Co-funded by the
Erasmus+ Programme
of the European Union

MOOC-Maker

Construction of Management Capacities of MOOCs in Higher Education

(561533-EPP-1-2015-1-ES-EPPKA2-CBHE-JP)

WPD1.6

Attrition and Retention Aspects in MOOC Environments

Version 5.0

(English)

August 20th, 2016

ABSTRACT

This study provides information concerning MOOCs. It discusses benefits of MOOCs, as well as issues which appear. The main goals are to raise awareness about the high dropout-rates and to come up with different models and solution to counteract this trend. Thus, it particularly addresses the alarming high dropout rates in MOOCs, which are often beyond 90% of all participants. For that, attrition and retention aspects will be discussed in-depth. Extended researches in this field of expertise will be presented as a result of a conducted literature review. The review will start by discussing the historical learning settings and then approach to the present, where finally MOOCs appear. The historical findings are applicable to some concerns faced in MOOCs as well and will reveal some improvements for MOOCs. Concerning aspects of attrition and retention, some analyses get performed, a proposal for an attrition model and a field study on attrition aspects will be presented. Furthermore, a survey has been conducted with MOOC participants, where several, different MOOCs were considered. As a result, a model has been created which groups the MOOC participants with respect to attrition and retention. A second survey with MOOC creators has been performed, which shall give insights from a MOOC creators' point of view. Both surveys will be thoroughly discussed and analysed in this work. The second survey is part of the best practices where also a meta analysis was conducted, which considers further literature, with special focus on sophisticated and profound best practices and recommendations. Finally, the report concludes with selected findings and recommendations, which were yielded throughout the report.

Keywords

massive open online course, MOOC, cMOOC, xMOOC, collaborative learning, cloud-based tools, dropout rate, funnel of participation, gamification on MOOCs, attrition, e-education, learning, learning strategies, learning engagement, online learning, learning analytics, motivation, retention, open education, web 2.0, cloud-based educational environment, predictive modelling

TABLE OF CONTENTS

I. INTRODUCTION

One primary measure about the quality in higher education has become the success of the involved students as well as the quality of students' online courses. The success of students can be defined in various ways, but commonly it gets strongly linked with the idea of achieving persistence to complete a selected student program. Therefore, one goal of educators and institutions is to establish courses with increased retention rates (Berge & Huang, 2004).

Within the last years massive open online courses (MOOCs) have become more and more popular. Students, educators, educational institutions and researchers are particularly interested in this new way of learning (Hernández, Gütl, & Amado-Salvatierra, 2014).

Originally, MOOCs were intended to provide introductory courses at universities for students all over the world. The first MOOC was designed by Stephen Downes and George Siemens at the University of Manitoba in Canada. In 2011 a MOOC with the title "Introduction to Artificial Intelligence" was offered by the Stanford University, where 160,000 participants enrolled. Now, online courses are provided in many different areas. The number of worldwide available MOOCs has grown steadily and in the middle of the year 2015 about 2400 MOOCs were expected worldwide. Figure 1: Growth of MOOCs (Sanchez-Gordon & Calle-Jimenez, 2015, p. 2) shall illustrate the worldwide, cumulative growth of the number of MOOCs over the years, within a timespan starting from 2012 and ending in mid 2015. As the number of provided MOOCs increased, also the number of students did who enrolled in an online course. Figure 2: Percentage of Students Enrolled in at least one online course (Sanchez-Gordon & Calle-Jimenez, 2015, p. 2) shows the percentage of students who enrolled in at least one online course, where a steadily growing trend can be observed (Sanchez-Gordon & Calle-Jimenez, 2015).

There are numerous advantages in using online courses, but also some potential drawbacks and issues. One particular problem is the unusual high dropout rate in various MOOCs. This issue will be particularly addressed in this work by conducting a broadly conceived attrition and retention analysis. Furthermore, several positive and negative aspects of MOOCs get discussed, where also possible solutions get described.

The following shall briefly outline the main chapters of this report:

Chapter "II. BACKGROUND" shortly describes the most important common terms, which show up throughout the report. Subsequently, a selection of various historical as well as present learning- and teaching-approaches is getting defined. The chapter finalizes with additional recommendations in literature for getting further insights concerning the discussed topics in this report.

Chapter "III. HISTORICAL RESEARCH ON ATTRITION AND RETENTION" primarily discusses historical learning settings, which have a strong influence on new learning methods in terms of improvements and the removal of some known drawbacks. It shows that issues like attrition or dropout were already concerns in most educational learning settings, which required respective measures.

With chapter “IV. ATTRITION AND RETENTION IN ONLINE-LEARNING AND MOOCs” the literature survey approaches to the present where MOOCs appear and aspects of attrition and retention get discussed in detail, by also considering findings of the previous chapter III.

Chapter “V. USER MODELS AND USER PREDICTIONS” presents practical results of experiments with a set of eleven MOOCs, offered by University of Galileo. It starts with a general overview of the MOOCs and with an analysis in terms of completers, non completers and dropout, which is the starting point for several classification experiments aiming to identify different classes of students. The chapter concludes with a discussion about feature selection and ranking.

Chapter “VI. BEST PRACTISES” introduces methods and profound suggestions for MOOCs, with a special focus on improving retention, which have proven over time and are well-established, derived from a conducted survey, from experiences and the elaborated literature. It comprises a performed meta-analysis from the literature and the survey was conducted with MOOC creators of the MOOC Maker Consortium.

Chapter “VII. RECOMMENDATIONS AND FINDINGS” describes selected findings and recommendations which result from all preceding chapters in order to improve MOOCs and counteract aspects of attrition and dropout.

Finally, chapter “VIII. SUMMARY” briefly summarizes and concludes the report.

II. BACKGROUND

In this chapter some important terms will be introduced among others. Despite the fact that some definitions widely vary in literature, commonly used clarifications will be utilized. Furthermore, various learning settings get listed, where some strongly overlap with others which makes a few of them difficult to clearly distinguish from each other. For that, some learning settings will rather be discussed coherently. Finally, the chapter concludes with selected and recommended works in literature.

Common Terms

There are some terms and properties which are relevant for many, different learning paradigms. First, those will be shortly presented:

Berge and Huang (2004) and Angelino, Williams and Natvig (2007) declare **persistence** as the result of a student's decision to continue the participation in a specific learning program, meaning it refers to the act of continuing a program towards an educational goal, like earning a degree or certificate.

According to Berge and Huang (2004) the definition of the term **retention** is rather complex and problematic due to the fact that different researches often contain inconclusive or contradictory results. It is said that retention studies usually cover the degree of completion and non-completion. To be more precise this only holds for some specific classes of students. In other cases it can be more important that the objectives of a participant get reached. This report will follow the definition of Berge and Huang (2004): **Retention** is the continued student participation in a learning program to completion. It is measured by the number of students that progress from one level to the next in a specific program until either completion or the student's personal goals are met. Further, they declare **attrition** as a decline in the number of students from the beginning to the end of a specific learning program (Angelino et al., 2007; Berge & Huang, 2004). Attrition can be seen as the observed event when a student “drops” from a learning program. Furthermore, attrition can be interpreted as a measure to determine the quality of education delivered by an institution: In case of a high attrition rate, the perception is that the institution is facing some problems in quality (as cited in Angelino, Williams, & Natvig, 2007).

Selection of Learning- and Teaching-Approaches

In order to give a small overview, some of the most important learning settings will be listed, followed by a brief definition. These definitions particularly apply to this report and may partially vary in literature:

Brick and mortar learning means that students get taught face-to-face in a traditional classroom setting (Harris & Nikitenko, 2014).

The term **in-person** is also commonly used to indicate traditional, brick and mortar, place-based education (Berge & Huang, 2004).

One of the most accepted descriptions in literature for **self-directed learning** is the definition of Knowles (1975), who described it as the process in which

> *individuals take the initiative, with or without the help of others, in diagnosing their learning needs, formulating learning goals, identifying human and material resources for learning, choosing and implementing appropriate learning strategies, and evaluating learning outcomes*.

Self-regulated learning (SRL) is sometimes used in the same meaning and context as the terms **autonomous learning**, **self-planned learning**, **self teaching** and **independent study**. Their subtle differences are rather inconsistently used, causing that they are used interchangeably by many researchers. Quite often **self-regulated learning** and **self-directed learning (SDL)** are also not clearly distinguished in literature (Saks & Leijen, 2014). Saks and Leijen (2014) conducted a study in order to distinguish the two latter ones. Firstly, they agree with the definition of **self-regulated learning**, as stated before, which was defined by Knowles (1975). For not overemphasising its distinction and to overcome its confusion, some researchers introduced a model where both terms are connected with the umbrella concept **self-direction in learning**. For the sake of completeness, also the description of **self-regulated learning** mentioned by Saks and Leijen (2014) will be listed:

> ***Self-regulated learning*** *is an active, constructive process whereby learners set goals for their learning and attempt to monitor, regulate and control their cognition, motivation, and behaviour, guided and constrained by their goals and contextual features on the environment*. (p. 191)

Distance learning and **distance education** refer to the educational principle where students are not physically present at university and can study at home. In this setting assisting technologies are necessary elements, which can comprise e.g. written material, radio, television, CD-ROMs, e-Mail, audio/video-conferencing, the Internet or any other technologies. Commonly an instructor plays an important role, delegating specific tasks the students need to perform. Since the technological capabilities continuously improved over the years so did the definition of distance education, whereas now also e.g. MOOCs and others could fall into this category of learning (Bozkurt et al., 2015), but this report will repudiate from this extended definition in order to define a more distinguishable margin between those learning settings. Here, **distance education** will be declared as the traditional learning setting where a student can learn at home by receiving the relevant information and material via a postal way or other means. The instructor is responsible to prepare the corresponding tasks and relevant learning material. As Moore, Dickson-Deane and Galyen (2011) put it, **distance education** is often described as the "*effort of providing access to learning for those who are geographically distant*" (p. 1). As a side-note it shall be mentioned that **distance education** could be seen as an umbrella term, comprising different stages of learning which evolved over time. This idea is also supported by Keegan (1996). Angelino et al. (2007) mentioned a definition of distance education which included educational or

training courses delivered to remote (off-campus) sites via audio, video (live or pre-recorded) and other means including synchronous and asynchronous instruction.

As the name already suggests, **computer-based learning (CBL)** uses a computer for instructional purposes, whereas the computer hard- and software, including all its peripherals like the input devices, are the main parts of the learning environment (Seel, 2012).

Web-based learning (WBL) encompasses all educational interventions which are actively using the Internet. Some further classifications can e.g. be tutorials and online discussion groups (Cook, 2007).

Online learning or **online education** means distance learning that is delivered using computer- or web-based systems (Berge & Huang, 2004).

Blended learning means a combination of in-person and online learning. It combines the best parts of online and face-to-face learning. It can also be called **hybrid learning** (Berge & Huang, 2004; Watson, 2008).

The **Flipped Classroom** represents a pedagogical approach, which uses asynchronous video lectures and practical problems as a homework and active, group-based activities for solving problems in the classroom, which results in a unique combination of learning theories: Active, problem-based learning tasks follow the constructivist ideology and instructional lectures are based on behaviourist principles (Lowell Bishop & Verleger, 2013).

In today's context **e-learning** mainly comprises online courses like web-based courses and learn management systems (LMS), which organize and deliver online courses (Downes, 2005).

Massive Open Online Courses (MOOCs) are a rather new development in the field of e-learning and distant education, which gain more and more popularity. Compared to traditional courses MOOCs are larger in scale, are globally accessible and have no restriction on participation (Adamopoulos, 2013).

In literature two different MOOCs often get distinguished: **cMOOCs** are based on "connectivism and networking", which uses connections, social networking and interaction in order to emphasise knowledge creation, creativity and autonomy. cMOOCs can also be called *network-based* MOOCs. **xMOOCs** are based on the "behaviorist" approach, focussing more on a traditional learning approach with content acquisition by using video presentations, short quizzes and testing, resulting rather in duplicating present knowledge. xMOOCs can also be called *content-based* MOOCs. *Task-based* MOOCs focus on the tasks the students are required to complete in order to acquire the necessary skills, meaning the learners are asked to complete certain types of work like e.g. assignments. These task-based MOOCs can be seen as a mixture of instructivism and constructivism (Gütl, Rizzardini, Chang, & Morales, 2014; Pisutova, 2012).

Small Private Online Courses (SPOCs) are a derivate of MOOCs, which are locally constrained, meaning the number of students is limited to on-campus students. The word was coined by professor Armando Fox in 2013. His idea was to use MOOCs as a supplement to classroom teaching, rather than considering them as a replacement. In

this way they can improve the instructor leverage, student throughput as well as the student engagement Fox (2013).

Besides those mentioned learning settings also others exist in literature.

Improvements of learning settings and new learning paradigms strongly evolved from the past. Due to the fact that they got especially influenced by previously well-established learning settings this report will discuss some of the most important ones in history in chapter “III. HISTORICAL RESEARCH ON ATTRITION AND RETENTION” of this report.

Recommendations in Literature

Tinto (1975, 1987) examined the phenomenon of student retention in traditional education and introduced a student integration model for college dropout, where he explained the process why students leave colleges and universities prematurely. Bean (1985) introduced a model for students' attrition and their intention of leaving an institution, considering different aspects like academic, social and personal factors. Avakian et al. (1982) studied student retention in terms of differences in race and sex at an urban university. Boudreau and Kromrey (1994) performed a longitudinal-study about retention, completion and academic performance of participants in a freshmen orientation course. Moore and Miller (1996) elaborated how using multimedia affects student retention and learning. Bedford and Durkee (1989) and Hyman (1995) described strategies helpful for retaining students at colleges and universities (Adamopoulos, 2013).

Carr (2000) was repeatedly cited with the “anecdotal evidence” that online courses resulted in higher attrition rates than traditional learning settings. According to this topic she elaborated challenges and suggestions in this field.

Moody (2004) heavily discussed aspects about the high attrition rates in distance education and their causes.

Hara and Kling (2000) presented insights for students' distress with web-based distance education and their work shall enhance the understanding of the instructional design issues as well as other issues in order to improve web-based distance education.

Serwatka (2005) addressed different ways how to improve retention in distance learning classes.

Angelino et al. (2007) referenced many noteworthy works in their paper and discussed various strategies to engage online students and to reduce attrition rates.

Willging and Johnson (2004) discussed different factors that influence students' decision to drop out of online courses, where they used logistic regression analysis to compare various factors of persistent learners and those who dropped out over time. One interesting result was that demographic variables do not predict the likelihood for dropping out of an online course.

Downes (2008) introduced the online course “Connectivision & Connective Knowledge”, which can be seen as the first established MOOC. From a historical point of view, this work gets recommended.

Moskal, Dziuban and Hartman (2013) critically discussed the blended learning setting, and could observe very good completion results for blended learning at the University of Central Florida, by means of an iterative loop of continuous quality improvements.

Cormier and Siemens (2010) and Masters (2011) explored the different dimensions of MOOCs, discussing the changes in the roles of a professor or instructor, where it is now more important to amplify, curate, aggregate, filter or select, model and stay present. Further Cormier and Siemens (2010) and Russell et al. (2013) discussed the retention problem in more detail and showed some future perspectives of MOOCs. Xu and Jaggars (2013) compared in which extent students perform in an online setting as well as in a face-to-face setting. Mackness, Mak and Williams (2010) stated ideas for designing MOOCs where participants collect positive experiences. Mak et al. (2010) analysed the blogs and discussion forums as a means of communication, as well as online learning tools in MOOCs (Adamopoulos, 2013).

Pappano (2012) is an often cited article discussing several aspects of MOOCs, giving more general insights and showing the interest on MOOCs as well its potential.

Hill (2012) tried to present an overall picture of online educational delivery models and nicely combined some of them in pictures.

Gené, Núñez and Blanco (2014) discussed aspects of integrating gamification elements into MOOCs. A model to motivate the students gets proposed, which is based on the most attractive and addictive elements of games in environments which are not for play and aim to get the students to complete the course.

Jordan (2015) covers MOOC completion rates by revisiting aspects like assessment, length and attrition in depth. The work is based on data which was collected of 221 MOOCs.

Ulrich and Nedelcu (2015) discussed hopes and worries concerning MOOCs provided at the University of Bucharest. It investigates perceptions and expectations linked with MOOCs, considering the students' and the faculty's perspective, where interviews, questionnaires and content analyses were conducted.

Pouezevara and Horn (2016) explored the potential for the international, educational development of MOOCs and online education. Furthermore, it compares typical MOOCs and online learning systems in detail.

It shall be noted that some findings in the listed recommendations resulted in different or even contradicting outcomes (Adamopoulos, 2013).

III. HISTORICAL RESEARCH ON ATTRITION AND RETENTION

Some of the most important learning settings from the past to the present have been listed in the previous chapter “II. BACKGROUND”, by giving a short definition of each. This chapter aims to elaborate them in more detail, primarily focussing on historical settings. The reason is that all previously well-established learning settings had a strong influence on a new learning setting in terms of improvements and the removal of some known drawbacks. This section shall raise awareness concerning different learning settings. Another important aspect of this section is the following observation: All different learning settings share some commonalities. In general, they want to introduce a good way for students to efficiently learn new concepts covering plenty different topics. Despite other similarities many of them had to deal with the same issues: Ever since, aspects like retention, attrition and persistence were hot topics because each learning setting had a particularly high dropout rate. Thus, this section will address different learning schemes, trying to find correlations with similar observations and phenomenons also found in present settings, like MOOCs. This section will supply information on different perspectives from the past and present prior challenges, improvements and ideas.

The problem of students' retention is not due to an isolated factor which could be fixed easily but several aspects must get addressed by carefully selecting interventions. Retention of students in courses, programs or degree levels has been a timeless concern of educators. The lack of retention, also called dropout, of students in educational systems was always a historical challenge: For the past 100 years the dropout rate in brick and mortar higher education was constantly between 40 to 50%. (Berge & Huang, 2004). Due to that, this survey firstly starts with a brief historical research on attrition and retention, which shall be useful to discuss and master the problems faced at present.

