborative e-learning university module. *Computers & Education, v. 57*, pp. 1865-1875. doi:10.1016/j.compedu.2011.04.006

Boe, T., Gulbrandsen, B., & Sorebo, O. (2015). How to stimulate the continued use of iCT in higher education: Integrating information systems continuance theory and agency theory. *Computers in Human behavior, 50*, 375-384. doi:10.1016/j.chb.2015.03.084

Cai, S., & Zhu, W. (2012). The impact of an online learning community project on university chinese as a foreign language students' motivation. *Foreign Language Annals, v. 45*(3), pp. 307-329. doi:10.111/j.1944-9720.2012.01204.x

Chang, R., Hung, Y., & Lin, C. (2015). Survey of learning experiences and influence of e-learning style preferences on user intentions regarding MOOCs. *British Journal of Educational Technology, v. 46*(3), pp. 528-541. doi:10.1111/bjet.12275

Chen, J. (2011). The effects of education compatibility and technological expectancy on e-learning acceptance. *Computers & Education, v. 57*, pp. 1501-1511. doi:10.1016/j.compedu.2011.02.009

Chen, Y., & Chen, P. (2015). MOOC study group: facilitation strategies, influential factors, and student perceived gains. *Computers & Education, v. 86*, pp. 55-70. doi:10.1016/j.compedu.2015.03.008

Cheng, Y. (2012). Effects of quality antecedents on e-learning acceptance. *Internet Research, v. 22*(3), pp. 361-390. doi:10.1108/10662241211235699

Cheung, C., & Lee, M. (2011). Antecedents and consequences of user satisfaction with an e-learning portal. *International Journal of Digital Society, v. 2*(1), pp. 1-8.

Chien, T. (2012). Computer self-efficacy and factors influencing e-learning effectiveness. *European Journal of Training and Development, v. 36*(7), pp. 670-686. doi:10.1108/03090591211255539

Chiu, C., Wang, E., Shih, F., & Fan, Y. (2011). Understanding knowledge sharing in virtual communities: An integration of expectancy disconfirmation and justice theories. *Online Information Review, v. 35*(1), pp. 134-153. doi:10.1108/14684521111113623

Chou, H., Lin, C., Woung, L., & Tsai, M. (2012). Engagement in e-learning opportunities: An empirical study on patient education using expectation confirmation theory. *Journal of Medical Systems, v. 36*, pp. 1697-1706. doi:10.1007/s10916-010-9630-9

Chou, S., Min, H., Chang, Y., & Lin, C. (2010). Understanding continuance intention of knowledge creation using extended expectation-confirmation theory: An empirical study of Taiwan and China online communities. *Behaviour & Information Technology, v. 29*(6), pp. 557-570. doi:10.1080/01449290903401986

Chow, W., & Shi, S. (2014). Investiganting students' satisfaction and continuance intention toward e-learning: An extension of the expectation confirmation model. *Procedia Social and Behavioral Sciences, 141*, 1145-1149. doi:10.1016/j.sbspro.2014.05.193

Chyung, S., & Vachon, C. (2013). An investigation of the profiles of satisfying and dissatisfying factors in e-learning. *Performance Improvement Quarterly, v. 26*(2), pp. 117-140. doi:10.1111/j.1937-8327.2005.tb00335.x

Clara, M., & Barbera, E. (2014). Three problems with the connectivist conbception of learning. *Journal of Computer Assisted Learning*(30), pp. 197-206. doi:10.1111/jcal.12040

Clarke, T. (2013). The advance of the MOOCs (massive open online courses): The impeding globalisation of business education? *Education+Training, v. 55*(5), pp. 403-413. doi:10.1108/00400911311326036

Darab, B., & Montazer, G. (2011). An ecletic model for assessing e-learning readiness in the Iranian universities. *Computers & Education, v. 56*, pp. 900-910. doi:10.1016/j.compedu.2010.11.002

Daskalakis, S., & Tselios, N. (2011). Evaluating e-learning initiatives: A literature review on methods. *International Journal of Web-based Learning and teaching Technologies, v. 6*(1), pp. 35-51. doi:10.4018/jwltt.2011010104

Davis, M., & Behara, R. (2015). Navigating disruptive innovation in Undergraduate Business Education. *Decision Sciences Journal of Innovative Education, v. 13*(3), pp. 305-327.

Duque, L., & Weeks, J. (2010). Towards a model and methodology for assessing student learning outcomes and satisfaction. *Quality Assurance in Education, v. 18*(2), pp. 84-105. doi:10.1108/09684881011035321

Erdogmus, N., & Esen, M. (2011). An investigation of the effects of technology readiness on technology acceptance in e-HRM. *Procedia Social and Behavioral Sciences, v. 24*, pp. 487-495. doi:10.1016/j.sbspro.2011.09.131

Franklin, D. (2015). Will the internet ever replace colleges and universities as we know it today? An internet discussion about the future of higher education. *Procedia Social and Behavioral Sciences, v. 176*, pp. 738-744. doi:10.1016/j.sbspro.2015.01.534

Gillani, N., & Eynon, R. (2014). Communication patterns in massively open online courses. *Internet and Higher Education, v. 23*, pp. 18-26. doi:10.1016/j.iheduc.2014.05.004

Gómez, F., Guardiola, J., Rodríguez, O., & Alonso, M. (2012). Gender differences in e-learning satisfaction. *Computers & Education, v. 58*, pp. 283-290. doi:10.1016/j.compedu.2011.08.017

Greene, J., Oswald, C., & Pomerantz, J. (2015). Predictors of retention and achievement in a Massive Open Online Course. *American Educational Research Journal*, pp. 1-31. doi:10.3102/0002831215584621

Hew, K. (2014). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology*, pp. 1-22. doi:10.1111/bjet.12235

Hew, K., & Cheung, W. (2014). Students' and instructors' use of massive open online courses (MOOCs): motivations and challenges. *Educational Research Review, v. 12*, pp. 45-58. doi:10.1016/j.edurev.2014.05.001

Ho, L., Kuo, T., & Lin, B. (2010). Influence of online learning skills in cyberspace. *Internet Research, v. 20*(1), pp. 55-71. doi:10.1108/10662241011020833

Hung, H., & Cho, V. (2008). Continued usage or e-learning communication tools: a study from the learners' perspective in Hong Kong. *International Journal of Training and Development, v. 12*(3), pp. 171-187.

Ismail, N. Z., Razak, .. R., Zakariah, Z., Alias, N., & Aziz, M. N. (2012). E-learning continuance intention among higher learning institution student's in Malaysia. *Procedia Social and Behavioral Sciences, 67*, 409-415. doi:10.1016/j.sbspro.2012.11.345

Kesim, M., & Altinpulluk, H. (2015). A theoretical analysis of MOOCs types from a perspective of learning theories. *Procedia Social and Behavioral Sciences, v. 186*, pp. 15-19. doi:10.1016/j.sbspro.2015.04.056

King, G., & Sen, M. (2013). The future of colleges and universities. *Symposium*, pp. 1-7.

Klaos, S. (2011). Factors influencing learners' satisfaction in an open e-learning environment. *Journal of Teacher Education for Sustainability, v. 13*(1), pp. 29-42. doi:10.2478/v10099-011-0003-3

Klobas, J. (2014). Measuring the success of scaleable open online courses. *Performance Measurements and Metrics, v. 15*(3), pp. 145-162. doi:10.1108/PMM-10-2014-0036

Lai, M. (2008). Technology readiness, internet self-efficacy and computing experience of professional accounting students. *Campus-Wide Information Systems, v. 25*(1), pp. 18-29. doi:10.1108/10650740810849061