As it could be seen in the previous chapter “II. BACKGROUND”, some learning settings are rather difficult to distinguish from each other and their definitions vary widely in literature. Thus, this report will introduce a more high-level view, where the particular learning settings are comprised. Focussing on the university domain and sorted by its manifestation, the high-level view will look as follows:

- Traditional face-to-face learning
- Distance education
- Computer-based learning
- Web-based learning
- Blended learning

- MOOCs and online learning

Traditional Face-to-face Learning

The problem of student retention was already an issue in traditional education: According to McNeely (1938), the first national retention study in the United States revealed a dropout rate of 45%. In this study 25 universities were involved.

Gütl, Rizzardini, et al. (2014) also mentioned dropout rates between 40 and 50% in the educational setting comprising physical presence.

Tinto (1982) has stated, that the percentage of students dropping out from traditional courses were constantly between 40 to 50% over the last century (McMahon, 2013).

McMahon (2013) reported an attrition rate of 70% among adult learners from the year 2002, whereas a more recent study from 2007 reported non-completion rates of 16% in the traditional learning mode.

Adamopoulos (2013) wrote that various methods had been deployed in order to study retention on traditional education. Methods like simple cross-tabulations, two-sample comparisons, linear regression, logistic regression, prohibit analysis as well as Markov processes had been employed. Further, he mentioned that in some prior works the applied methods for analysing the issue of student retention were criticized concerning consistency, efficacy and suitability, which might show up reasons for the conflicting findings reported in literature. In his work, he gave examples which led to erroneous findings. Additionally, he proposed suggestions in order to minimize erroneous results which he applied to his research. Those methods were elaborated in more detail in his work, which will not be discussed in this report. Adamopoulos (2013) referred to qualitative studies from the past which suggested that students prefer taking difficult courses in a traditional setting rather than an online setting (Jaggars, 2012). He stated, that his work was also in accordance to prior research concerning difficulty, workload and team projects on course retention in traditional education.

Adamopoulos (2013) referenced other works like (Tinto, 1975) and others, which showed that social life has a significant, positive impact on the institutional fit for each traditional class. Furthermore, it could be observed that students affect other students' attitudes much greater than faculty members do. He recommended studying the interactions between individual, academic and social systems. Further, he rounded off his suggestions by stating that researches should have a closer look on individual goals of students in terms of e.g. the importance of completing a course, attributes like abilities, educational background, academic performance, family background, faculty contact and interaction and other environmental factors like for instance finances.

In traditional face-to-face classes some students are prevented from pursuing a degree because of physical and temporal barriers (Gleason, 2004).

Diaz reported a dropout rate of 7.2 percent for traditional students. Murray reported that the Washington State Community College online program claimed a retention rate of 85 percent for traditional students. Kalher found for the Brevard Community College completion rates of 84.7 percent. According to Murray and Carr, there were no standardized national retention rates available those time (as cited in Gleason, 2004, p.

395) whereby anecdotal evidence suggested retention rates in face-to-face classes well above 60 percent (Gleason, 2004). Since the retention rates for traditional classes vary by intuition, Murray stated that many people disagree with the validity of the corresponding statistics. It was mentioned that some institutions do not count students who drop out during the first few weeks of the semester, whereas others do. Due to that it was said that considering the retention rate might be more useful at a local level, instead of a national one because of many influential factors (as cited in Gleason, 2004, p. 395).

Tinto (1975) suggested that retention in traditional education, where not only a single course but several courses are conducted, can be improved by increasingly integrating students into the academic and social system of a college.

Gleason (2004) noted that online-learning is not for everyone, meaning that some students do not have the motivation or self-discipline to complete a corresponding learning setting. Others also simply prefer traditional classroom environments.

One mentioned favoured suggestion to increase retention was some sort of student orientation. Over time it has proven successfully to offer orientation sessions at traditional colleges in order to make a fruitful transition to college. On the other side it has been exposed over time that there were also many issues involved in designing an orientation seminar with respect to decisions whether it is e.g. supposed to be optional or mandatory, about the format, as well as its length. Positive effects from those seminars were that students were able to meet each other in person. Further, they were able to get an impression of the faculty and got a better idea of the provided course. Thus, they were able to overcome their fears and concerns about the program and could develop relationships (Gleason, 2004).

Distance Education

Distance education already existed almost two centuries ago. Time after time, it steadily evolved. Meanwhile, it has undergone drastic changes in terms of learning and communication. In the beginning postal services delivered the relevant material. Now uncountably many tools are available through the Internet and new, faster communications arose as the time passed. History has shown that authors and researchers used inconsistent definitions for distance education or distance learning, which makes it difficult to compare different studies or to build upon results from previous ones (J. L. Moore et al., 2011). This section shall mainly focus on the setting where the relevant learning material was sent via a postal way or radio and television systems were used to broadcast instructions from an educator.

One general disadvantage is that people are facing issues of isolation and disconnect in distance education (Gütl, Rizzardini, et al., 2014).

Due to the fact, that there are no national statistics for completion rates of students in the distance education setting, it is believed that the dropout rate is between 10 to 20 percent higher than for the in-person learning setting (as cited in Berge & Huang, 2004).

As attrition is a major issue in higher education, it also affected distance education. According to Angelino, Williams and Natvig (2007), attrition rates for classes taught through distance education were 10 to 20 % higher than classes taught in a face-to-face

setting. Due to that, educators should try to engage students early and often by using various learning strategies which should be customized to the content of the class as well as the students' existing pre-knowledge. Angelino et al. (2007) suggested that the goal of a professor should be to develop relationships with the students so that they feel comfortable in their environment. Furthermore, they suggested that a professor should facilitate the learner-learner integration and collaboration, which would result in learning from each other and expanding their knowledge base together. In 2003 national studies were conducted by the Nation Center for Education Statistics (NCES), which showed that more and more institutions started offering distance education courses. In the same way, the number of students who enrolled in distance education courses was increasing. Although more students were choosing this educational setting, it could be observed that attrition rates were commonly higher than in face-to-face courses. It was mentioned that finding ways to decrease attrition in distance education are important for both sides, from the economical and the qualitative viewpoint. High attrition rates result in a negative economic impact for universities: The development, delivery and assessment of the courses costs money and the lost tuition revenue results in wasted expenditures. Thus, it is critically necessary to identify the reasons students drop or fail courses in determining which services and delivery methods an institution would need to provide in order to guarantee successful completion in distance education programs (Angelino et al., 2007).

Nash (2005) tried to discover reasons why students dropped or failed distance education courses. The study revealed that the number one issue was time management. "*Students either tried to accomplish too much in one semester or they had difficulty managing their time*" (p. 5). Other mentioned issues were that some course assignments were too difficult, the directions of the assignments were not clear or students were not able to get help when they needed it (Angelino et al., 2007).

Martinez (2003) mentioned that in case of attrition there were often personal reasons involved such as family problems, finances, child care, distractions or job needs. Moody (2004) further stated that students may underestimate a course and that it would be harder than originally estimated.

Hara and Kling (2000) stated that students often reported confusion, anxiety and frustration caused by unclear feedback or ambiguous instructions from educators. The feeling of social isolation also showed up repeatedly.

Computer-based Learning

As the name "computer-based learning" suggests, this learning method uses various aspects of computer technology to support individuals in learning for different educational purposes. With computers it is possible to represent information in several ways, like texts, diagrams, graphs, audio and video files, to name but a few. Furthermore, computers enable interlinking different and related sources or representations in a way never seen before, which can also be called "hypermedia content". Additionally, computers also allow interactive representations which can be directly manipulated by the user. This gives the learners the possibility to explore data in a way they prefer, which is a learning-by-doing approach or an active learning strategy and an important way for fostering learning. The manipulation of data representations can be called simulations or microworlds (Winters, Greene, & Costich, 2008).

Computer-based learning environments (CBLE) enabled the learner a higher degree of learning control as well as opportunities for self-directed learning. However, empirical studies showed that students had troubles with CBLEs. Besides the advantages mentioned before, learners often need to determine the most helpful representation of data. The learners' decision making is then based on their self-knowledge, experiences and beliefs, as well as motivational factors, prior knowledge, task definitions, goals and strategic knowledge, which partly defines a self-regulated learning approach. Thus, researchers started trying to better understand the specific processes of self-regulated learning in order to better integrate self-regulated learning approaches into CBLEs (Winters et al., 2008). Winters et al. (2008) concluded with finding evidence that academically successful students tend to use more effective learning strategies in CBLE and that students with much prior knowledge tend to engage in more cases of planning and monitoring. Self-regulated students were performing well when provided with a high amount of learner control. Additionally, it could be observed that collaborative working students may provide regulation for each other. In terms of computer-based learning, tutors could be seen very supportive for students, as they outline particular strategies and hints before the students engage in the assignment.

When considering self-regulation with computer-based learning, it encompasses many aspects of traditional learning settings. Those aspects are for example the isolation of the students and the requirement of a cyclical, recursive and active process with motivation, behaviour and context. Winters et al. (2008) suggested that a think-aloud methodology is a potentially more accurate approach for checking the students' understanding than questionnaires or interview methods. For that, it is not sufficient that the students should just verbalize their thoughts, but try to explain their cognitions.

In contrast, Pane, Corbett and John (1996) were assessing the dynamics of computer-based instructions and evaluated a multimedia educational software system which included text, graphics, animations and simulations. When they compared this system with an informational equivalent environment which only used text as well as carefully selected still images, they only found little evidence that dynamic presentations enhanced the students' understanding of the information covered in this lesson. They argued, that it cannot be expected from students to take full advantage of the exploratory possibilities, which are offered by computer-based instructions. Furthermore, usability problems of the supplied simulations had negative impact on the performance of students, which suggests that designers of educational software should consider usability aspects in a higher extent. As the mere usage of animations and simulations did not guarantee students' learning, they suggested to also use well designed static graphics and text, which might be as effective as well as much cheaper to use and to produce. On the other hand, movies and simulations could be seen more motivating, which would subsequently result in spending more time on the educational content, even outside a provided laboratory setting, which would imply learning more.

Welsh, Wanberg, Brown and Simmering (2003) referenced other works which found completion rates of 95% in a traditional classroom setting, but only a completion rate of 64% for those attending the course via a computer. The coherent completion rates for a reported career advancement course resulted in a less drastic difference, between these two learning settings, namely a completion rate of 95% in the traditional classroom setting and a rate of 90% in terms of an equivalent computer-based course. Further, the observation showed up, that when the learners have "*a clear rationale for completing a*

course, learners generally seem to do so" (Welsh et al., 2003, p. 254). This could be confirmed with a study conducted by Wisher and Priest, where no difference between instructor-led and technology-delivered training could be found, where both groups had a completion rate of 100%. One possible explanation for a completion rate of 100% for both groups was that the learners participated in a course which enabled the transition to a new, promising career. This shows that motivation is one primary factor for achieving persistence (as cited in Welsh et al., 2003). Welsh et al. (2003) further stated that computer-based learning can result in positive experience, however, in case the used technologies do not work well, the potential is growing rapidly to turn out in a very negative experience. One mentioned example in this context was that teachers were very frustrated with CD-ROM training when the training was facing technical issues. In case the problems could not be solved within about ten minutes, the teachers "*quit with extremely strong feelings of frustration or anger*" (as cited in Welsh et al., 2003, p. 255). Furthermore, in case the technology did not work well the first time a person used it, the person would most probably turn resistant and takes another course instead.

Web-based Learning

Web-based learning (WBL) can be seen as the next step after computer-based learning in the evolution of learning settings. This learning paradigm overcame barriers of distance and time and introduced novel instruction methods. Negative consequences included social isolation, upfront costs and technical issues. Web-based learning can be used to deploy individualised instruction (Cook, 2007). Cook (2007) argued that many WBL designs were failing in the desired principle of learning efficiently. He further stated that WBL was often just used for the sake of technological progress instead of enhancing the learning experience. He suggested that the WBL approach should be accepted as a potentially powerful instructional tool and that it is more important to focus on learning, considering when and how it is useful to apply WBL, instead of trying to evaluate whether WBL outperforms other instructional media. Manifestations of WBL can for example be online tutorials or discussion groups, enhanced with features like multimedia, links to online resources and other content within the course, as well as self-assessment tools. In WBL teachers can often be seen in the role of facilitators, who define the scope of a discussion, monitor and guide the discussion and provide or help students finding complementary course material. The communication between course members can either be synchronously or asynchronously. On important aspect is to properly design WBL.

One of the most obvious advantages is the fact, that WBL overcomes physical distances, which can be seen as the main feature separating WBL from other computer-based learning methods. From another perspective resources can be shared among different universities, which can reduce redundancy in developing course material. Other advantages of WBL are flexible scheduling in timing of participation, resources can be easily updated and learning can be individualised by giving the learners more control in selecting from multiple learning opportunities within a course. Students who struggle with some content can use provided supplementary content and more experienced students can quickly move on to other chapters. It also allows novel instructional methods: As the learners have access to the Internet anyway, they can additionally use the web for further research. The learners can also perform learning exercises with interactive elements as well as various games. As in the previous section "Computer-based Learning", various multimedia elements can be used to enrich a course in a proper

way, which would be rather difficult by using a plain textbook. The asynchronous communication system allows learners to first think deeply about given issues and then answer in a thoughtful and proper way. This is very contrary to face-to-face settings where a discussion could continue before some learners had time to reflect about it and then participate in the debate. Furthermore, WBL enables online assessments which allow immediate and customized feedback (Cook, 2007).

Besides all positive aspects of WBL, there are also many disadvantages. In fact, many of those are interlinked with the previously mentioned advantages. One very big disadvantage is the perception of social isolation, which comes along with flexibility in time and location. This means, that it was quite common studying alone in this setting. In an era where team work and team learning is getting increasingly valued, this can be seen as some step backwards. A question which showed up in this context was, whether WBL supports or reduces critical interpersonal relationships and communication skills. Another disadvantage is the observation of de-individualised instruction: Besides the promising individualised instruction, as mentioned before in the advantages, many WBL systems failed to offer the individual needs of participants. It was mentioned that in those time adaptive WBL appeared to be more a vision than a proper reality and that the instruction was rather predetermined than personalised. Another disadvantage shows up by considering the costs: Developing WBL environments, considering aspects of effective instruction, can be very expensive. By using the technological possibilities for web-based learning environments, several technical problems can arise, which may have the potential to completely disrupt a WBL course. Even minor problems may result in serious impediment, which can decrease the participants' satisfaction, as well as the course participation. The instructional design is another very important criterion for a successful WBL course. If a poor quality in the design is supplied, acceptance will decrease. In contrast to face-to-face courses, where a potential talented instructor can teach with minimal effort in preparation, the instruction in WBL must be explicitly planned and created. But as research has shown, most WBL courses had an inferior instructional design. In some cases, it could be observed that technology in courses were only used for technology's sake, meaning the instructors just want to keep up to date, less considering achieving an educational goal, which has often led to poor instructional designs (Cook, 2007). Another aspect which Cook (2007) mentioned was the statement, that on the one hand well-designed graphics and animation improve the way of learning, however, non-essential multimedia content can result in the opposite and may distract learners as well as decrease learning. Additionally, he mentioned that the frame rate of video clips and their colour did not seem to impact satisfaction or understanding of the learners, which means that high visual fidelity is no vital necessity. Finally, he suggested that educators shall not replicate a face-to-face course or written material on the web. Converting a course to the WBL format requires substantial revision in order to minimize the stated disadvantages of WBL and to unveil its full potential (Cook, 2007).

Carr (2000) stated that attrition rates in undergraduate distance education programs, which encompassed web-based teaching, were between 10 and 20% higher than in face-to-face settings. This information is based on "*anecdotal evidence*" (p. A39). Furthermore, it was stated that comparing attrition rates between institutions may be misleading because the measurements of retention rates vary widely.

Terry (2001) evaluated the graduate level business courses from Texas and tried to determine the attrition rates for online and face-to-face courses. He concluded that there existed some business disciplines like accounting, economics, computer

information systems, marketing and management which had online attrition rates comparable to face-to-face counterparts. One key finding was that online courses concerning business statistics and finance resulted in attrition rates between 33 and 38%, while face-to-face settings ranged between 13 and 23%.

Gütl et al. (2014) also stated that the online dropout rates may be 10 to 20% higher, compared to the higher educational setting with physical presence.

McMahon (2013) mentioned a study by Meister (2002) which reported non-completion rates of 21% in online classes in the context of e-learning. Furthermore, McMahon (2013) referenced a study of Masters students from the US, where dropout rates for online courses were between four and six times higher than those in traditional learning settings. Others further reported varying attrition rates from 18 to 80%. Despite this varying observations, the consensus in literature is "*that the attrition levels among online programmes are significant*" (McMahon, 2013, p. 4).

Angelino et al. (2007) introduced and discussed several strategies to reduce attrition in the context of distance education, augmented with web-based technologies. They extracted four major strategies to counter this ongoing trend:

- Student integration and engagement,
- Learner-centered approaches,
- Learning-communities and
- Accessibility to online student services.

In the following the found strategies of Angelino et al. (2007) will only be outlined briefly. For more details the original work gets recommended:

Tinto (1975) already suggested integrating students into the university setting in an academical as well as social way, stating that student engagement and integration are key elements to student persistence. It was suggested that student integration and engagement can for example be fostered by initiating contact with students via phone calls or other technical means. Other suggestions are to conduct a pre-course orientation or to facilitate informal online charts through the course website. In order to improve the learner-centered approach, suggestions were to stimulate the communication between students and educators from the very beginning of a course, which engages the students in a better way. Furthermore, this gives students a possibility for shaping instructional content. This also enables teachers getting to know the pre-knowledge and experiences of the participating students. Additionally, the students shall also get to know each other. Students could for instance post their introduction or their expectations about the course online. Chats were mentioned being a good means for "breaking the ice" between students themselves, as well as between students and teachers. The third vital idea was to establish learning communities, which may help students in overcoming physical separation, the feeling of isolation, the lack of support, as well as the feeling of being disconnected. By conducting group projects and assignments students will engage in the learning process and might develop a sense of community, which counteracts and reduces high attrition levels. The last point which will be outlined is the suggestion of online student services, which focus on the learners'

needs. Among others, those services shall encompass online access to assessments, educational counselling, registration, technical support, study skills assistance, career counselling, library services, as well as students' rights and responsibilities (Angelino et al., 2007).

In general, in order to increase persistence in this educational setting, "*educators need to engage students early and often*" (Angelino et al., 2007, p. 8). Students should feel comfortable in the educational setting and have the opportunity to develop relationships with colleagues. It was stated, that "*distance learners typically like to work independently [from each other] but will participate in collaborative work if initiated by their instructor*" (Angelino et al., 2007, p. 8). This is why educators should design courses that engage students and enable learner-centered approaches.

Blended Learning

Watson (2008) explains blended learning as the convergence of online and face-to-face education. It can be seen as a way of enriching the content taught in class as well as to extend learning beyond the walls of an educational institution. Commonly it gets affiliated to computer-based instructional material, whereas the upcoming of the Internet enabled further possibilities, which greatly increased the quality of digital classroom resources. The Internet especially fostered blending online learning and face-to-face instructions. Watson (2008) stated that blended learning was likely to become the predominant model of the future and that it could turn far more important as both learning approaches considered distinctively. He declared that there is no single type of blended education and that over time one can expect that all the different approaches, from fully online to fully face-to-face, would get filled. As online teaching is different to face-to-face teaching, blended learning is also unique and requires tailored instruction methods and course design principles. This for instance means that content from face-to-face teaching cannot just be used for online programs and vice versa. Another challenge in this mixed approach is that specific assessments need to be designed in order to properly test the presented content. An additionally mentioned development was that text-based content would be less effective than multimedia content, like animations, videos and simulations. Teachers would need to access online content in a fast way, so that the classroom instruction flow keeps moving. Further it was stated that blended learning requires some course management system or a learning management system (LMS) in order to organize its content and the needed communication. Finally, he concluded that blended learning might face some challenges for research and policy because it can vary in so many ways. As an outlook he mentioned the Science Leadership Academy, which used a project-based approach to achieve five primary goals: inquiry, research, collaboration, presentation and reflection. Herein, qualified teachers created a new learning environment within a traditional brick and mortar environment: All students had notebooks and could access a course management system and a program for social networking for better collaboration. The students performed researches followed by subsequent presentations of their projects in the classroom, as well as online. Watson (2008) suggested that in this new environmental teaching setting teachers need to approach their role differently. Their role should rather be seen as guides and mentors. Additionally, he wrote that classrooms must be transformed to flexible learning environments, where learning should also be possible outside their own country. Chris Lehmann, principal and co-founder of the Science Leadership Academy stated, that blended learning is about a new paradigm for students and teachers which is

no longer constrained to four walls. Nicholas Negroponte, founder and chairman of the One Laptop per Child association, said that true personalization in learning is a vital requirement (Watson, 2008).

Moskal, Dziuban and Hartman (2013) conduced a survey comprising blended learning, fully online and face-to-face settings, as well as other learning settings with the University of Central Florida (UCF). First, a course rating was presented, where over one million students responded for the academic years 2008 to 2011, which were indexed by the corresponding course modality. It turned out that blended learning enjoyed the highest percentage (52%) of "excellent" responses in terms of students' satisfaction. This learning modality was followed by 48% in terms of a fully online course. The face-to-face setting also achieved 48% of "excellent" responses. They concluded, that irrespective of the course mode in which students used to learn, in case the educator facilitated learning, communicated well and respected the students, then the students would rate the course excellently to a high degree. Further, statistical data about the success rates was presented: Blended learning had a success-rate of 90%, the fully online program achieved 88.3% and the face-to-face setting reached 78.7%. Dropout rates were stated with 2.8% for the blended approach, 3.1% for the face-to-face setting and 4.3% for the fully online course. These achievements required alignment of institutional, faculty and student goals. Additionally, a reliable and robust infrastructure was required in order to support the students as well as the institution. It was also mentioned, that "*continuous evaluation can effectively track the impact of blended learning on students, faculty and the institution*" (Moskal et al., 2013, p. 15), which could be used to improve and develop corresponding, supportive structures, enhancing the success rates of the participants. An iterative loop of continuous quality improvement also helped to achieve these results (Moskal et al., 2013).

MOOCs and Online Courses

While MOOCs can be seen as a quite recent development, the original ideas can be tracked back to the early 1960s, where an industrial-scaled, educational technology was proposed (Adamopoulos, 2013).

As it could be seen in chapter "III. HISTORICAL RESEARCH ON ATTRITION AND RETENTION", MOOC-courses evolved from a sequence of different learning settings, so did the meaning of the learning approach "distance education" because, many researches augmented the meaning of distance education over time, when new communication technologies showed up. The appearance of the Internet put the main foundation so that a new model dubbed as Massive Open Online Courses (MOOCs) could emerge. In 2008 the term "MOOC" was coined, which was used to describe an open online course, offered by the University of Manitoba in Canada. From then on, a range of different topics and platforms emerged and the term "MOOC" was declared as "the education buzzword of 2012" by Daniel (2012), which was intended to reflect the widespread interest in this concept (Liyanagunawardena, Adams, & Williams, 2013).

From the beginning of computing, academics were always sharing digital information. A few years ago sharing open educational resources (OER) became more and more interesting and turned to a relevant source of information for students as well as educators. In 2011, the Massachusetts Institute of Technology (MIT) introduced OpenCourseWare (OCW), where the material of the courses was permanently published on the open web, where different licences were defined for its usage. Many other

universities then started following this approach, like the Open University in UK. This movement was beneficial for students, who could use the content for learning, as well as educators, using the material for their own courses. A significant amount of open educational resources were however limited in use, since they were designed being a particular portion within a larger educational course. When ambitious learners used those portions for their own learning requirements, some became frustrated because the partial content was extracted from another overall scope. Similar problems could also be observed with educators aiming to natively reuse parts of the content. With the rising of MOOCs, the concept of open access was pushed into a rather different direction. MOOCs connect learners and experts who want to foster learning by publishing freely available online content. Commonly, social networking is used as a means of establishing connectivity. In general, MOOCs are designed in a way so that there are no prerequisites, fees, formal accreditation or the requirement of a particular degree of participation. They shall enable attendants a completely voluntarily participation, depending on their actual interests. The online course "Connectivism and Connective Knowledge", which was offered at University of Manitoba in 2008 and was supported by George Siemens and Stephen Downs, can be seen as the first MOOC. This course was using the approach of connectivisim. Later on, international co-operative partnerships with many universities were created, such as Coursera[1], edX[2] and Udacity[3] (Downes, 2008; Liyanagunawardena et al., 2013).

By studying several works in literature, McMahon (2013) concluded that there is the general opinion that the attrition rate is higher in courses delivered online and that researchers reported varying levels of attrition among online learners.

MOOCs and online courses can be seen as the latest evolution in the different learning settings. In the next chapter "IV. ATTRITION AND RETENTION IN ONLINE-LEARNING AND MOOCs" attrition and retention aspects will be discussed in depth by always keeping its original history in mind.

1 www.coursera.org

2 www.edx.org

3 www.udacity.com

IV. ATTRITION AND RETENTION IN ONLINE-LEARNING AND MOOCs

This chapter covers several aspects of online-learning and MOOCs, with special focus on attrition and retention. As MOOCs and online-learning share many commonalities and are hardly to distinguish, both will be discussed in a rather coherent way. In general, MOOCs aim to be freely available, with no limitation in the number of participants, whereas online learning settings are rather limited for a particular amount of students. Parts of this chapter are based on the paper of Gütl, Chang, Hernández Rizzardini and Morales (2014).

According to Gütl, Chang, Hernández Rizzardini and Morales (2014), MOOCs have gained significant publicity and popularity following the rapid development of open and online courses to a massive crowd of learners. There is an abundance of MOOC courses that are free, scalable, accessible and distributed with unrestrictive participation and in recent times the list of MOOC platform providers has also grown (Adamopoulos, 2013). Prominent MOOC platform providers such as Stanford University's Coursera had raised a total of $85 million and edX had a total of $60 million funding by MIT and Harvard at the end of 2013 (Bersin, 2013). The student population had reached 5 million students by December 2013 for Coursera and more than 1.3 million students for edX (Fowler, 2013). Other providers such as Udacity and Udemy have each raised some $21 million and $16 million (Holdaway, 2014). Udacity had reached a student population of 1.6 million in 2013 (Chafkin, 2013).

MOOCs aim to revolutionize the way of education. They cover lots of different areas like Mathematics, Engineering, Computer Science and many others. Plenty of renowned universities already noticed the potential of this new way of learning and started offering lots of courses. For example Stanford University offered MOOCs which attracted tens of thousands of participants. The same holds for companies offering MOOCs professionally, like e.g. edX, Coursera and Udacity, as already seen previously. Students have many benefits in attending MOOCs: Many MOOCs are publicly open, so everybody can register. There is no physical presence needed, meaning you can participate from anywhere at any time. Technologically progressed open educational tools enable professors and educational institutions to provide MOOCs with a high number of students, where geographic and economic barriers vanish. Furthermore, MOOCs basically enable an unlimited number of students to participate. MOOCs don't require any particular pre-knowledge, meaning students with various different backgrounds are free to join any course. Because of these and other reasons MOOCs expanded rapidly and became significantly popular for students as well as educators. Besides the big boom and acceptance of MOOCs there are still various issues which must be resolved: MOOCs face extremely high dropout rates. Additionally, the actual needs of students should get considered in more detail (Adamopoulos, 2013).

MOOCs continued the trend of the *Open Education* movement which considers how e.g. educational tools, resources and knowledge can be used to enhance quality of

education, which was accompanied with very low retention rates of students. On the other hand it could be observed that students' interest in MOOCs is already extraordinarily high, based on the observation that lots of students start a specific MOOC. For that it is necessary to find out causes and derive corresponding feasible solutions (Adamopoulos, 2013). Gütl, Chang, et al. (2014) stated that *Open Education* has been in existence since the early 1990s and it is gathering prominence, given the development of the digital era (Oecd, 2007). *Open Education* gives the participants the opportunity to attend, participate and access learning content. Some institutional-based learning in *Open Education* has no or minimal barriers to entry, for example, no admission requirements and participants are allowed to complete a part or the entire degree program. Some programs are free, some are priced at a low cost, others may include the costs of acquiring a certificate or providers may offer digital badges as an accreditation credential (Carey, 2012). The dynamic nature of MOOCs provides both advantages and disadvantages. MOOCs are generally open and free, usually have no or limited entry barriers and can have a large group of participants who collaborate to discuss course content. On the other hand, if not managed and controlled properly, simultaneous conversations, postings and streams of discussion may overload some participants. Technical issues with Information Communications Technology (ICT) infrastructure such as broadband, hardware and software must also be considered when accessing MOOCs. Participants' access to MOOCs can vary depending on technical connectivity, quality of course content, and most importantly the participant's motivation and desire to learn (EDUCAUSE, 2011). The notion of choice is an important aspect when it comes to studying with a MOOC, as well as the participants' real intent to study with a MOOC. It has been reported that some participants had no intention to complete a course but purely to discover, explore and find out more about the content (Kolowich, 2013). Early data has also begun to unearth the participants' motivation to access and study with MOOCs. Nevertheless, massive enrolments have resulted in massive attrition and low completion rates (Carey, 2012; Rosen, 2012). The question to ask is whether MOOC dropouts are viewed as a sign of deficient quality or whether they are an expression of an individual's choice. In case the latter holds, then this must be factored in the design of MOOCs. This also clearly indicates to researchers and instructors alike, that there is more to completion and attrition rates than simply focusing on low MOOC completion rates.

Within the last years MOOCs rapidly expanded and so special online communities emerged to reduce the required effort associated with finding the appropriated course or platform. Those communities offer user-generated course reviews, which is an important source of information for both, students who consider attending one and researchers, aiming to understand the phenomenon behind MOOCs (Adamopoulos, 2013).

In general, as it could also be seen in the historical learning settings, dropout rates appear to be significantly higher in terms of online learning than for traditional courses. (Berge & Huang, 2004).

Gleason (2004) concluded years before MOOCs even appeared that there will always be students who drop out of programs, whether this is in case of an online or face-to-face environment. He stated, that it is basically very unlikely that 100 percent retention rates can ever be achieved. That time specific online courses had already been offered

for over 20 years. Even after this period, this teaching concept was still new to many students and it was different in such a way so that all participants, starting from the supportive staff up to the students, needed to adapt in order to achieve the expected outcome.

Adamopoulos (2013) had difficulties in comparing his research on MOOCs with prior literature related to traditional education because of the novelty of MOOCs and the different levels of analysis. He stated that future research needs to inspect how online socialization influences the dropout decision of students, participating in MOOCs. Further, he concluded that student attrition is a continuously ongoing problem with sometimes conflicting outcomes. According to him not much has been done in order to effectively improve retention during the last centuries (Glass & Garrett, 1995). He criticized that in the past the majority of attrition studies did not examine any individual course characteristics, but rather focused on finding potential students who are performing well for later recruitment, which strongly differs from the nature of MOOCs, highlighting general openness (Adamopoulos, 2013).

Adamopoulos (2013) tries to tackle the problem of the very high dropout rates by employing the Grounded Theory Method (GTM) with user-generated online reviews. It resulted that professors mostly influence retention in online courses and that they have the highest impact on students to successfully complete a specific course. The course assignments and the course material also positively influence the successful completion of a course. Further, the study revealed that using a discussion forum had a negative impact on completing a course, but a positive effect on the probability that the course gets completed partially. Self-paced courses had a negative influence, when they got compared with other courses which had a fixed time-table. Additionally, the difficulty and the workload, as well as the duration of a course, had a negative impact on its completion. On the other hand, if a course had been designed more difficult and with a higher workload, a self-paced timetable with a longer duration over weeks resulted in a higher probability that the corresponding course will be completed successfully. Further, it was found out that final exams as well as final projects and peer assessments also have a positive influence. Considering textbooks which complement the course, they only have a positive impact if they are freely available, otherwise a large portion of the students cannot access the contained information, which would result in the opposite effect. The study showed that a certificate, which gets handed out if a course had been completed, also affects retention. Furthermore, the reputation of a university positively influences the completion rate. In terms of the content of courses, it was found out that academic disciplines like Business and Management, Computer Science and Science in general, significantly increased the completion rates, whereas disciplines like Engineering, Humanities and Mathematics had reduced rates in completion. Additionally, student characteristics, like gender or formal education, did not correlate with attrition. These findings resulted in recommendations and other implications: It was suggested that the course characteristics, like difficulty, workload, duration, manual or automated grading, etc., are strong indicators for the final satisfaction of students, which could be seen as guidelines for designing a course. As an example, it was stated that MOOCs should rather have a fixed timetable which has been set by an instructor. In case of a more difficult course, the students should be allowed to plan their own time-schedule. Further, it was found out that the form of certification could be improved so that it gets

more useful for students, which would also affect their motivation to complete a course. The analysis of the used discussion forums showed that there was room for improving the completion rates. It was suggested that other mechanisms or complementary technologies are required in order to connect, assists and advice the students. Herein, it was recommended, that a wiki-system could be a solution. The results with the high variance of completing a course, depending on the taught scientific discipline, suggested that some disciplines, like e.g. Computer Science, are more appropriate for online courses or students accept them more (Adamopoulos, 2013).

As in the section “Computer-based Learning” mentioned, the usage of computers enables representing information in very different ways, which supports learning. Moreover, this technology enables interlinking the different representations in a way it was never possible before. Interlinking information gives the learner the possibility to easily acquire related knowledge and supports the learning process in general (Winters et al., 2008). Due to that one recommendation is to represent data in a corresponding manner and to provide cross-linked information, which gives the learner the possibility of constructing a broader and profound knowledge.

In Pappano (2012) it was mentioned that some learners prefer local learning groups for MOOCs. Thus, some try to organize local evening discussions of the week's material followed by a social hour. It was mentioned that especially those MOOCs got completed successfully, where local, weekly meetings of student groups could be achieved. This shows that learners try on their own to overcome the feeling of disconnect. Nevertheless, this approach is highly limited to a particular location, which means that more effort should be put in reducing the feeling of disconnect and introducing ways to enable students interacting more socially.

In the following, different, selected contributions in literature will be discussed consecutively, in order to better illustrate their observations, findings and models in terms of attrition and retention aspects in online courses. Further, relevant information concerning MOOCs will also be covered.

An Attrition Analysis of Open Courses and MOOCs

Gütl, Chang, Hernández Rizzardini and Morales (2014) discussed various aspects concerning attrition and stated the general question, whether one should be concerned about the massive dropout rates in MOOCs. Firstly, a MOOC experiment shows different types of attrition, namely healthy and unhealthy attrition. Students who successfully complete a course are considered as persistent learners, also coined as completers. Students with healthy attrition are defined as those who complete a specific part of a course. Some mentioned main issues of unhealthy attrition get related to poor technical infrastructure, the lack of support from employers, poor time management, lack of background knowledge and skills, a bad learning experience, a lack of personalization, which partially resulted in feeling isolated and disengaged. Furthermore, it is also mentioned that a required high level of interactivity may contribute to the students' feeling of having less control about the course. Information overload is another issue

which can show up. In this section, some relevant parts of their work will be covered in detail:

The rapid development of digital media, Internet and Web technologies as well as advanced technological devices have strongly enhanced the provision of education. As already seen in chapter “III. HISTORICAL RESEARCH ON ATTRITION AND RETENTION”, access to education is now much more flexible and open, firstly with open access to learning content, and then free access to online courses which led to open virtual classes for the masses. MOOCs have become the new hype in e-education. Educators and educational institutions are grappling with this new and powerful learning concept that has attracted thousands of learners. This new way of learning has no or lower entry barriers. It can bring diverse and an international group of learners together. MOOCs have the potential to be a disruptive technology if not managed carefully. Given the expected high volume of learners, MOOCs may become unmanageable with the amount of students' contributions and insufficient guidance and support by the instructors. Educators and educational institutions might also find the very high dropout rates of more than 90% alarming. To manage the expectations of MOOCs, Gütl, Chang, et al. (2014) investigated the reasons of such high dropout rates. Findings from literature and experimentation for their study revealed that attrition may be grouped into two classes. The first is classed as “healthy” attrition, which concerns a group of learners who is selectively picking content of courses. The second class is named “unhealthy” attrition, which subsumes all students initially wanting to finish a MOOC but fail for various reasons. Students who complete a MOOC successfully are classified as “persistent” learners. This distinction is important as it enables strategies to be developed to deal with issues and causes surrounding each group. Many educators would agree that the second class of “unhealthy” attrition is the one that they are most concerned with. When the MOOC attrition class distinction is made, the results show that the actual dropout rates are significantly lower.