Lambropoulos, N., Faulkner, X., & Culwin, F. (2012). Supporting social awareness in collaborative e-learning. *British Journal of Educational Technology, v. 43*(2), pp. 295-306. doi:10.1111/j.1467-8535.2011.01184.x

Lee, B., Yoon, J., & Lee, I. (2009). Learners' acceptance of e-learning in South Korea: theories and results. *Computers & Education, v. 53*, pp. 1320-1329. doi:10.1016/j.compedu.2009.06.014

Lee, M. (2010). Explaining and predicting user's continuance intention toward e-learning: An extension of the expectation-confirmation model. *Computers & Education, v. 56*, pp. 506-516. doi:10.1016/j.compedu.2009.09.002

Lee, Y., Hsieh, Y., & Hsu, C. (2011). Adding innovation diffusion theory to the technology acceptance model: Supporting employees' intentions to use e-learning systems. *Educational Technology & Society, v. 14*(4), pp. 124-137.

Li, Y., Duan, Y., Fu, Z., & Alford, P. (2012). An empirical study on behavioural intention to reuse e-learning sytems in rural China. *British Journal of Educational Technology, v. 43*(6), pp. 933-948. doi:10.1111/j.1467-8535.2011.01261.x

Liaw, S. (2008). Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education, v. 51*, pp. 864-873. doi:doi:10.1016/j.compedu.2007.09.005

Lin, K. (2011). E-learning continuance intention: Moderating effects of user e-learning experience. *Computers & Education, v. 56*, pp. 515-526. doi:10.1016/j.compedu.2010.09.017

Lin, T., & Chen, C. (2012). Validating the satisfaction and continuance intention of e-learning systems: combining TAM and IS success models. *International Journal of Distance Education Technologies, v. 10*(1), pp. 44-54. doi:10.4018/jdet.2012010103

Lin, W. (2012). Perceived fit and satisfaction on web learning performance: IS continuance intention and task-technology fit perspectives. *International Journal of Human-Computer Studies, 70*, pp. 498-507. doi:10.1016/j.ijhcs.2012.01.006

Lin, W., & Wang, C. (2012). Antecedences to continued intentions of adoptiong e-learning system in blended learning instruction: A contigency framework based on models of information system success and tas-technology fit. *Computers & Education, v. 58*, pp. 88-99. doi:10.1016/j.compedu.2011.07.008

Lin, Y., Lin, H., & Hung, T. (2015). Value hierarchy for massive open online courses. *Computers in Human Behavior, v. 53*, pp. 408-418. doi:10.1016/j.chb.2015.07.006

Lindsey, B., Rhoads, R., & Lozano, J. (2015). Virtually unlimited classrooms: pedagogical practices in massive open online courses. *Internet and Higher Education, v. 24*, pp. 1-12. doi:10.1016/j.iheduc.2014.07.001

Loya, A., Gopal, A., Shukla, I., Jermann, P., & Tormey, R. (2015). Conscientious behaviour, flexibility and learning in massive open on-line courses. *Procedia Social and Behavioral Sciences, v. 191*, pp. 519-525. doi:10.1016/j.sbspro.2015.04.686

Luaces, O., Díez, J., Betanzos, A., Troncoso, A., & Bahamonde, A. (2015). A factorization approach to evaluate open-response assignments in MOOCs using preference learning on peer assessmets. *Knowledge-Based Systems, v. 85*, pp. 322-328. doi:10.1016/j.knosys.2015.05.019

Ma, J., Zheng, J., & Zhao, G. (2015). The applicable strategy for the courses alliance in regional universities based on MOOC platform. *Procedia Social and Behavioral Sciences, v. 176*, pp. 162-166. doi:10.1016/j.sbspro.2015.01.457

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education, v. 80*, pp. 77-83. doi:10.1016/j.compedu.2014.08.005

McAndrew, P., & Scanlon, E. (2013). Open learning at a distance: lessons for struggling MOOCs. *Science, v. 342*, pp. 1450-1451.