While MOOCs are scalable and the courses can attract a large number of students world-wide, early studies have highlighted problems with low retention and high dropout rates. The majority of courses reported a completion rate of less than 10% and with an average of about 7% (Jordan, 2014). Gütl, Rizzardini, et al. (2014) stated that MOOC settings have a very low retention rate which is usually between 3 and 8%, whereas a retention rate of e.g. 0.4% can also occur (Liu et al., 2013). Some issues of high dropout rates relate to a poor learning experience and a lack of personalization which attributed to students feeling isolated and disengaged. Other issues, as reported by Brinton, Chiang, Jain and Lam (2013), relate to the design of the learning activities with the level of interactivity. While the high volume of interactions may be seen as a promising learning engagement pedagogy, the inability to organize, manage and control interactions and the level of discussion may result in information overload. In this regard, poor navigation tools may have made it difficult for participants to follow postings and the thread of discussion (Brinton et al., 2013). Other negative experience included inadequate technical infrastructure to access the course material (Gütl, Rizzardini, et al., 2014).

There are many reasons why the retention rates are low (Clow, 2013; Downes, 2012; Lewin, 2013). Often a big percentage of people who signed up for a MOOC have varying intentions, some are completers, others are interested in the subject's content and

there are also those who intend to complete but fail to, for various reasons. Given these combinations, an understanding of the actual needs of students is a good indication to appreciate the different types of MOOC attritions. Also, given the popularity of MOOCs and that MOOCs have existed for a while now, there is rich data which can reveal the student's real intent (Fox & Patterson, 2012; Pappano, 2012).

Retention, Attrition and Persistence

Dropout from educational programs is a systematic problem and concerns educators and researchers for a very long time. From a historical perspective, the percentage of students leaving higher education programs is about 40 to 45% for the past 100 years on average. The retention of students must be viewed in a holistic way and is not only caused by a few issues (Berge & Huang, 2004; Park & Choi, 2009). In the context of dropout research, three concepts are important, namely attrition, retention and persistence, which were all defined in the subsection “Common Terms” (Berge & Huang, 2004). Research on attrition is focused on “brick and mortar” or physical educational settings, distance education and also increasingly focused on e-learning settings in recent years (Levy, 2007), which was the reason why this report started with historical learning settings in the beginning. Gütl, Rizzardini, et al. (2014) also mentioned that research on attrition and retention is an active process spanning across eight decades, which includes learning settings with traditional, physical presence up to settings with distance education and e-learning. Research in e-learning settings revealed that dropout experiences may lead to frustration and lower confidence in learning. Given the significance of research on attrition, retention and persistence, a number of theoretical frameworks have been constructed ever since the beginning of education, aiming to understand and describe the effects of attrition. Four perspectives introduced by Berge and Huang (2004) included the social influences, organizational characteristics and processes, economic influences, and psychological characteristics. A selection of the relevant and important models will be introduced in the remainder of this section.

An early model developed by Tinto (1993) and further improved in the following two decades is the Longitudinal Model of Individual Departure. The student's decision for dropout or persistence is modelled by pre-entry attributes, goal commitments, formal and informal experiences on academic and social aspects, as well as academic and social integration. The original intention was to model traditional students (Berge & Huang, 2004; Park & Choi, 2009). A conceptual model for non-traditional students was developed by John P. Bean and Metzner (1985). This model covers factors on academic performance, psychological aspects, background and environmental aspects. It was emphasized that non-traditional students are more influenced by the external environment on their dropout decision than traditional students. Although the focus of this model was on non-traditional students, the model covers aspects for both traditional and distance learners (Park & Choi, 2009).

Given the limitations with the existing dropout models, Rovai proposed the composite persistence model (CPM) in the context of online learning in 2003 (Rovai, 2003). As shown in Figure 1: Rovai's Composite Persistence Model (CPM) (Lee, Choi, & Kim, 2013, p. 330), the model is based on previous models (e.g. John P. Bean and Metzner (1985)) and influenced by Tinto (Lee et al., 2013). The CPM is composed of four components.

“Student characteristics” and “student skills” are both dimensions prior to admission. The “internal factors” and “external factors” are the dimensions relevant after admission. According to Lee, Choi and Kim (2013), the research community has not found consensus about the relevance on student characteristics for modelling attrition.

Park (2007) has reviewed Rovai's CPM, and further developed a theoretical framework for adult dropout in online learning (see Figure 1: Theoretical framework for adult dropout in online learning (Park & Choi, 2009, p. 209)). Notably, Park found that in particular the aspects of learner skill have little empirical support for dropout in studies. In Park's model external factors are relevant prior to, but also during the course (Park & Choi, 2009).

Berge and Huang (2004) have compiled relevant variables influencing attrition and retention, applicable in various learning contexts, like physical presence, online learning and blended learning. As shown in Table 1: Variables Addressed in the Various Retention Models (Berge & Huang, 2004, p. 13), the variables are grouped into three classes: personal variables, institutional variables and circumstantial variables. “Personal variables” subsume demographic aspects (such as age, ethnicity and economic status), individual aspects (such as academic skills and abilities, motivation) and prior educational experiences (such as academic achievements). “Circumstantial variables” distinguish institutional interaction (such as academic and social interaction) and external interaction (such as life and work circumstances). “Institutional variables” comprise social aspects (such as mechanisms for social integration), academic aspects (such as structural and normative systems) and bureaucratic aspects (such as mission, policy and institutional funding) (Gütl, Rizzardini, et al., 2014).

Motivated by continual concerns on dropout issues, a number of strategies have been developed to mitigate attrition and encourage persistence in learning activities. Chyung (2004) suggested the SIEME model, a five step model to reduce dropout rates, which can also be seen in Figure 1: Chyung's SIEME model (McMahon, 2013, p. 6). Step one introduces the notable notion of the separation of healthy and unhealthy attrition. Healthy attrition subsumes cases where students find that the learning event is not in line with their expectations. Unhealthy attrition includes cases where learning events are in line with students' expectations but dropout decisions are caused by other factors. The other steps focused on unhealthy attrition and are concerned with the identification, change and evaluation of hygiene factors (e.g. course organization relating to course design, facilitation skills) and motivational factors (e.g. course satisfaction), in order to reduce dropout issues (McMahon, 2013).

Yang, Sinha, Adamson and Rose (2013) emphasized that unlike research on attrition of other forms of online learning, MOOCs raise new research questions. This is supported with an increased freedom of the learners to choose what, when, where and how to learn. Learners can control their own learning and in many settings they can personalize and select their own learning content and choose the preferred (learning) tools. Usually, there are no pre-requisites or a financial burden to enrol in open courses. The entry barrier is low and no penalty is applied if one leaves the environment. The high dropout rates reported by various sources are emphasized as one of the major drawbacks of this learning context. This finding is both disillusioning and misleading. A study by Jordan (2014) reported a median value of 6.5% completion rate across 39 courses which had a

broad range between 0.9% and 36.1%. The study also revealed a decreasing completion rate for increasing course length.

Some selected research on attrition in the context of open learning may shade some light on the reasons for the very high dropout rates. Clow (2013) proposed the Funnel of Participation Model which is inspired by the "marketing funnel" to model the process of a customer from taking notice to buying an asset. This usually happens in four phases: awareness, interest, desire, and action. In the context of MOOCs, the analogies to these phases are: awareness, registration, activity, and progress. Each phase is characterized with a large fraction of dropout. The "funnel" notion can be applied for the density of contributions in the activity phase. The effect can be illustrated by actual attrition numbers of MOOCs. For example, in the "Introduction to Infographics and Data Visualization" offered by the College of Communications at the University of Texas, Austin, Liu et al. (2013) reported that about 5000 students started the course, 44% interacted in the forum, 33% completed the first quiz, 26% the second quiz, and finally only 0.4% completed successfully. In particular in the context of the activity phase, researchers (such as Balakrishnan and Coetzee (2013) and Yang et al. (2013)) have looked into the learners' interaction pattern on learning activities to predict dropout or persistence with the learning tasks.

As already mentioned earlier, Adamopoulos (2013) followed a completely different approach, which will be stated for the purpose of comparison: Based on the Grounded Theory Method (GTM), a content analysis (text and opinion mining) of user-generated online reviews had been performed. The proposed model for online course retention suggests the following categories: Student course evaluation, course characteristics, university characteristics, platform characteristics, student characteristics.

Kizilcec, Piech and Schneider (2013) criticized the monolithic view of disengagement in MOOC settings regarding attrition research and discussion. Based on the engagement in terms of interaction patterns, they have suggested a classification method and identified four classes of engagement: The class "completing" groups learners who complete a majority of activities and eventually finish the course. The class "disengaging" describes patterns of students who take assignments in the beginning but stop over time and completely leave the course or still consume some content without taking further assignments. The class "auditing" is characterized by students taking infrequent assessments but they engage by consuming learning content. The class "sampling" includes learners who selectively consume content.

Gütl, Rizzardini et al. (2014) conducted a study to understand reasons and factors for leaving the MOOC in an "Introduction to e-learning" course offered by Galileo University in Guatemala in 2012. 1680 students enrolled in the course, only 143 (8.5%) participants completed the course and a total of 1537 (91.5%) left the course. A questionnaire was sent out to the group of students who did not finish the course and 134 students completed the questionnaire. The respondents were 69 (51.49%) male and 56 (48.51%) female. As shown in Figure 5: Dropout students and the motivation to enrol in a MOOC (Gütl, Chang, et al., 2014, p. 4), a variety of reasons were uncovered in order to understand the motivation to enrol in the MOOC.

Interestingly, only 30 students (22.39%) expressed that their objective was to “complete the course”. The rest, making up 77.61%, gave the following reasons: 45 users (33.58%) indicated that they wanted to experience the MOOC environment. 24 enthusiastic learners (17.91%) wanted a “sneak preview” into the topics. 12 users (8.96%) indicated, they wanted to audit the MOOC by learning only the content that they were interested in without having to finish the course. Five participants (3.73%) were interested in the content without formally completing the course. Eighteen participants (13.43%) had given “other reasons”, these included “having a quick view of the subject”, “deepening knowledge on a subject”, “contributes to my job activities”, “refresh and update the knowledge in a subject” and “learn about the methodology”.

Proposal for an Attrition Model

Inspired by various literature, for example Adamopoulos (2013), Chyung (2004), Clow (2013), Kizilcec et al. (2013) and Yang et al. (2013) on MOOCs' attrition, retention and completion rates, Gütl, Chang, et al. (2014) proposed an Attrition Model for Open Learning Environment Setting (AMOES), by additionally considering their own research in the same area. This Attrition Model for AMOES is shown in Figure 6: Attrition Model for Open Learning Environment Setting (AMOES) (Gütl, Chang, et al., 2014, p. 5) and is proposed to understand and differentiate the reasons for attrition.

As shown in Figure 6: Attrition Model for Open Learning Environment Setting (AMOES) (Gütl, Chang, et al., 2014, p. 5), the AMOES model is divided into three sections. These are the “attrition and retention factors”, “the open online learner group classification” and the “funnel of involvement in the open learning setting”. As the aim of attrition analysis is on the learners, the AMOES group learners are split up into three classes, namely “healthy attrition”, “unhealthy attrition”, and “persistent learners”. Interlinked with these different types are factors comprising external, internal and student factors which may contribute to a learner belonging to a healthy, unhealthy or persistence class of learners. Another contributing link is the administrative (awareness and registration) and pedagogical (activities and success) aspects of the MOOC which is termed the “funnel of involvement” in the learning setting. The final evolution of this model was based on different studies of Gütl, Chang, et al. (2014) in MOOC uptake and dropout and a deep analysis on attrition, retention and persistence of MOOCs over recent years.

Examples of external factors are competing courses that are offered in the MOOC space, varying technological infrastructure in different countries, cultural aspects and others. As these factors appear outside of a MOOC provider, institutions could identify strategies that may curb some of the external aspects. Examples for internal factors are aspects related to the organization of the MOOC which are under the control of the MOOC provider. Student factors are influences related to student's individual desire to study a MOOC, prior knowledge of the study area, as well as other varying reasons. For example, some students enrol in a MOOC because of their job, some for general interest and others for credential aspiration. Depending on the students' intention and motivation, they ultimately form different types of learners such as the enthusiast learner, content learner, restricted learner, disengager and completer, as listed in Figure 6: Attrition Model for Open Learning Environment Setting (AMOES) (Gütl, Chang, et al., 2014, p. 5).

The funnel of involvement in open learning is a modified version of the “funnel of participation” of Clow (2013). This last section of the AMOES model is closely interlinked with the external, internal and student factors, along with the different classification types of the students. “Awareness” is closely linked with the “External Factors” in which a MOOC must exist. This is followed by the “Registration” step where students sign up and then participate in the MOOC's “Activities”. At this stage, the MOOC provider (“Internal Factors”) plays a pivotal role in controlling the amount of activities that are balanced with interactive, engaging and contributing participation, which would lead to a satisfying and ultimately successful experience.

Implied in the Activities' funnel of involvement of the offered MOOC is the dependency upon the availability, compatibility and reliability of the Information Communication Technology (ICT), which touch both, external and internal factors. Finally, the measure of “Success” is based upon the contributing student factors and the different classes of healthy, unhealthy and persistence learners.

Field Study on Attrition Aspects

In order to apply the previously introduced AMOES model, Gütl, Chang, et al. (2014) conducted a field study of a MOOC experiment. The MOOC experience and topic of the course was “Introduction to e-learning”. There were four learning modules and each module contained a short video representing the main learning content. Other resources included pre-readings and hyperlinks. The learning activities included a set of recorded tutorials and specific instructions to complete the tasks. Online discussion forums were the main source of collaboration. The MOOC used a gamified approach to motivate students and they were awarded with digital badges in case of contributions and achievements. To ensure a sustainable model, a peer-assessment approach was used in the learning activity. The complete analysis of this study is reported by Rizzardini, Gütl, Chang and Morales (2014) and Gütl, Rizzardini, et al. (2014).

1680 participants registered for the 4-weeks' MOOC “Introduction to e-learning”. The dropout rate for the MOOC was very high, where only about 8.5% of the enrolled learners successfully completed the course. Table I: AMOES Funnel of Involvement (Gütl, Chang, et al., 2014, p. 6) shows the breakdown of the attrition rates according to the funnel of involvement. There was about 33.01% active participation in the online forums in Week 1 and only about 21.60% completed the learning tasks for Week 1. The second week showed a reduced completion rate of 13.80% in the learning tasks, 26.02% of the learners respectively completed the learning tasks and actively participated in the forums. Another decrease in the participation to about 10.24% was recorded for completing the learning tasks in Week 3, while a reduction to 18.05% resulted of those who participated in the forums. All in all, only 8.50% of the learners successfully completed the MOOC.

The high attrition rate led to a study of those participants who did not complete the MOOC, namely 1,537 students in total or 91.5% of dropout students. A contact list from the “Registration” step was used to email those who did not complete the course. 437 (28.4%) responded to the email and 303 respondents (69.34%) agreed to complete a post survey. However, only 134 (44.22%) of the 303 respondents successfully completed and

returned the survey. Like most other MOOC experiences, a high attrition rate and a low completion rate had been a major source of uproar and debate among institutional scholars and educators. Due to the fact that the research is limited in this area, Gütl, Chang, et al. (2014) developed the AMOES model, where students get separated in accordance with the different learner types.

Chyung (2004) conducted various dropout studies on distance education students and indicated that dropout rates can be misleading if a deeper analysis and the reasons for attrition were not understood in the first place. For this, an investigation of the high dropout rates in MOOCs resulted in dividing attrition into healthy and unhealthy attrition. Internal and student factors may contribute to some causes of attrition that resulted in students belonging to one of these attrition groups.

The AMOES model shown in Figure 6: Attrition Model for Open Learning Environment Setting (AMOES) (Gütl, Chang, et al., 2014, p. 5) is used to determine the attrition and retention rate of a MOOC experiment conducted at Galileo University in Guatemala in collaboration with Graz University of Technology in Austria and Curtin University in Australia. The focus of this study was to examine the motivation of students and the reasons why they have started studying with a MOOC. The students were categorized into the classes of attrition (healthy and unhealthy) and the retention class (persistence) accordingly.

The healthy attrition group may include those students who are enthusiastic learners and wanted a preview of the course to gain a quick understanding of the topic. These courses may be topical, popular, innovative and current. In this group of students those are included who are eager to experience a new learning environment. This group of learners would usually try to find out if this is a learning setting which they could adopt as an alternative to traditional face-to-face or online learning. In this case, the attrition based on these aspects can be characterized as good attrition. In the MOOC course “Introduction to e-learning”, “enthusiast learners”, as part of the AMOES model, were the combination of those students who wanted a preview of the MOOC topics and those who wished to experience the MOOC. The combination of these resulted in 51.5% enthusiastic learners.

As seen in Figure 7: Group of MOOC learners according to AMOES (Gütl, Chang, et al., 2014, p. 6), the second class of good and healthy attrition is the “content learner”. These learners are selective learners and choose what they wish to learn from a list of topics. They deliberately choose to study interesting content, without the requirement to complete the entire course. Two external factors which contributed to this decision are low or no entry barriers for enrolling in MOOCs and the free access to MOOCs. This gives learners the advantage to consume content and to gain knowledge and skills. In the course “Introduction to e-learning” about 3.73% of the learners were part of this class.

The final, third class of the healthy attrition is the “Restricted Learner”. This is a group of learners who decide to audit the entire course without the requirement to formally complete the course or to complete the assigned work and/or participate in gaining digital badges. In case of the MOOC “Introduction to e-learning”, this scenario occurred to about 8.96% of the learners. The three classes “enthusiast”, “content” and

"restricted learners" are those learners who get classified to healthy attrition, as these learners have personalized their learning and they have also gained considerable knowledge and skills according to their learning requirements and needs.

In terms of unhealthy attrition, the learners failed to continue due to a variety of reasons. This may be caused by external factors such as technological limitations, student work or job commitments. Other causes can be internal factors, such as the institutional operation of the MOOC, the organization of the course itself, the lack of supportive services and student factors, such as the lack of prerequisite knowledge required to pursue the course, the inability to raise time to study the material or engage in online discussion forums. Another contribution to this attrition group is with the "Activities" of the funnel of involvement. The activities may be poorly designed, it may be difficult to access the materials and other issues related to the pedagogical design of the course might show up. The causes of unhealthy attrition require thorough investigations. In the MOOC "Introduction to e-learning", the unhealthy attrition was caused by students who were disengaged, which is the reason why they are now known as the "Disengager" learners. About 22.39% of the learners in this MOOC indicated that they would like to complete the course and for a variety of reasons but were unable to.