McGill, T. J., Klobas, J. E., & Renzi, S. (2014). Critical success factors for the continuation of e-learning intuitives. *Internet and Higher Education, v. 22*, pp. 24-36. doi:10.1016/j.iheduc.2014.04.001

Merino, P., Valiente, J., Hoyos, C., Sanagustín, M., & Kloos, C. (2015). Precise effectiveness strategy for analyzing the effectiveness of students with educational resources and activities in MOOCs. *Computers in Human Behavior, v. 47*, pp. 108-118. doi:10.1016/j.chb.2014.10.003

Mohammadyari, S., & Singh, H. (2015). Understanding the effect of e-learning on individual performance: The role of digital literacy. *Computers & Education, v. 82*, pp. 11-25. doi:10.1016/j.compedu.2014.10.025

Ossiannilsson, E. (2012). Quality enhancement on e-learning. *Campus-Wide Information Systems, v. 29*(4), pp. 312-323. doi:10.1108/10650741211253903

Paechter, M., Maier, B., & Macher, D. (2010). Computers & Education. *Computers & Education, v. 54*, pp. 222-229. doi:10.1016/j.compedu.2009.08.005

Pereira, F. A., Ramos, A. S., Gouvêa, M. A., & Costa, M. F. (2015). Satisfaction and continuous use intention of e-learning service in Brazilian public organizations. *Computers in Human Behavior, 46*, 139-148. doi:10.1016/j.chb.2015.01.016

Perna, L., Ruby, A., Boruch, R., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). moving through MOOCs: Understanding the progression of users in Massive Open Online Courses. *Educational Researcher, v. 43*(9), pp. 421-432. doi:10.3102/0013189X14562423

Phelan. (2015). The use of E-learning in social work education. *Social Work, v. 6*(3), pp. 256-264. doi:10.1093/sw/swv010

Pinho, J., & Macedo, I. (2008). Examining the antecedents and consequences of online satisfaction within the public sector: the case of taxation services. *Transforming Government: People, Proccess and Policy, v. 2*(3), pp. 177-193. doi:10.1108/17506160810902185

Porter, S. (2015). The economics of MOOCs: a sustainable future? *The Bottom Line: Managing Library Finances, v. 28*, pp. 52-62.

Pynoo, B., Devolder, P., Tondeur, J., Braak, J., Duyck, W., & Duyck, P. (2011). Predicting secondary school teachers' acceptance and use of a digital learning environment: A cross-sectional study. *Computers in Human Behavior, v. 27*, pp. 568-575. doi:10.1016/j.chb.2010.10.005

Raaij, E., & Schepers, J. (2008). The acceptance and use of a virtual learning environment. *Computers & Education, v. 50*, pp. 838-852. doi:10.1016/j.compedu.2006.09.001

Radford, A., Coningham, B., & Horn, L. (2015). MOOCs: not just for college students - How organizations can use MOOCs for professional development. *Employment Relations Today*, pp. 1-15. doi:10.1002/ert.21469

Reich, J. (2015). Rebooting MOOC research. *Science, v. 347*.

Sawang, S., Newton, C., & Jamieson, K. (2013). Increasing learners' satisfaction/intention to adopt more e-learning. *Education + Training, v. 55*(1), pp. 83-105. doi:10.1108/00400911311295031

Scott, P. (2014). The reform of English higher education: Universities in global, national and regional contexts. *Cambridge Journal of Regions, Economy and Society, v. 7*, pp. 217-231. doi:10.1093/cjres/rst021

Shen, C., & Kuo, C. (2015). Learning in massive open online courses: evidence from social media mining. *Computers in Human Behavior, v. 51*, pp. 568-577. doi:10.1016/j.chb.2015.02.066

Spector, M. (2014). Emerging educational technologies: Tensions and sinergy. *Journal of King Saud University - Computer and Information Sciences, v. 26*, pp. 5-10. doi:10.1016/j.jksuci.2013.10.009

Steffens, K. (2015). Competences, learning theories and MOOCs: recent developments in lifelong learning. *European Journal of Education, v. 50*(1). doi:10.1111/ejed.12102

Summak, M., Baglibel, M., & Samancioglu, M. (2010). Technology readiness of primary school teachers: A case study in Turkey. *Procedia Social and Behavioral Sciences, v. 2*, pp. 2671-2675. doi:10.1016/j.sbspro.2010.03.393

Teo, T. (2010). Development and validation of the e-learning acceptance measure (ELAM). *Internet and Higher Education, v. 13*, pp. 148-152. doi:10.1016/j.iheduc.2010.02.001

Terras, M., & Ramsay, J. (2015). Massive open online courses (MOOCs): insights and challenges from a psychological perspective. *British Journal of Educational Technology, v. 46*(3), pp. 472-287. doi:10.1111/bjet.12274

Udo, G., Bagchi, K., & Kirs, P. (2011). Using SERVQUAL to assess the quality of e-learning experience. *Computers in Human Behavior*, pp. 1-12. doi:10.1016/j.chb.2011.01.009

Ulrich, C., & Nedelcu, A. (2015). MOOCs in our university: Hopes and worries. *Procedia Social and Behavioral Sciences, v. 180*, pp. 1541-1547. doi:10.1016/j.sbspro.2015.02.304

Ventura, P., Bárcena, E., & Monje, E. (2014). Analysis of the impact of social feedback on written production and student engagement in language mooc. *Procedia Social and Behavioral Sciences, v. 141*, pp. 512-517. doi:10.1016/j.sbspro.2014.05.089

Wong, L., Tatnall, A., & Burgess, S. (2014). A framework for investigating blended learning effectiveness. *Education + Training, v. 56*, pp. 233-251. doi:10.1108/ET-04-2013-0049

Wu, J., Tennyson, R., & Hsia, T. (2010). A study of student satisfaction in a blended e-learning system environment. *Computers & Education, v. 55*, pp. 155-164. doi:10.1016/j.compedu.2009.12.012

Yu, T., & Yu, T. (2010). Modelling the factors that affect individuals' utilisation of online learning systems: An empirical study combining the task technology fit model with the theory of planned behavior. *British Journal of Educational Technology, v. 41*(6), pp. 1003-1017. doi:10.1111/j.1467-8535.2010.01054.x

Zhang, L., Wen, H., Li, D., Fu, Z., & Cui, S. (2010). E-learning adoption intention and its key influence factors based on innovation adoption theory. *Mathematical and Computer Modelling, v. 51*, pp. 1428-1432. doi:10.1016/j.mcm.2009.11.013

Zhuhadar, L., Kruk, S. R., & Daday, J. (2015). Semantically enriched Massive Open Online Courses (MOOCS) platform. *Computers in Human Behavior, v. 51*, pp. 578-593.

Appendix A: Script in R

```
#Descriptive analysis
```{r echo = T, results='asis',message=FALSE}
knitr::opts_chunk$set(echo = TRUE)
dad <- read.table(file="dad_tese.csv", sep = ";",header = T)
  dadint <- dad %>%
  select(num_range("int",1:6)) %>%
  summarise_all(funs(min = min, median = median,max = max,mean = mean, sd = sd))
dadint2 <- dadint %>% gather(stat, val) %>%
  separate(stat, into = c("var", "stat"), sep = "_") %>%
  spread(stat, val) %>%
  dplyr::select(var, min, median, max, mean, sd)
dadint3 <- dadint2 %>%
  mutate(CV = sd/mean)