In the data analysis of the MOOC "Introduction to e-learning", 13.43% selected "Other Reasons" for their inability to complete the MOOC. Some reasons such as "having a quick view of the subject", "deepening knowledge on a subject", "contributes to my job activities", "refresh and update the knowledge in a subject" can be easily classified into "healthy attrition". As already mentioned earlier, about 8.5% of the enrolled learners were completers or persistence learners. Table II: MOOC Learners according to AMOES (Gütl, Chang, et al., 2014, p. 7) and Figure 8: MOOC learners according to AMOES (Gütl, Chang, et al., 2014, p. 7) show the breakdown of healthy attrition, unhealthy attrition and persistence learners. Figure 9: Group of MOOC learners according to AMOES (Gütl, Chang, et al., 2014, p. 7) shows the group of learners according to attrition, comprising healthy and unhealthy attrition, as well as retention which is based on persistence. Other learners are also shown. It shall be noted that in all three illustrations (Table II: MOOC Learners according to AMOES (Gütl, Chang, et al., 2014, p. 7), Figure 8: MOOC learners according to AMOES (Gütl, Chang, et al., 2014, p. 7) and Figure 9: Group of MOOC learners according to AMOES (Gütl, Chang, et al., 2014, p. 7)) the completers with 8.5% refer to all participants who registered at the beginning. That is why this value cannot just be summed up with the other values. It has been illustrated this way to show all different classes of learners.

Lessons Learned from Dropout Students

The previous findings of section "An Attrition Analysis of Open Courses and MOOCs" resulted from a former work conducted by Gütl, Rizzardini, et al. (2014). Some selected key facts of this work are, that one main motivation for enrolling in the MOOC was to find out what the MOOC actually was about and to collect experiences with it. In contrast, the study's aim was to investigate the motivations of those students who enrolled in a MOOC with the initial intention of completing it, but turned to non-completers for various reasons in the end. The majority of students who fell into this category mentioned personal reasons such as changes in their job, lack of employer's

support or health reasons. In terms of academic reasons for not completing MOOCs, difficulty in juggling work and study, technical inability, poor course design and a high workload got mentioned. Surprisingly, 98% of those participants who did not complete the MOOC still have a positive attitude of it, considering a MOOC as a useful way for studying. Furthermore, they would use a MOOC in the future. Many of these registrants see a MOOC as a flexible alternative which eliminates geographical barriers in terms of time and distance. Moreover, they see a MOOC as another way of acquiring new skills and knowledge. A used System Usability Scale showed that the perceived usability is lower for non-completers, which is not surprising. Concerning the Intrinsic Motivations Measure, the results indicated a high motivation for both groups to learn, whereas students, who dropped out later on, were less keen in learning to use new tools.

In this subsection, some important details of Gütl, Rizzardini, et al. (2014) will be covered in order to present more details and findings, where its content is based on their work:

The same survey, as mentioned previously, with 134 students who had not completed the MOOC, revealed that only 22% of the students intended to complete the MOOC but were unable to due to various factors, including academic and personal reasons. A big majority of the students indicated that changes in their job, insufficient time, difficulty with the subject matter and unchallenging activities are some reasons for the dropout. The following is primarily denoted to investigate those students who had not finished the MOOC.

As mentioned earlier, the used MOOC focused on the topic of “e-learning” and was organized in four parts, which were to be completed within four consecutive weeks: (1) “Introduction to e-learning”, (2) “Technological platforms for e-Learning”, (3) “How to create a fascinating e-learning course”, and (4) “Developing an e-Learning course”.

The overall goal of the presented research was to uncover the motivation for enrolling in the MOOC, the reasons for leaving the MOOC and how students organized (when and where) to work on the MOOC. The comparison between these groups, those who had finished and those who left the MOOC mid-way, will be presented later on. The study also determined insights on motivational, emotional and usability issues. Some important and interesting findings got especially highlighted in the work.

The experimentation procedure for both groups included the following steps, which can also be read in the paper of Rizzardini et al. (2014): (1) students enrol in the MOOC, (2) students complete a pre-questionnaire to gather demographic details, (3) students undertake an orientation week to familiarize themselves in the MOOC environment, (4) students access four weeks of learning activities, participate in online collaborations and complete assessments. Finally, (5) students who had completed the MOOC were asked to participate in a post-questionnaire to evaluate their own performance and the overall MOOC experience. To complement the initial research, as reported by Rizzardini et al. (2014), an additional questionnaire focusing on the dropout aspects was developed for this study. This questionnaire was sent to those students who did not complete the MOOC. Consequently, two questionnaires were made available for step five, one specifically for “completers” and the other for “non-completers”. The investigation included the MOOC tools, content, cloud-based tools, the surveys, user behaviour and

user collaborative contributions over online forums, data entries from the peer assessment process, views and experiences from the instructors/professors and interviews with the tutors and students.

The pre-questionnaire contained questions on demographics and learning preferences. The post-questionnaire for the group who had finished the MOOC contained the standard measurements as listed above. Open-ended questions captured the learners' opinions about cloud-based tools and the overall MOOC experience. For the group who left the MOOC, questions on dropout aspects were adapted from Willging and Johnson (2004). For both groups the following standard measurement instruments were used: The Computer Emotions Scale (CES) by Kay and Loverock (2008), the Intrinsic Motivations Measure (IMM) by Tseng and Tsai (2010) and the System Usability Scale (SUS) by Brooke (1996). For CES and IMM, a four point Likert scale was used and a 5-point Likert scale for SUS.

For the four weeks' course 1680 learners were enrolled from 30 different countries. As the MOOC organizers were located in Guatemala, the majority of the participants were located in Guatemala (76.60%), followed by Spain (5.11%), U.S.A. (3.63%), Honduras (3.09%), Mexico (2.20%) and others (9.04%).

The following findings are based on the data collected from the 134 students who had not completed the course but answered the questionnaire of the follow-up study. Guatemala had a response rate of 76.12%, 4.47% from the USA, and 2.24% from Spain and Mexico. On average the users were M=39.95 (σ=11.32) years old, with an age ranging from 17 to 63 years. Fifty users (or 37%) had MOOC experiences, 84 (63%) have never been exposed to this experience.

In order to get a better understanding of the reasons why the participants did not finish the MOOC, they were asked about personal, academic, support and learning environment reasons. Among the personal issues (see Figure 2: Personal reasons for leaving the MOOC (multiple answers possible) (Gütl, Rizzardini, et al., 2014, p. 7)), 92 participants (or 69.40%) indicated the main reason was a change in the job responsibilities during the course. This was followed by 20 (14.93%) indicating “personal health problems”, 18 (13.43%) had the opinion that the program had not met their expectations, 11 (8.21%) stated “family problems”, 10 (7.46%) raised “financial difficulties”, and 4 (2.99%) said the “company did not support MOOC participation”. In terms of academic reasons for dropping the MOOC (see Figure 3: Academic reasons for leaving the MOOC (multiple answers possible) (Gütl, Rizzardini, et al., 2014, p. 7)), 94 participants (or 70.15%) indicated it was difficult to work and study at the same time. Twenty participants (or 14.93%) indicated they were “not technically prepared for this program”, 12 (8.96%) stated the program was too difficult, and in contrast 10 (7.46%) emphasized that the “program was not challenging”. Nine respondents (6.72%) indicated that “classes were poorly taught” and 5 (3.73%) said the course was poorly designed. Finally, four (2.99%) found they were “not academically prepared for this program”.

In terms of reasons concerning help and support (see Figure 4: Insufficient support reasons for leaving the MOOC (multiple answers possible) (Gütl, Rizzardini, et al., 2014, p. 7)), 47 participants (35.82%) indicated the main issue was that they did “not get enough encouragement/support to continue from colleagues, family or employer”. This

is followed by poor feedback (i.e. “have not received useful feedback on assignments and tests”) by 43 participants (or 32.09%). Thirty participants (22.39%) indicated they “have not received the necessary training to use the technologies required in the course” and “not enough support from the technical staff” was raised by 24 participants (17.91%). The learning environment aspect (see Figure 5: Learning environment reasons for leaving the MOOC (multiple answers possible) (Gütl, Rizzardini, et al., 2014, p. 7)) revealed a number of reasons. Forty-four (32.84%) participants selected “other”, which was the most selected category and included issues such as “slow Internet connection”, “too many forums which caused confusion” and “lengthy and boring videos”. This was followed by 38 (28.36%) participants who argued that they had too little interaction with other students, 33 (24.63%) perceived “too little interaction with the instructors”, 22 (17.16%) found their “typing skills were not sufficient enough to interact with the class”, and 20 (14.93%) emphasized that the “learning environment was not personalized”.

To learn more about how the students had participated in the MOOC, they were asked when they had set time to work on the course (see Figure 6: When students set time aside to work on the course (Gütl, Rizzardini, et al., 2014, p. 8)). Sixty-eight (or 50.75%) participants had indicated “at home after work”, followed by 20 (14.93%) who spent the “weekends” going over the learning tasks. Fifteen (or 11.19%) worked on the course “during lunch time” and 14 (10.45%) allocated time “at work”. Seventeen (or 12.69%) gave “other” reasons which included “at work and home”, “at night” or “did not have time”. The allocated “time” to work on the course was also quite low (see Figure 7: Average hours per week allocated to spend on the MOOC (Gütl, Rizzardini, et al., 2014, p. 8)). Sixty-one participants, which are almost half of the respondents (or 45.52%), did not allocate more than one to two hours and only 15 respondents (11.19%) spent five hours or more studying the course each week. Twenty-four (17.9%) did not specify their effort.

A set of questions using a 5-point Likert scale (from totally dislike to totally like) was created to determine the overall perception of the MOOC according to the features of the course's content. Table 1: MOOC core feature student's perception (Gütl, Rizzardini, et al., 2014, p. 8) shows the details. Among all subjects, there was neither a clear preference for specific media nor dislike of the media.

In order to uncover the differences between the groups of users who had and had not finished, a comparison of the students' perception on learning activities, emotional and motivational aspects was conducted. The results for the group of students who had completed the MOOC are also reported in Rizzardini et al. (2014).

A 5-point Likert scale (from totally disagree to totally agree) was used to determine the overall perception of the MOOC experience between those who had and had not finished the MOOC (see Table 2: MOOC Learning Activities student's perception (Gütl, Rizzardini, et al., 2014, p. 9)). The perception on the various aspects of the group who had finished is notable better than from the other group.

Focusing on the emotional aspects, the Computer Emotional Scale (CES) of Kay and Loverock (2008) was applied. Four different emotions, namely happiness, sadness, anxiety and anger, are described by 12 items. The item “helpless” (on the “anxiety” dimension) has not been included in the analysis because of inconsistencies with the

data. Table 3: MOOC Computer Emotions Scale with 4-point Likert scale from 0 to 3 after Kay and Loverock (2008) (Gütl, Rizzardini, et al., 2014, p. 9) shows the results for both groups. The findings revealed that MOOC participants perceived low anger and sadness as well as significantly higher happiness while performing learning activities. The difference between both groups is marginal and is slightly better for those group who had completed the MOOC (reverse calculated for negative emotions).

For the motivational aspects, the intrinsic motivation measures according to Tseng and Tsai (2010) were applied to assess the learners' perception of the MOOC learning experience (Rizzardini, Amado-Salvatierra, & Gütl, 2013). More specifically, Table 4: Intrinsic Motivations Measure of cloud-based tools after Tseng and Tsai (2010) (Gütl, Rizzardini, et al., 2014, p. 10) shows the motivational attitude with learning a new set of tools, utilizing the tools to finish the learning tasks and reflecting the knowledge gained from completing the learning activities. The findings reveal for both groups a remarkable high intrinsic motivation to learn with and learn about the tools, although the dropout group had a less motivated response in learning to use new tools.

With respect to the usability aspects, the System Usability Scale (SUS) by Brooke (1996) showed good results with $M=77.46$ ($\sigma=16.28$) for students who had finished the course and $M=59.94$ ($\sigma=16.51$) for students who had dropped out. The perceived usability is significantly lower for the group of students who had not finished the MOOC, compared with the other group.

Focusing on overall aspects of the MOOC, the 134 students belonging to the group who had not finished the MOOC were asked to answer open-ended questions. Responses on the question what they “did the most like” confirmed the findings above but also revealed advantages of “flexible schedule”, “methodology” and “easy access to new knowledge”. Answers to the questions what they “did not like at all” emphasized issues on “the length of the videos”, “the lack of monitoring and feedback from tutors”, “participation in forums” and the “effort to master activities”. Suggestions on the organization could improve to support the MOOC experience, including “more time and flexibility for finishing assignments”, “less content and assignments per week” and “monitoring and feedback by tutors”. In the open-ended section, the participants indicated that they needed to enhance their overall effort to succeed. Comments such as “discipline”, “focus”, “time management and planning”, “developing digital abilities”, “active communication” were some mentioned initiatives required to complete the MOOC.

Interestingly, from all 134 students who had not finished the MOOC, 131 (97.76%) are considering MOOCs as a useful way to study online courses, and 132 (98.51%) would also consider attending a MOOC in the future. An illustrative selection of reasons included “setting is flexible”, “eliminates distances and optimizes time”, “learning methodology is effective and innovative”, and “it enables to acquire new knowledge”, which were some intrinsic desires.

V. USER MODELS AND USER PREDICTIONS

This chapter presents user models and user predictions which are the result of practical experiments with a set of eleven MOOCs. It starts with a general overview, followed by a preliminary analysis in terms of completers, non completers and dropout. Subsequently, classification experiments will be presented which aim to identify different classes of students. Finally, the chapter concludes with a discussion about feature selection and ranking.

Description of the MOOCs

In total, the preliminary analysis focused on 11 different MOOCs, offered by the University Galileo in Guatemala. Each MOOC had different goals and an individual focus. There were more technical ones as "Android" and "Digital Interactive TV", as well as more theoretical ones as "Community Manager" and "E-Learning". The considered MOOCs can be found in Table 1, where each MOOC is briefly described by its content and intended target groups.

MOOC	Content	Target
Android	Learn basics of developing mobile apps for Android Highly practical course with weekly tasks, demos and code Object-oriented programming and preferably Java are required knowledge	Students
Cloud Based Learning	Overview of the characteristics of cloud computing applied to education	Professionals, Teachers
E-Learning	Learn how to develop content in an effective virtual modality A various set of different educational and technological tools will be presented and used Premise of the course is that there exists no perfect recipe to ensure successful implementation	Professionals, Teachers
Community Manager	Learn professional management skills, tools to position brands, strategies and promotion on the web	Professionals
Medical Urgencies	Understand and apply the basic and essential knowledge of first aid Participants will learn how to face those first crucial minutes with maximum security. Interactive content, videos, references, forums and case studies will enrich the learning experience	No specific target
Client Attention	Learn and develop the key features and characteristics of professional customer service	Professionals
Digital Interactive TV	Theory and practice necessary for the design and development of interactive applications for digital TV A new interactive concept of television is introduced and discussed The course will include the use of the software Eclipse and the installation of Ginga-NCL	Professionals
E-Learning (Tools)	Learn to design interactive virtual courses easily Discover Open Source alternatives and commercial tools to develop HTML5 content for Moodle platforms Introduction to eXelearning and Rapid E-learning	Professionals, Teachers
User Experience	Provides a general introduction on efficient design and implementation of user interfaces Gives a firm foundation based on scientific theories and experience, and a selection of tools Practical focus will be based on evaluation and understanding examples of Web and mobile interfaces	Students

Table 1: *Considered MOOCs with a brief description*

Each MOOC had a duration of eight weeks. The evaluation of each student consisted of different activities and quizzes throughout the eight weeks' duration, culminating with a final examination. Some MOOCs also required submitting a project and weekly quizzes, which were meant as a form of self assessment for the students. The final score was a summation of the grades of all assignments, the project and the final examination. Table 2 shows typical activities within the MOOCs.

Type of Assignment	Description
Activity	Weekly tasks or assignments students have to submit
Quiz	Self assessments for the students (no points given)
Project	Normally to present before the end of the MOOC as single or group work
Final Test	Last assignments within a MOOC

Table 2: *Different MOOC assignments*

At the end of each MOOC, a survey was sent to the students who didn't manage to complete the course. The survey had the purpose to collect feedback from the students in order to understand the reasons which lead to the dropout of the MOOC. The questions were tailored to the individual, different types of students, according to what their expectations and goals for the MOOCs were.

As expected, not all the students returned a filled out survey, only a small part of them did. Because of this, it was possible to split the students enrolled to a MOOC into three different groups: “completer”, “non completer answered survey” and “non completer unanswered survey”. In the same way, also the overall MOOC log file was split into three files, one for each group. It is important to notice that students who enrolled but never interacted with the MOOCs don't show up in any of these logs and are therefore not considered in the conducted analysis and experiments. Each of the three logs consisted of the following five properties:

- UserId: The system id to identify a student
- Origin URL: The URL from which the request came from
- Request URL: The requested URL
- Timestamp: The date and time at which the request was made
- Tool: A categorization of the kind of resource which was requested

The students who answered the survey were analysed and a second classification was performed. These students were not only split according to healthy and unhealthy attrition, but the healthy attrition group was even further subdivided into the subgroups “enthusiast learner”, “content learner” and “restricted learner”, as described in “IV. ATTRITION AND RETENTION IN ONLINE-LEARNING AND MOOCs”. This second classification was used in the attrition experiment, described in detail in the section “Experimental Setup”, within this chapter.

The MOOC platform consisted of different types of tools, where each type indicated the kind of interaction the students had with the platform. Overall, there were 19 different types of tools. However, it was discovered that eight out of 19 tools were accounted for more than 99% of the total requests. Thus, the analysis was conducted focusing only on those eight tools which are described in the following:

- Assessment: Tool, used to measure the knowledge or satisfaction of a student in the MOOC. The assessment tools get classified into two types: online quizzes and surveys. Online quizzes are usually the most used tools in MOOCs.
- Assignment: Link to the list of assignments and their description, grouped by the type of assignments (e.g. tasks, projects, participation, etc.)
- Course board: Page, which has a description of the topics that the students will learn per week
- Evaluation: Tool, used to upload, download or view tasks assigned in the MOOC. It also shows the grades of the students, together with the evaluation for each single task the students have submitted.
- File storage: Tool, which contains all the files used in the course, such as documents or resources. The files are presented as a list with information such as file size and last modification.
- Forums: Link to show the last threads of the MOOC forum and the student's threads within it
- Learning content: Tool, used to show the content uploaded by the instructors. It shows the whole content of the MOOC, including not only links to the files and resources but also videos, audio, mind maps, images and others could be included.
- Peer evaluation: Tool, used to make a student peer review

Preliminary Analysis: Completer, Non Completer and Dropout

With this as a basic background, a general analysis regarding the number of students was conducted, were completers and non completers were analysed. Table 3 shows a summary concerning the number of students enrolled for a particular MOOC and the dropout, both calculated with all enrolled students and only with the active students as well. Even only considering the active students, it is obvious that the dropout rate (except for three MOOCs) is always above 50%, reaching up to 87% for the MOOC “Android”. When the whole number of registered students is considered, these rates increase dramatically and are always higher than 90%.

MOOC	Registered	Active	Completer	Non Completer	Overall Dropout Rate	Active Dropout Rate
Android	8852	593	77	516	99%	87%
Cloud Based Learning	2055	279	123	156	94%	56%
E-Learning	2141	245	81	164	96%	67%
Community Manager	9145	821	320	501	97%	61%
Medical Urgencies	2112	118	49	69	98%	58%
Client Attention	2915	91	60	31	98%	34%
Cloud Based Learning (Edu)	1350	182	99	83	93%	45%
Cloud Based Learning (Tools)	2045	317	131	186	94%	59%
Digital Interactive TV	847	121	63	58	93%	47%
E-Learning (Tools)	1679	257	101	156	94%	60%
User Experience	2150	189	62	127	97%	67%
Total	35291	3213	1166	2047	96%	58%

Table 3: *Summary of the enrolled students*

Despite the high difference in the number of completers and non completers, the total number of interactions with the platform of the former group is (as expected) higher than the one of the latter group. This is also shown in Figure 1 to Figure 5. Figure 1 and Figure 2 refer to the MOOCs “Cloud Based Learning” and “Medical Urgencies” respectively and they show the number of requests per day and type of students as well as the average, each over the whole MOOC duration.

It is important to notice that the number of requests for the completers is almost always higher than those of the non completers. It is also noteworthy that during the first and the final week very few interactions are happening, meaning that most of the activities in the MOOCs are concentrated in the middle of the eight weeks.

Figure 3 and Figure 4 refer to the MOOCs “Community Manager” and “Digital Interactive TV” respectively and show the total number of requests in relation to the first four weeks of the course for both, the completers and the non completers. Figure 5 refers to all MOOCs which were considered (the same as in Table 3). Even at a first glance an obvious difference in the amount of interactions can be observed, which resulted by the two different groups, and this is already the case just some days after the start of the MOOCs.

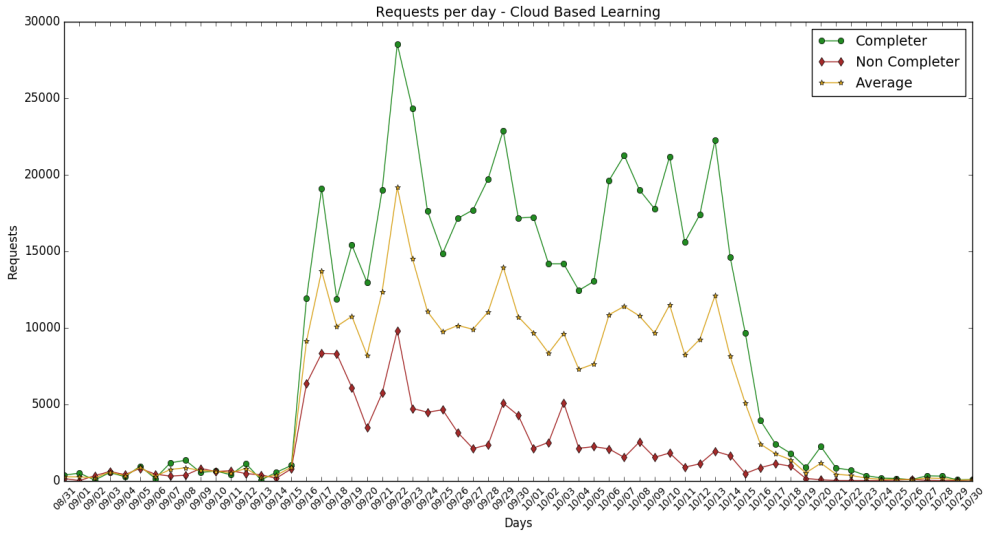

Figure 1: *Requests per day - Cloud Based Learning*

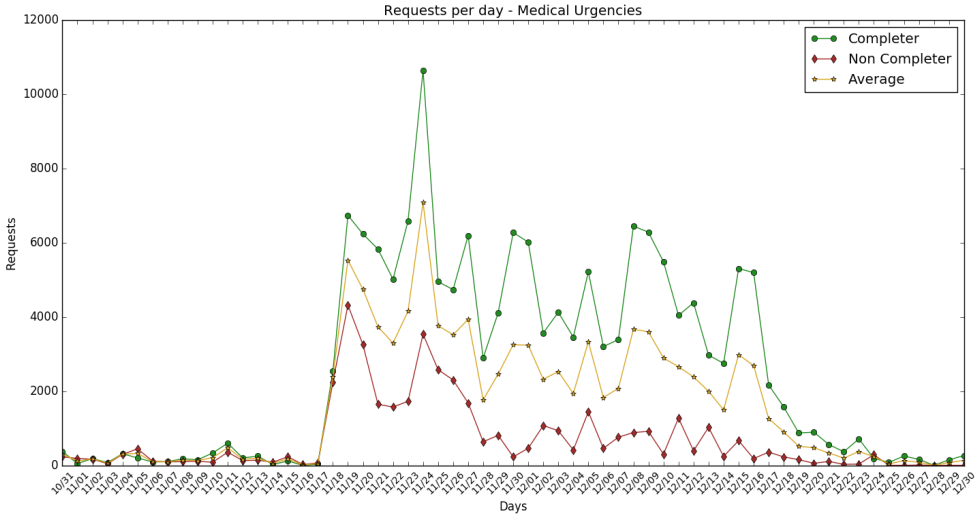

Figure 2: *Requests per day - Medical Urgencies*

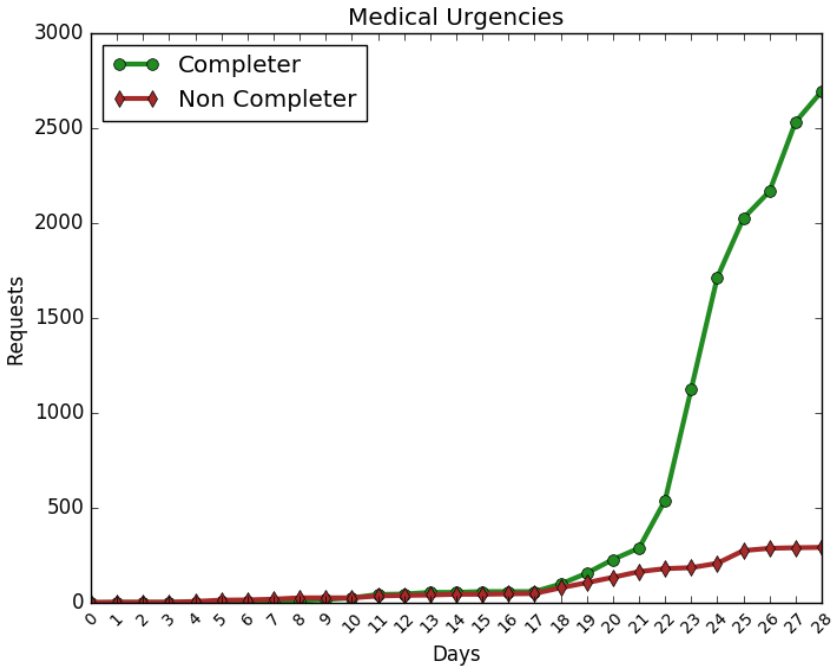

Figure 3: *Requests for "Medical Urgencies"*

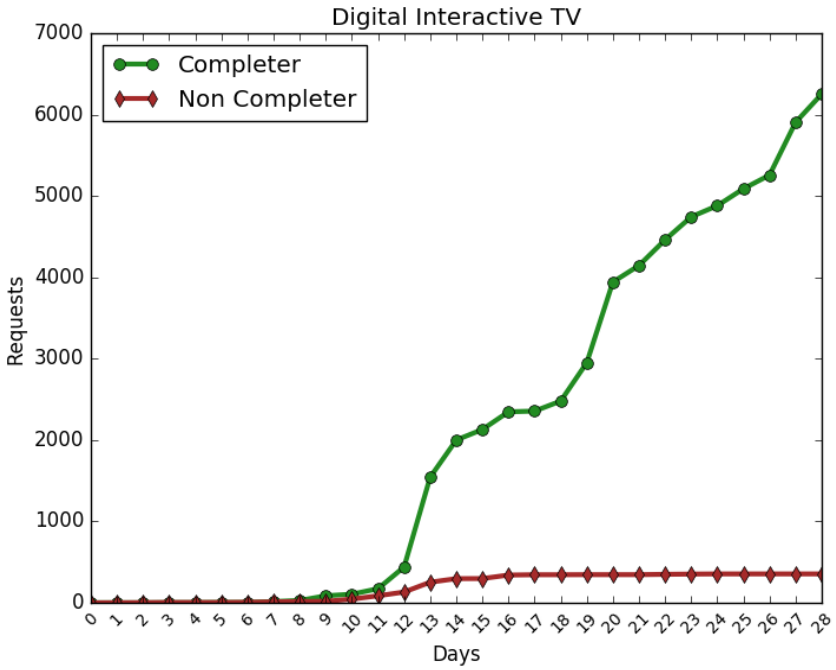

Figure 4: *Requests for "Digital Interactive TV"*

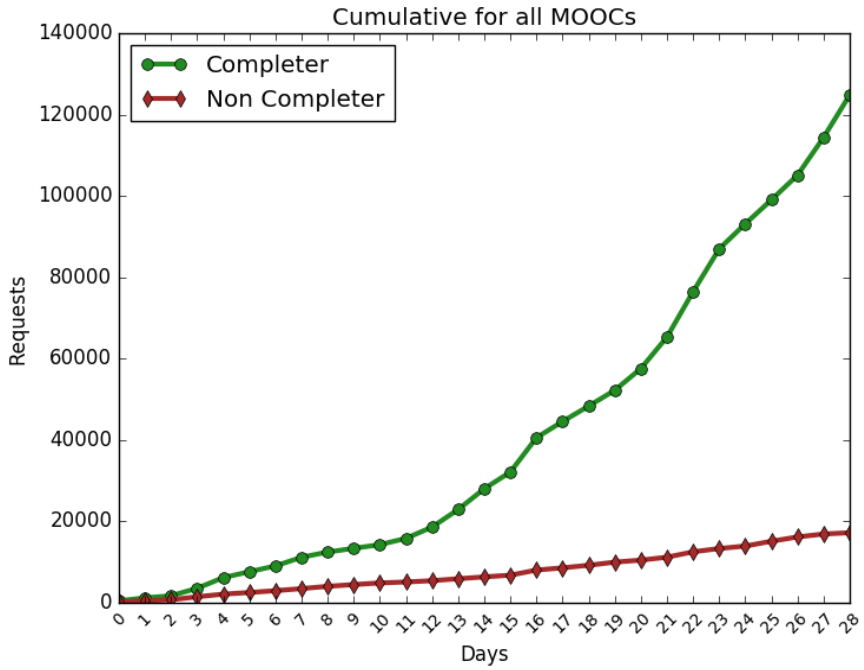

Figure 5: *Cumulative requests for all MOOCs*

Experimental Setup

There were two main experiments which were conducted. The goal of the first experiment was to predict which students were likely to complete a MOOC and those who were likely to fail (completer-non completer experiment). In the second experiment the idea was not only to separate completers and non completers, but also to further categorize the non completers according to their behaviour and to find the reasons that brought them to eventually drop out of the program (attrition experiment). Each of these experiments will be described in detail in the sections “Completer-Non Completer Results” and “Attrition Results” respectively.

SVM

In order to run the experiments a general classifier (namely Support Vector Machine, or SVM) was chosen to make predictions regarding students' behaviours and their final results. SVM is a supervised classifier, a family of algorithms that firstly has to be initialized (trained) over a part of the data we want to classify (training set). Once the algorithm has been initialized, it is possible to use the rest of the data (testing set) as an input on which predictions are made. A common splitting for training/testing sets is normally 80%/20%, which is the one that has also been used in this case.

The number of classes within the training set determines the number of classes in which SVM will try to split the testing set. It is therefore important to create the training and testing set in such a way that they correctly represent the overall students' population. A

way to achieve this is by means of a stratified split which returns stratified randomized sets, where both of them are created by preserving the percentage of samples for each class. For example, a stratified split of a dataset consisting of 100 students where 90% are non completers and 10% are completers (9:1 ratio) with a training-testing splitting of 80%/20% will create a training set consisting of 72 non completers and eight completers and a testing set consisting of 18 non completers and two completers. It can be easily seen that in this way the 9:1 ratio of the dataset remains for the training and testing set.

SVM tries to classify the data by maximizing the distance of the closest point of each class to the decision line, which is the line splitting the sets of points into classes. A simple example of such process is shown in Figure 6.

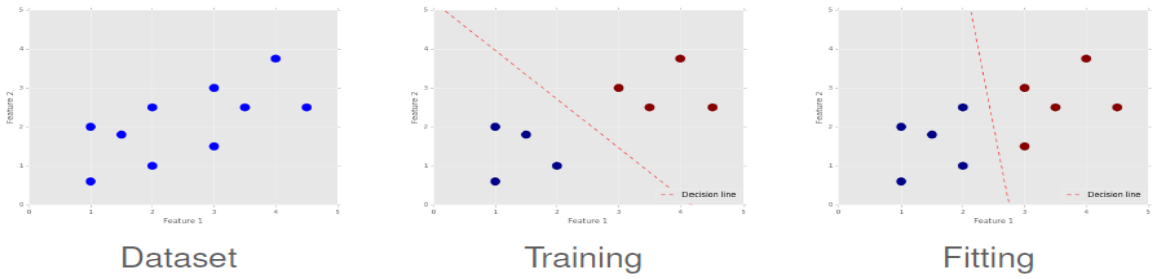

Figure 6: *SVM training and fitting*

In the training phase the SVM gets initialized with the training set and the best fitting line (the one that maximizes the distance of the closest point of each of the two classes to the line itself) is computed. In the fitting phase two new instances get considered and the SVM tries to add them to the predicted class and thus updates the best fitting line. In this simple example only two features are visualized but in general the number of features is significantly higher and since each feature increases the dimension, it is hard to visualize all features together if their number is higher than four.

Feature Extraction

In order to train the SVM classifier, starting from the log data, different features were constructed and used to initialize the algorithm.

From the timestamps, information per student session was constructed. With the initial settings a session had a timespan of 30 minutes. After 30 minutes or more of inactivity a new session gets started. This is a general, common duration for any kind of online session. As a further step different session lengths could be considered (e.g. 15 and 60 minutes) to see if the length of a session can influence the results to some degree. From the per-student session, the timestamp information and the tool type, it was possible to construct various features that can be classified into three main groups:

- General information: number of sessions, number of requests, average number of requests, total session length, average session length, total timespan within clicks, average timespan within clicks, active days, average requests per day, average requests per active day
- Requests per week: total number of requests for each week (from W1 to W8)
- Requests per tool: total number of requests for each tool (only considering the 8 tools mentioned before)

A description of all the features is shown in Table 4.

Feature	Description
Number of Sessions	Total number of (30 minutes) sessions
Number of Requests	Total number of requests (a request is a single entry in the log file)
Average Requests per Session	Average number of requests per session (Total Requests/Total Session)
Session Length	Total length of the whole session
Average Session Length	Average length for session (Total length/Total Session)
Total Timespan	Sum of timespans within consecutive clicks
Average Timespan	Average timespan within two consecutive clicks
Active Days	Total days in which the student interacts (at least 1 request) with the platform
Average Requests Per Day	Average number of requests per day (Total Requests/MOOC Length)
Average Requests Per Active Day	Average number of requests per active day (Total Requests/Active Day)
Assessment	Total number of requests for Tool Assessment
Assignment	Total number of requests for Tool Assignment
Course Board	Total number of requests for Tool Course Board
Evaluation	Total number of requests for Tool Evaluation
File Storage	Total number of requests for Tool File Storage
Forum	Total number of requests for Tool Forum
Learning Content	Total number of requests for Tool Learning Content
Peer Evaluation	Total number of requests for Tool Peer Evaluation
Number of Requests W1 to W8	Total number of requests for each weeks (8 weeks in total, so 8 different features)
Aggregate up to W7 of all of the above	E.g. AssessmentW3 means sum of Assessment up to W3

Table 4: *Description of all the features*

Interesting and valuable to notice is, that this approach is relatively general and not domain dependent. In fact, the information that was used to create the listed features just required the timestamp and tool information, which is information that every MOOC's log should offer.

Evaluation Metric

Before discussing the experiments and the obtained results it is necessary to give a brief explanation about the metric used to evaluate the quality of the prediction. A widely used metric is the F1 Score which is defined in relation to Precision and Recall:

- Precision: Is the fraction of retrieved instances that are relevant
- Recall: Is the fraction of relevant instances that are retrieved
- F1 Score: Is the harmonic mean of Precision and Recall

With this definition, the F1 Score is bounded between 0 and 1, where 0 represents totally wrong predictions, while 1 represents perfect predictions. Practically, Precision, Recall and finally the F1 Score are computed for each class singularly and then the average over the all classes is getting calculated.

Completer-Non Completer Results

In this first experiment the goal was to predict which students would complete the MOOCs and which would eventually drop out. Thus, it was a classification experiment with only two classes, completers and non completers.