dadcol <- dad %>%
  dplyr::select(num_range("col",1:7)) %>%
  summarise_all(funs(min = min, median = median,max = max,mean = mean, sd = sd))
dadcol2 <- dadcol %>% gather(stat, val) %>%
  separate(stat, into = c("var", "stat"), sep = "_") %>%
  spread(stat, val) %>%
  dplyr::select(var, min, median, max, mean, sd)
dadcol3 <- dadcol2 %>%
  mutate(CV = sd/mean)

dadqua <- dad %>%
  dplyr::select(num_range("qua",1:6)) %>%
  summarise_all(funs(min = min, median = median,max = max,mean = mean, sd = sd))
dadqua2 <- dadqua %>% gather(stat, val) %>%
  separate(stat, into = c("var", "stat"), sep = "_") %>%
  spread(stat, val) %>%
  dplyr::select(var, min, median, max, mean, sd)
dadqua3 <- dadqua2 %>%
  mutate(CV = sd/mean)

dadusa <- dad %>%
  dplyr::select(num_range("usa",1:6)) %>%
  summarise_all(funs(min = min, median = median,max = max,mean = mean, sd = sd))
dadusa2 <- dadusa %>% gather(stat, val) %>%
  separate(stat, into = c("var", "stat"), sep = "_") %>%
  spread(stat, val) %>%
  dplyr::select(var, min, median, max, mean, sd)
dadusa3 <- dadusa2 %>%
  mutate(CV = sd/mean)

dadval <- dad %>%
  dplyr::select(num_range("val",1:6)) %>%
  summarise_all(funs(min = min, median = median,max = max,mean = mean, sd = sd))
dadval2 <- dadval %>% gather(stat, val) %>%
  separate(stat, into = c("var", "stat"), sep = "_") %>%
  spread(stat, val) %>%
  dplyr::select(var, min, median, max, mean, sd)
dadval3 <- dadval2 %>%
  mutate(CV = sd/mean)