Table 5 presents the obtained results for each week, where each week also includes the information regarding any previous week (e.g. week 4 will include all the information from the start of the MOOC until week 4; week 2 will include only information regarding the first 2 weeks and so on).

Analyzed weeks		Result		
		Precision	Recall	F1 Score
Week 1	Completer	0.75	0.60	0.67
	Non Completer	0	0	0
	Average	0.62	0.50	0.56
Week 2	Completer	0.7	0.88	0.78
	Non Completer	0	0	0
	Average	0.51	0.64	0.57
Week 3	Completer	0.60	0.55	0.57
	Non Completer	0.29	0.33	0.31
	Average	0.49	0.47	0.48
Week 4	Completer	0.67	0.90	0.77
	Non Completer	0.60	0.25	0.35
	Average	0.64	0.66	0.61
Week 5	Completer	0.76	0.57	0.65
	Non Completer	0.44	0.67	0.53
	Average	0.65	0.60	0.61
Week 6	Completer	0.82	1	0.90
	Non Completer	1	0.58	0.74
	Average	0.88	0.86	0.85
Week 7	Completer	0.88	0.91	0.89
	Non Completer	0.83	0.77	0.80
	Average	0.86	0.86	0.86
Week 8	Completer	0.92	1	0.96
	Non Completer	1	0.85	0.92
	Average	0.95	0.94	0.94

Table 5: *Obtained results for each week*

As expected, the more weeks are getting considered the better the prediction becomes. From week 6 it is already possible to obtain a F1 Score higher than 0.8. On the other hand the fewer weeks are getting considered the worse the score turns out. This is due to less information available when constructing the features.

Moreover, for some analysed MOOCs the first interactions happened after two/three weeks from the official start of the course. Thus, when running the experiment focusing only on the initial weeks, the features coming from these particular MOOCs could have potentially caused a worsening of the overall predictions. This shows how MOOCs with a different organization and course structure could potentially represent a problem for such predictions (at least when considering only the initial weeks while building up the features).

Attrition Results

A further classification experiment that was conducted focused on attrition. Attrition can be defined as the decline in the number of students from the beginning to the end of the MOOC. Attrition shouldn't be simply considered as a negative aspect, because the decision of students to drop or their failing in the MOOCs could be due to different causes. Rather than just completing a MOOC students might be interested in only covering some topics discussed in a course. Other students may simply enrol to get a

preview of a course in order to gain an understanding of the discussed topics. On the other hand, students who were interested and tried to successfully complete the MOOC but failed, represent the other side of the medal. Therefore, as already discussed in chapter "IV. ATTRITION AND RETENTION IN ONLINE-LEARNING AND MOOCs", it is possible to differentiate two kinds of attrition: healthy and unhealthy attrition.

In order to determine which students dropped out due to healthy or unhealthy attrition the surveys were taken into account. According to the answers the students provided, it was possible to split the group "non completer answered survey" into healthy and unhealthy attrition. Moreover, a third class constituted by the completers was considered during the experiment.

It is important to notice that this classification was only performed with those students who answered the survey, while no information was known for those students who didn't return the survey, who were actually the majority. Table 6 summarizes the number of students per MOOC according to their labelling.

MOOC	Classes	Amount	MOOC	Classes	Amount
Android	Completer	77	Digital Interactive TV	Completer	63
	Healty	46		Healty	16
	Unhealty	38		Unhealty	4
	Unlabeled	432		Unlabeled	38
Client Attention	Completer	60	Elearning	Completer	81
	Healty	6		Healty	27
	Unhealty	3		Unhealty	12
	Unlabeled	22		Unlabeled	125
Cloud Based Learning	Completer	121	Elearning Tools	Completer	101
	Healty	39		Healty	20
	Unhealty	15		Unhealty	18
	Unlabeled	104		Unlabeled	178
Cloud Based Learning (Edu)	Completer	99	Medical Urgency	Completer	49
	Healty	16		Healty	6
	Unhealty	13		Unhealty	4
	Unlabeled	54		Unlabeled	59
Cloud Based Learning (Tools)	Completer	131	User Experience	Completer	62
	Healty	74		Healty	28
	Unhealty	22		Unhealty	10
	Unlabeled	90		Unlabeled	89
Community Manager	Completer	320	Global	Completer	1164
	Healty	78		Healty	356
	Unhealty	59		Unhealty	198
	Unlabeled	364		Unlabeled	1495

Table 6: *Number of students according to labelling*

In Table 6 it can be observed that unlabelled students are almost in every MOOC higher or equal to the total amount of labelled students. Such a scenario introduces some problems. Firstly, although it could still be possible to run a classification experiment using the labelled students as a training set and the unlabelled ones as a testing set, the splitting will be similar to 50%/50%, which is not a proper setting, because normally a splitting of 80%/20% would be used. This will lead to a training phase for the SVM with a training set consisting of too few students compared to the whole, with subsequently

bad prediction results. Besides, without knowing the true label for the unlabelled students, it will not be possible to precisely evaluate the results of such predictions. Because of this situation it was only possible to run the experiment and evaluate the results using the labelled students only. In doing so, all but the students who did not return the survey were considered. The testing and training sets were constructed with a normal 80%/20% splitting and the classifications were evaluated in terms of the F1 Score. Table 7 shows the evaluation of the classifications.

MOOC	Attrition Classes	Precision	Recall	F1 Score	MOOC	Attrition Classes	Precision	Recall	F1 Score
Android	Completer	0.94	1	0.97	Digital Interactive TV	Completer	0.92	0.92	0.92
	Healty	0.50	0.67	0.57		Healty	0.50	0.67	0.57
	Unhealty	0.25	0.12	0.17		Unhealty	0	0	0
	Average	0.65	0.70	0.67		Average	0.79	0.82	0.81
Client Attention	Completer	0.83	0.83	0.83	Elearning	Completer	1	0.94	0.97
	Healty	0	0	0		Healty	0.75	0.60	0.67
	Unhealty	0	0	0		Unhealty	0.25	0.50	0.33
	Average	0.71	0.71	0.71		Average	0.88	0.83	0.85
Cloud Based Learning	Completer	0.96	0.96	0.96	Elearning Tools	Completer	0.94	0.85	0.89
	Healty	0.75	0.75	0.75		Healty	0.43	0.75	0.55
	Unhealty	0.33	0.33	0.33		Unhealty	0.33	0.25	0.29
	Average	0.86	0.86	0.86		Average	0.78	0.75	0.76
Cloud Based Learning (Edu)	Completer	0.88	0.75	0.81	Medical Urgency	Completer	1	0.90	0.95
	Healty	0.33	1	0.50		Healty	0.33	1	0.50
	Unhealty	0	0	0		Unhealty	0	0	0
	Average	0.72	0.69	0.68		Average	0.86	0.86	0.83
Cloud Based Learning (Tools)	Completer	0.96	1	0.98	User Experience	Completer	1	1	1
	Healty	0.77	0.67	0.71		Healty	0.60	0.50	0.55
	Unhealty	0.20	0.25	0.22		Unhealty	0	0	0
	Average	0.83	0.83	0.83		Average	0.78	0.75	0.76
Community Manager	Completer	0.97	0.98	0.98	Global	Completer	0.96	1	0.98
	Healty	0.71	0.62	0.67		Healty	0.77	0.67	0.71
	Unhealty	0.54	0.58	0.56		Unhealty	0.20	0.25	0.22
	Average	0.87	0.87	0.87		Average	0.83	0.83	0.83

Table 7: *Obtained evaluation results for the attrition experiment*

In this experiment the prediction was made using log information from the whole duration of each MOOC (eight weeks). In some cases the results per MOOC are quite poor, like for example in the case of “Client Attention”, “Digital Interactive TV” and “Medical Urgency”. This could potentially be due to the small number of students considered, which is directly dependent on the number of students who returned the survey: The lower the number of returned surveys, the lower is the number of students considered in the prediction experiment and in general a small set could create some difficulties for the SVM. As a matter of fact, those were all MOOCs with a low level of non completers and thus returned surveys. However, when the dimension of the considered dataset grows, the F1 Score grows consequently. This is the case when considering all the MOOCs together (see: “Global” row in the table).

It is also interesting to notice that the scores for the completers are always higher when compared to those having healthy and unhealthy attrition. This could suggest that students in the group completers are more similar with each other than the students assigned to a healthy and unhealthy attrition group. Thus, the SVM can easily detect those but struggles with the other two classes. Another possible explanation could be that the considered features are good to differentiate completers from non completers, but not sufficiently good to further split the non completers into the groups of healthy and unhealthy attrition. Because of these limitations a classification involving sub-

grouping healthy attrition into enthusiast learner, content learner and restricted learner would result in noticeably worse outcomes than the ones presented here.

Feature Ranking and Selection

The goal of this experiment was to analyse whether it is possible to weight features according to their importance for the prediction. Obtaining a higher accuracy by using a smaller set of features makes the model more flexible and potentially adaptable to other domains as well. The smaller the set of features, the easier it can be shaped for different applications. In this experiment a feature selection step was introduced before running the classifier. In the preprocessing phase common univariate, statistical tests were run and the 10% top performing features were kept and used in training and predicting. As usual, the training set was constructed with 80% of the whole examples. The classification was done by only considering the first half of a MOOC's duration for two main reasons. Firstly, considering the second half of a MOOC would not have been so revealing, since at this stage normally most of the non completers have already dropped out. Thus, it is easier for the classifier to differentiate between the classes. Secondly, the goal is generally to differentiate completers and non completers while the MOOC is still ongoing. The following Table 8 shows the selected features for each MOOC, subdivided into “Aggregate”, “Tools” and “Others”.

MOOC	Select 10%
Android	Aggregate: ActiveDaysW4, AveReqPerTotalDayW4, TotReqW4, EvaW4, TotReqW4 Tools: AssignmentsW4 Other: AveReqPerTotalDay, TotReq
Client Attention	Aggregate: ActiveDaysW2, ActiveDaysW3, ActiveDaysW4 Tools: Other: RequestW1, TotReqW1, AveReqPerTotalDayW1, ActiveDaysW1, ActiveDays, TotSess
CloudBased Learning	Aggregate: ActiveDaysW4, AveReqPerTotalDayW4, TotReqW4 Tools: Other: ActiveDays, ActiveDaysW4, AveReqPerTotalDay, RequestW4, TotReq, TotSess
CloudBased Learning Edu	Aggregate: ActiveDaysW2, ActiveDaysW3, ActiveDaysW4, TotReqW2, TotReqW3, TotReq4 Tools: Other: ActiveDays, AveReqPerTotalDay, TotReq
CloudBased Learning Tools	Aggregate: ActiveDaysW4, AveReqPerTotalDayW4, TotReqW4 Tools: Other: ActiveDays, ActiveDaysW4, AveReqPerTotalDay, RequestW4, TotReq, TotReqW4, TotSess
Community Manager	Aggregate: ActiveDaysW3, ActiveDaysW4, LEARNINGCONTENTW4, ASSESSMENTW3, ASSESSMENTW4, EVALUATIONW4 Tools: Other: ActiveDays, ActiveDaysW3, ActiveDaysW4
Digital Interactive TV	Aggregate: ActiveDaysW4, ASSESSMENTW4, PEEREVALUATIONW4, LEARNINGCONTENTW4 Tools: Other: ActiveDays, ActiveDaysW4, AveReqPerActiveDayW4, AveReqPerTotalDayW4, RequestW4, TotReqW4
ELearning	Aggregate: ASSIGNMENTSW4, TotReqW4 Tools: ASSIGNMENTSW4, FILESTORAGEW4 Other: ActiveDaysW4, AveReqPerTotalDay, AveReqPerTotalDayW4, RequestW4, TotReqW4
ELearning Tools	Aggregate: ActiveDaysW4, TotReqW4, EVALUATIONW4 Tools: EVALUATIONW4 Other: ActiveDays, ActiveDaysW4, AveReqPerTotalDayW4, RequestW4, TotReqW4
Medical Urgencies	Aggregate: TotReq4, EVALUATIONW4 Tools: EVALUATIONW4, FILESTORAGEW4, LEARNINGCONTENTW4 Other: ActiveDaysW4, AveReqPerTotalDayW4, RequestW4, TotReqW4
User Experience	Aggregate: AveReqPerTotalDayW4, TotReqW4, PEEREVALUATIONW4 Tools: PEEREVALUATIONW4 Other: AveReqPerTotalDay, AveReqPerTotalDayW4, RequestW4, TotReq, TotReqW4

Table 8: *Selected features for each MOOC*

The features coloured in red are those referring to Week 4 (the last considered week within the experiment) and they are predominantly selected in almost all considered MOOCs. This indicates that the closest time window, for the moment which is

considered, carries a lot of useful information for the classifier. Moreover, the earlier the week, the less performing it seems to be. On the other hand, the presence of aggregate features, spanning over all considered weeks, indicates that the overall temporal behavior carries precious information when distinguishing different learners. Some of these aggregate features are almost always picked, as for example “ActiveDaysW4”, “AveReqPerTotalDayW4” and “TotReqW4”. Thus, a general pattern seems to exist and should be further analysed in the future.

VI. BEST PRACTISES

This chapter introduces several valuable methods and suggestions for MOOCs, with a special focus on improving retention, which are derived from a conducted survey with MOOC creators, from practical experiences and the elaborated literature.

Survey with MOOC Maker Partners

An online survey with some MOOC Maker partners of the consortium has been conducted with the goal to collect information of MOOC creators concerning attrition and retention aspects, statistical data, as well as best practises and recommendations about MOOCs. The survey was sent to eight different partners. Five out of eight partners successfully completed the survey. The survey was subdivided into the following, subsequent groups:

0. Declaration of Consent
1. General Questions
2. General Questions about the Lab's or Institution's Experience
3. Information about Lab's or Institution's offered MOOCs
4. Issues and Recommendations
5. Closing

Concerning the five completed surveys, one partner stated having no experiences in creating MOOCs, not actively offering MOOCs and also not wanting to offer some in the future, but deploying online courses for only residential students instead, which shall be continued in the future. Four partners have lots or at least a few experiences in creating MOOCs. In more detail, one partner has a few experiences, two partners have medium and one partner has lots of experiences, as seen in Figure 7.

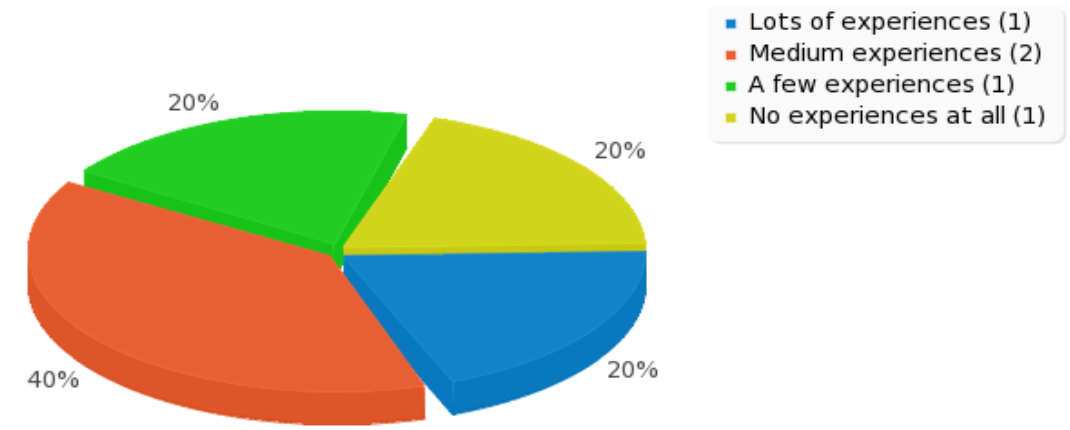

Figure 7: *Lab's or institution's experiences in creating MOOCs*

In the following, only those four partners will be considered who have experiences in creating MOOCs:

One institute is currently not actively offering MOOCs but will, all others do and will continue it. Two out of four are also offering online courses for residential students only and will continue in doing so, whereas two partners do not restrict the MOOCs to residential students only.

The following Tables 9 to 11 shall give statistical data of the four partners. The tables are constructed according to the groups mentioned above. Not all questions are stated in the tables for ensuring clarity. For reasons of completeness, the corresponding primary questions of the conduced survey are listed in the APPENDIX I.

1. General Questions

	Partner A	Partner B	Partner C	Partner D
Experiences in creating MOOCs	A few experiences	Medium experiences	Medium experiences	Lots of experiences
Currently actively offering MOOCs	No, but we will.	Yes, and we will continue.	Yes, and we will continue.	Yes, and we will continue.
Currently actively offering online courses for residential students only	Yes, and we will continue.	No	No, and we don't want.	Yes, and we will continue.

Table 9: *General questions*

2. General Questions about the Lab's or Institution's Experience

	Partner A	Partner B	Partner C	Partner D
Number of years creating and offering MOOCs	2	1	3	4
Number of different, created MOOCs	6	10	4	14
Total number of offered MOOCs	8	10	12	18
Used MOOC Platforms	Open Education powered by Blackboard	Coursera, Open edX	iMOOC, EMMA	iMOOC, EMMA
Types of offered MOOCs	xMOOC	cMOOC	iMOOC	xMOOC

Table 10: *General questions about the lab's or institution's experience*

3. Information about the Lab's or Institution's offered MOOCs

	Partner A	Partner B	Partner C	Partner D
Duration of provided/ created MOOCs in weeks	4	5-6	6-8	5-9
Anybody can register to these courses	Yes	Yes	Yes	Yes
Languages of the offered MOOCs	Spanish	Spanish	Portuguese, English	Spanish, English
Fields/Subjects of the offered MOOCs	Education, Social Sciences	Mathematics, Engineering, Education	Education, Environmental Sciences, History	Physics, Art, Education, Engineering, Humanities, Computer Sciences, Law
Introductory courses offered	Yes	Yes	Yes	Yes
Sophisticated courses offered	No	Yes	No	No
Prerequisites for joining a course	No	No	No	Yes/No (Depends on the MOOC)
Additional course material available	No	Yes	Yes	Yes/No (Depends on the MOOC)
Final exam	No	Yes	Yes (In case a participant wants to earn ECTS credit points; No otherwise)	Yes
Group work	No	Yes/No (Depends on the MOOC)	No	Yes/No (Depends on the MOOC)

	Partner A	Partner B	Partner C	Partner D
Self-paced elements	Yes, most blocks of the courses are designed this way.	Yes, the whole courses are designed this way.	Yes, there are some elements integrated in the courses.	Yes, most blocks of some courses are designed this way.
Discussion Forum	No	Yes	Yes	Yes
Participants can contact an educator	Yes	Yes/No (Depends on the MOOC)	Yes	No
Certificate after completion	No	Yes/No (Depends on the MOOC)	Yes	Yes
Virtual badges or something similar offered	No	No	Yes	No
Motivational elements used	No	N/A	"Bootcamp[4]" module at the beginning of the course	Simulations, interactive activities
Fee for participation	No	No	No	No (Yes, only for receiving a verified certificate)
Courses based on books	No	Yes, some of them.	Yes, some of them.	No
Limited number of participants	No	No	No	No
Students partially grade other students	Yes/No (Depends on the MOOC)	Yes/No (Depends on the MOOC)	Yes	Yes/No (Depends on the MOOC)
Mandatory course for some participants	No	No	No	No

Table 11: *Information about the lab's or institution's offered MOOCs*

4 The courses started with a so-called "bootcamp" module, which can last one or two weeks and gave the participants the opportunity to get acquainted with the environment.