dadsatcon <- dad %>%
  dplyr::select(sat1,sat2,sat3,con1,con2,con3) %>%
  summarise_all(funs(min = min, median = median,max = max,mean = mean, sd = sd))
dadsatcon2 <- dadsatcon %>% gather(stat, val) %>%
  separate(stat, into = c("var", "stat"), sep = "_") %>%
```
```

```
  spread(stat, val) %>%
  dplyr::select(var, min, median, max, mean, sd)
dadsatcon3 <- dadsatcon2 %>%
  mutate(CV = sd/mean)

kable(dadint3, caption = "Interactivity measures")
kable(dadcol3, caption = "Collaborative learning measures")
kable(dadqua3, caption = "Quality measures")
kable(dadusa3, caption = "Usability measures")
kable(dadval3, caption = "Value measures")
kable(dadsatcon3,caption = "Satisfactions and Continuance measures")
```
```
require(dplyr)
require(ggplot2)
require(ltm)
require(tidyr)
require(randomForest)
require(caret)
require(e1071)
require(lavaan)

#data
setwd("~/Dropbox/2017.1/Tese/Analise")
dad = read.table(file="dad_tese.csv", sep = ";",header = T)
#creating samples with 50 and 100 observations
dadb <- dad[sample(nrow(dad),100), ]
dadc <- dad[sample(nrow(dad),50), ]
#removing levels with count equal 0
i=1
for(i in 1:39) {
  dadc[,i] <- factor(dadc[,i])
}
for(i in 1:39) {
  dadb[,i] <- factor(dadb[,i])
}
for(i in 1:39) {
  dad[,i] <- factor(dad[,i])
}

#irtmodel <- grm(dadltm, constrained = FALSE, IRT.param = TRUE, Hessian = FALSE,
#start.val = NULL, na.action = NULL, control = list())
#model for each construct
int <- grm(dad[c(1:6)])
col <- grm(dad[c(7:13)])
qua <- grm(dad[c(14:19)])
usa <- grm(dad[c(20:25)])
val <- grm(dad[c(26:31)])
satcon <- grm(dad[c(32:37)])
#comparing with constrained models
int2<- grm(dad[c(1:6)],constrained=T)
anova(int2,int)
col2<- grm(dad[c(7:13)],constrained=T)
anova(col2,col)
qua2<- grm(dad[c(14:19)],constrained=T)
anova(qua2,qua)
usa2<- grm(dad[c(20:25)],constrained=T)
anova(usa2,usa)
val2<- grm(dad[c(26:31)],constrained=T)
anova(val2,val)
satcon2<- grm(dad[c(32:37)],constrained=T)
```

```
anova(satcon2,satcon)

#results without coefficients
int_res<- summary(int2); int_res[2:9]
col_res <- summary(col); col_res[2:9]
qua_res <- summary(qua); qua_res[2:9]
usa_res <- summary(usa); usa_res[2:9]
val_res <- summary(val); val_res[2:9]
satcon_res <- summary(satcon); satcon_res[2:9]

#creating grm models with samples
intb <- grm(dadb[c(1:6)])
colb <- grm(dadb[c(7:13)])
quab <- grm(dadb[c(14:19)])
usab <- grm(dadb[c(20:25)])
valb <- grm(dadb[c(26:31)])
satconb <- grm(dadb[c(32:37)])

intc <- grm(dadc[c(1:6)])
colc <- grm(dadc[c(7:13)])
quac <- grm(dadc[c(14:19)])
usac <- grm(dadc[c(20:25)])
valc <- grm(dadc[c(26:31)])
satconc <- grm(dadc[c(32:37)])

#construct reliability
cronbach.alpha(dad[,1:6])
cronbach.alpha(dad[,7:13])
cronbach.alpha(dad[,14:19])
cronbach.alpha(dad[,20:25])
cronbach.alpha(dad[,26:31])
cronbach.alpha(dad[,32:37])

#amount of information
information(int2,c(-4,4))
information(col,c(-4,4))
information(qua,c(-4,4))
information(usa,c(-4,4))
information(val,c(-4,4))
information(satcon,c(-4,4))

#fit statistics with residuals in upper diagonal and p in lower
margins.grm(int)
margins.grm(col)
margins.grm(qua)
margins.grm(usa)
margins.grm(val)
margins.grm(satcon)

#Item Characteristic Curves, changing objects: "int", "col", "qua", "usa", "val" and "satcon"
#Item Information Curves, changing objects: "int", "col", "qua", "usa", "val" and "satcon"
par(mfrow = c(2,2))
plot(int, type = "IIC", legend = TRUE, cx = "topright", lwd = 2, cex = 1,main = "IIC with 890")
plot(intb, type = "ICC", legend = TRUE, cx = "topright", lwd = 2, cex = 1, main = "IIC with 100")
plot(intc, type = "IIC", legend = TRUE, cx = "topright", lwd = 2, cex = 1, main = "IIC with 50")
plot(int, type = "IIC", items = 0, lwd = 2)
info1 <- information(int2, c(-4, 0))
info2 <- information(int2, c(0, 4))
text(-1, 4, labels = paste("Total Information:", round(info1$InfoTotal, 3),
"\n\nInformation in (-4, 0):", round(info1$InfoRange, 3),
```

```
paste("(", round(100 * info1$PropRange, 2), "%)", sep = ""),
"\n\nInformation in (0, 4):", round(info2$InfoRange, 3),
paste("(", round(100 * info2$PropRange, 2), "%)", sep = "")), cex = 1)

#Pearson correlations
rcor.test(dad[,1:6],method="pearson")
rcor.test(dad[,7:13],method="pearson")
rcor.test(dad[,14:19],method="pearson")
rcor.test(dad[,20:25],method="pearson")
rcor.test(dad[,26:31],method="pearson")
rcor.test(dad[,32:37],method="pearson")

#factor scores for person
fint <- factor.scores(int)
colt <- factor.scores(col)
quat <- factor.scores(qua)
usat <- factor.scores(usa)
valt <- factor.scores(val)
satcont <- factor.scores(satcon)

#kernel density estimation for factor scores (KDE)
par(mfrow = c(2,3))
plot(fint, include.items = TRUE, main = "KDE for Interactivity")
plot(colt, include.items = TRUE, main = "KDE for Collaborative Learning")
plot(quat, include.items = TRUE, main = "KDE for Quality")
plot(usat, include.items = TRUE, main = "KDE for Usability")
plot(valt, include.items = TRUE, main = "KDE for Value")
plot(satcont, include.items = TRUE, main = "KDE for Satisfaction and Continuance")

#correlations panel
panel.cor <- function(x, y, digits = 2, prefix = "", cex.cor, ...)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(0, 1, 0, 1))
  r <- abs(cor(x, y))
  txt <- format(c(r, 0.123456789), digits = digits)[1]
  txt <- paste0(prefix, txt)
  if(missing(cex.cor)) cex.cor <- 0.8/strwidth(txt)
  text(0.5, 0.5, txt, cex = cex.cor * r)
}
pairs(dad[1:6], lower.panel = panel.smooth, upper.panel = panel.cor)
pairs(dad[7:13], lower.panel = panel.smooth, upper.panel = panel.cor)
pairs(dad[14:19], lower.panel = panel.smooth, upper.panel = panel.cor)
pairs(dad[20:25], lower.panel = panel.smooth, upper.panel = panel.cor)
pairs(dad[26:31], lower.panel = panel.smooth, upper.panel = panel.cor)
pairs(dad[32:37], lower.panel = panel.smooth, upper.panel = panel.cor)
panel.hist <- function(x, ...)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(usr[1:2], 0, 1.5) )
  h <- hist(x, plot = FALSE)
  breaks <- h$breaks; nB <- length(breaks)
  y <- h$counts; y <- y/max(y)
  rect(breaks[-nB], 0, breaks[-1], y, col = "cyan", ...)
}
pairs(dad[1:6],panel = panel.smooth, cex = .5, pch = 24, bg = "light blue",
     diag.panel = panel.hist, cex.labels = 1.5, font.labels = 1.5)
pairs(dad[7:13],panel = panel.smooth, cex = .5, pch = 24, bg = "light blue",
     diag.panel = panel.hist, cex.labels = 1.5, font.labels = 1.5)
pairs(dad[14:19],panel = panel.smooth, cex = .5, pch = 24, bg = "light blue",
```

```
    diag.panel = panel.hist, cex.labels = 1.5, font.labels = 1.5)
pairs(dad[20:25],panel = panel.smooth, cex = .5, pch = 24, bg = "light blue",
    diag.panel = panel.hist, cex.labels = 1.5, font.labels = 1.5)
pairs(dad[26:31],panel = panel.smooth, cex = .5, pch = 24, bg = "light blue",
    diag.panel = panel.hist, cex.labels = 1.5, font.labels = 1.5)
pairs(dad[32:37],panel = panel.smooth, cex = .5, pch = 24, bg = "light blue",
    diag.panel = panel.hist, cex.labels = 1.5, font.labels = 1.5)

#composite reliability
dadi <- dad[,32:37]
items <- paste(names(dadi),collapse = "+")
model <- paste("extraversion", items, sep = "=~")
fit <- cfa(model, data = dadi)
sl <- standardizedSolution(fit)
sl <- sl$est.std[sl$op == "=~"]
names(sl) <- names(dadi)
re <- 1-sl^2
sum(sl)^2 / (sum(sl)^2 + sum(re))
```

The supplementary documents to this thesis include a brief presentation on power point file, the research instrument, results of the confirmatory factor analysis, results of the structural model, search results in Google Trends and the executable code in R.

Supplementary documents to this thesis can be found at the link:
https://github.com/fernandopcmm/Thesis