The used learning activities offered in the MOOCs of the partners were video content, video lectures, quizzes, peer reviews, animations, simulations and external tools which were integrated into the course to enhance learning.

Concerning the question how the partners plan and create MOOCs in terms of the team members and the involved effort, etc., three responses were received. Partner B stated, that three people participate in creating the MOOC, additionally the teacher and a teaching assistant are involved. Partner C mentioned that a team of teachers designs the MOOC, which gets assisted by a pedagogical design expert and a technological infrastructure expert. Another team is responsible for producing the video material. On average a MOOC is designed and produced in three months. Furthermore, a team of tutors is involved in supporting the learners during the course. It was mentioned, that the institution has produced its MOOCs using EU-funded means. Partner D stated, that the university has an annual call for MOOCs. The teaching staff, who is interested in creating a MOOC, needs to make a proposal. In case it gets selected, they receive support from a special unit on audio-visual production, IT and management. The temporal resources depend on a particular MOOC, but the workload is quite high for most of them.

4. Issues and Recommendations

Mentioned issues were that it is actually hard to find teachers who want to use a MOOC. Additionally, it was noted that it is difficult to deal with the involved teachers during the process of producing a MOOC. A commonly observed issue is the low completion rate in MOOCs, where one partner subsequently stated that a specific number of learners have a lack in engagement. Furthermore, it was mentioned that some main issues concern the required research on the design of an effective learning experience. Problems which were linked with the use of social media in a formal institutional context also showed up. Furthermore, making MOOCs is considered time consuming, where most of the teams do not meet the deadlines for delivering course materials. In some courses a low activity in the forums could be observed. Furthermore, it was explicitly stated that the dropout rates tend to be very high. Main issues which should be faced concern time issues and the lack of resources. It was also noted that a way should be found to revise self-paced MOOCs.

Improvements which could be made in the courses cover e.g. the design of the MOOCs in order to make them more active, the increased usage of gamification elements and including more engaging and interactive activities. Some errors within two rounds of the same course could also be corrected. Furthermore, a proper marketing of some courses could be achieved as well.

Further potential improvements are that the processes of producing MOOCs need to be better defined and a better teacher training for MOOC design and delivery is needed. Putting more effort into the marketing of the courses and implementing more engaging and interactive activities are improvements for the future. Additionally, it was stated

that more strict deadlines for the teaching staff should be defined, where a course shall not be announced until a good part of the learning material is actually created.

A positively mentioned experience is that the teachers who once finished MOOCs are really happy about that. Other positive experiences concern the learners' high interactivity, their consistent engagement, very high satisfaction rates, very successful courses with good results on the final surveys and satisfied learners.

One recommendation mentioned in section 3 of the survey was that learning should be enhanced by the creation of artefacts like texts, videos, presentations, slidecasts, mind maps, mash-ups, etc., which should be freely accessible online. These artefacts shall demonstrate the learner's knowledge and competencies regarding the studied material. Another recommendation mentioned in the same section of the survey was to introduce so-called “bootcamp” modules, which last for one or two weeks and give the participants the opportunity to get acquainted with the whole, new environment of a MOOC, covering the spaces, tools and services, as well as the processes of work and communication, which will be used in the ongoing course.

Activities to support retention and overcome the high dropout rate are using self-regulated learning approaches and trying to foster these skills, together with embedding many interactive activities in the MOOCs to engage the learners. Carefully produced videos are other means used to further engage the learners. It was also mentioned that the tutors and the team of facilitators are an important factor for achieving retention. Special pedagogical design elements like gamification can also improve retention rates.

One mentioned best practice for creating MOOCs was the recommendation to check the iMOOC pedagogical model. It was stated that the creation of a MOOC is approximately a one-year project in which teachers get firstly trained to understand what a MOOC is, the platform they are going to use and the available resources. There are regular meetings with them to see their advances and discuss potential improvements of the course. Another best practise is that the involved team pushes and motivates the teachers in the production process. Furthermore, workshops with the teachers involved in the production process are recommended.

Meta Analysis

This subsection is intended to present further best practices found in literature.

Stephanie Richter from the Northern Illinois University elaborated in Richter (2013) some general tips for designing a MOOC, which resulted from the experiences of creating the MOOC “Perspectives on Disability”. One important aspect is to clearly determine the targeted audience, which requires that the course gets designed so that the audience's needs are met. Important factors range from the course's content up to the used language and from the teaching strategies to the assessments. All those factors shall be designed according to the given requirements and the supposed pre-knowledge of the primary group of targeted participants. As their course was designed for individuals with little or no prior-knowledge about the content and it was primarily intended for middle school students, the used language was kept simple and some introductory content was prepared. MOOCs can be accessed worldwide, thus, the prerequisites should be clearly

stated beforehand and some supplementary learning resources can be provided as well. Another mentioned best practise for creating MOOCs is to firstly define the learning outcomes for a course and to initially describe what the students will actually learn by attending this course. Subsequently, the corresponding learning activities can be developed in order to achieve those goals, by creating e.g. appropriate learning activities and assessments which shall reflect whether the students reached the desired outcomes. Furthermore, another noteworthy recommendation is that the number of the desired outcomes should be appropriate to the length of the course. The length and the timing of a course are vital ingredients. Despite the fact that traditional courses commonly last a whole academic semester, MOOCs don't have this necessity. Richter (2013) stated that there is currently no research for determining the ideal length for MOOCs, but must observed courses lasted somewhere between four and eight weeks. A few MOOCs could be found which also reached a length of 10 to 12 weeks. It was mentioned that assessments are no absolute requirement for MOOCs. Many MOOC creators focus on group works and collaborations rather than individual tests and written exercises in order to enable the course for a massive crowd, otherwise the course might not scale any more and the assessments would become impractical. Automated grading should be used wherever possible. Other best practices like funding, promoting, planning MOOCs and designing a course in teams are also mentioned, which are rather out of the scope of this work but recommended for MOOC creators (Richter, 2013).

Dave Hallmon is an Instructional Designer and Strategist who outlined in Hallmon (2014) some best practices for MOOCs. His design considerations are referred to a mini lecture[5] of Tina Seeling, who is a professor of the Practise in the Department of Management Science and Engineering (MS&E) at Stanford University and a faculty director of the Stanford Technology Venture Program (STVP) ("Tina Seelig's Profile | Stanford Profiles," n.d.). One important fact mentioned is that students who attend MOOC programs often want to continue their education, even though some of them already have a college degree or more advanced degrees. An issue which should be considered is that the MOOC participants have very different experiences in terms of technical sophistication. Some have a lot of practise in e.g. using technical tool, making videos, collaborating with others, who are also coined as "Digital Natives". In contrary, some do not bring along those skills, which means there are students who will struggle with these challenges and will have difficulties in completing the corresponding MOOC. Thus, MOOCs should be designed in a way that these barriers are rather easy to overcome for a large, diverse group of potential learners. This can e.g. be achieved by considering usability aspects, embedding introductory lessons, providing tutorials for the tools, etc., which will result in a reduced dropout of the students in the very beginning. It is noteworthy that all registrants have different reasons and motivations for attending a MOOC. For a specific portion of students MOOCs represent a very important way of learning, which will also be reflected in the engagement and persistence of those students. According to Seeling, other students are more like "tourists", who are rather just checking out what is behind these online courses and auditing a specific course. It is also supposed that there are participants who just register in order to retrieve some

5 https://stanfordconnects.stanford.edu/watch/crash-course-creativity-more-44000-students (retrieved on May 29th, 2016)

specific content they are interested in, without the initial intention to finish the MOOC. Others could sign up for the course but having very low commitment to it, which may lead to the fact that they will not even login to the MOOC another time. Some also have the intention to proceed in a MOOC, hoping having more time later on. Furthermore, it is noted that students need clear expectations for each integrated activity. This is a reasonable argument, because in general in online learning an instructor is not readily available. In brick and mortar learning environments the instructor can be easily asked e.g. after class. In case of MOOCs, messages must be sent where the students must wait for a qualified answer, which is sometimes seen as a time consuming and cumbersome approach (Hallmon, 2014).

In the following, a list of highlighted recommendations for designing an online course will be presented, which shall improve the learning experience of the MOOC participants. These suggestions have been replicated from Hallmon (2014), which were initially presented in the mini lecture of Seeling, referenced above:

- Online instructors should personally email their students prior to the course start date to remind, welcome and extend.
- Online students should introduce themselves to the class in an innovative way other than just by using a thread in the general discussion area to get the creative juices flowing and a means of connection beyond superficial facts.
- Online students should receive the course's content each week through a well-thought-out and well-produced 5 min video lecture, especially in a MOOC and doubly for cMOOCs.
- Online lecture material should be supplemented and supported with additional external resources, readings, and online discussions to broaden and deepen the learner's experience with learning the content.
- Online courses should have at least one creative challenge each week, either as an individual or a team to increase the students' engagement within the class and keep them interested.
- Online courses should have major assignments broken into smaller chunks and instructors should provide formative feedback, which will remove ambiguities and help the understanding.
- Online students should publicly upload their creative challenge submissions to the course for public evaluation.
- Online instructors should model how students should provide feedback to their peers.
- Online instructors should give individual assignments first to see who is actually in the course and going to participate and would be of value to a group later in the course.
- Online instructors should setup a collaborative community where students feel comfortable asking questions and receiving answers from the whole class.

- Online instructors should see themselves as facilitators and make sure that they keep the conversation moving forward on the right track.
- MOOCs hosted by a university should provide a "special code" that will allow students and alumni from the hosting university to connect with each other specifically.

More details and supportive arguments can be found in the original article of Hallmon (2014).

Joosten (2013) from the University of Wisconsin-Milwaukee published an article to assist those interested in creating and deploying MOOCs. Herein, guidelines are presented and several questions get stated which MOOC creators shall answer themselves for their requirements. In the following some relevant content will be covered. MOOC creators can also refer to the original article. Again the advice shows up to find out the suitable audience for a given topic provided in a MOOC. Furthermore, it again gets stated that designing and facilitating MOOCs can be done best in a peer or a group. Additionally, it is mentioned that MOOC participants should get assessed regularly and get feedback on their strengths and weaknesses frequently. The learning objectives should always be kept in mind while creating a MOOC and those objectives shall be stated to the participants from the very beginning, even before their official registration, which should reduce early dropouts. Joosten (2013) declares that "*although some MOOCs (MOOCx) tend to focus only on content delivery, a truly effective learning experience requires interactivity, communication and community*", which means among others, that tools for interaction between the participants are an important means and should be implemented in some way. It is also noted that students sometimes have difficulties acclimating to the site of the online course or to other instructional tools, thus, some counter-measures should be implemented. Another mentioned issue is the fact that students frequently face difficulties scheduling the required work in an online course and the corresponding time. As a consequence, it is recommended to provide some technical means to foster students' time management and their organizational skills (Joosten, 2013).

Alario-Hoyos et al. (2014) described in their work recommendations after teaching the course "Digital Education of the future" at the Universidad Carlos III in Madrid, Spain. One recommendation is to choose the platform for the new MOOC either based on institutional agreements or according to the targeted learners. It is noted that sometimes interested latecomers would like to join an already active MOOC course, which results in the recommendation that a flexible time schedule of the course should be defined so that late enrolments are still possible. There was a general agreement on the fact that regular communication between the instructors and the participants is a necessity in order to raise the awareness and keep everyone up to date. In their deployed course the facilitator was explicitly responsible for regular communication, as well as acting as a mediator between the participants and the teaching personnel. Thus, e.g. notifications about pending tasks or latest news were sent out on a regular basis. Another recommendation is to create short video lectures with clear tone, which easily explain the corresponding covered content. In the mentioned course the teaching staff employed videos with a length of about ten minutes. The short video lectures resulted in

overall positive feedback of the MOOC registrants. Providing supplementary, easily understandable learning material for the participants is another recommendation, which complements the teachers' talks, presentations and videos. These materials, e.g. comprised of slides, can also be used to study the content offline. Partially, this approach was already pursued in their MOOC realization, were some modules' videos were enriched by supporting slides. In another module a supplementary book to the course's content was suggested to the participants. The participants' feedback showed that 69% preferred a video format which was based on slides, where a small screen of the teacher can be seen in of the corners of the video. Only 23% of the registrants preferred seeing the teacher in the foreground of the video, without any slides. Regarding students' assessments, it was mentioned that the essentially acquired competences of the participants by attending this course should be explicitly defined. In this case, those competences had been defined beforehand, which was also a recommendation of previous findings in literature. Furthermore, the recommendations stated to define assessments from the very beginning of the course and to clearly state the defined assessment policies as well as how the final scores will be calculated. Providing immediate feedback to taken assessments is also seen as a good practice, so that the participants exactly know their current performance during the course. Assessments like e.g. quizzes can be used to achieve this goal. Social learning is another factor which should be promoted. Despite the fact that supporting several social tools at once in a MOOC environment results in more effort for the educators, it has the advantage that the students can choose their own preferred means for communication and interaction which they feel most appropriate and easy to use. This is a recommendation which rather varies in literature. Sometimes students get irritated by having to decide which tool to use. Another disadvantage can be that the learners could be more isolated by their different choices of social tools. In their supplied online course the forum was the most popular tool for discussions of the participants. This was followed by e.g. the social tools Facebook, a Question and Answer system and Twitter. Furthermore, it is recommended to inform the learners at the beginning of the course to what degree the educators will interact with the students by using social tools, for answering questions, etc. For the taught course a facilitator spent three to four hours a weekday and one hour per day at weekends interacting via social tools. Nevertheless, some participants complained that the support by using social tools was not sufficient. Those complaints were especially distinctive at the beginning of the course (Alario-Hoyos et al., 2014). For further and other details the original work gets referred.

Spyropoulou, Pierrakeas and Kameas (2014) created MOOC guidelines which are based on best practices, covering content for designing and developing suitable educational material, content about the course's curriculum with discussing the structure of MOOCs, its configuration and content about the course implementation, presenting principles for supporting the students' learning process. In order to collect these guidelines six popular, major MOOC platforms got analysed, namely Coursera, FutureLearn[6], Udemy[7], Udacity, edX and Iversity[8]. Furthermore, it was stated that the resulted guidelines had already been used for designing and implementing four MOOCs in the field of Computer Science.

6 www.futurelearn.com

7 www.udemy.com

To guarantee an objective overview, they grouped the best practices according to the aforementioned three main components, namely (1) Course Curriculum Development, (2) Educational Material Development and (3) Implementation. Their best practices resulted by attending respective courses and searching good practices which the creators stated on their platforms. For each of the three main classes defined, the MOOC providers get compared with each other. Some of the best practices are grouped by additional sub-categories, where examples are given, showing which MOOC providers use which techniques. The analysed courses had durations between six and 14 weeks. Some courses had a workload of two to three hours a week, others between five and eight hours. It gets recommended planning a weekly workload between three to four hours, up to ten hours a week, which should be stated in the description of the course. Many recommendations overlap with those stated in this report so far, others cover new insights, as well as fine-grained details, which are also useful. Projects, quizzes and peer assessments are the most commonly used evaluation methods of the six MOOC providers. Some providers make their educational videos accessible via a university's YouTube video channel. Four out of six providers enable downloading the educational material for offline usage. It is noteworthy, that all six platforms provide a student's progress page. The survey points out that it is useful to divide a course into sections including activities where each of them can be completed within one week. It is stated that wiki-systems are also a useful means, which allows participants to submit additional resources and concepts for a course. Three out of six MOOC providers recommend integrating popular social networks like Facebook and Twitter, where announcements can be published and students can fetch updates, state questions or initiate discussions without any delay (Spyropoulou et al., 2014). Many other noteworthy guidelines get mentioned, which makes this work highly recommendable for creating and improving MOOCs.

Gamage, Fernando and Perera (2015) discuss factors which lead to effective MOOCs, considering the participants' perspective. Thereby, the Grounded Theory (GT) methodology got applied and ten dimensions were extracted which result in an effective MOOC, namely interactivity, collaboration, pedagogy, motivation, network of opportunities/future directions, assessment, learner support, technology, usability and content. The research found the unique dimension “network of opportunities/future directions”, where students can apply their acquired knowledge in industrial projects. Furthermore, the networking opportunity shall indicate that some students might collaborate in the future, concerning lifelong learning (Gamage et al., 2015). For further and detailed information the original work gets referred.

Yousef, Chatti, Schroeder and Wosnitza (2014) conducted an empirical examination of criteria in order to assure the design quality of MOOCs and wanted to find out which factors drive a successful MOOC. It was argued, that diverse reports showed a high dropout rate, which is on average about 95% of the course participants, where pedagogical problems showed up as well, which was their reason for investigating the quality of MOOCs' design. Several specific criteria could be identified for designing and implementing MOOCs. The study contained a large survey where learners and professors were involved, considering their MOOC experience. Finally, 74 indicators could be identified, which got grouped into two main dimensions, namely pedagogical and

8 www.iversity.org

technological, which were further scattered over six categories. The categories for pedagogical criteria were called “Instructional Design” and “Assessment” and the categories for the technical criteria were “User Interface”, “Video Content”, “Learning and Social Tools” and “Learning Analytics” (Yousef et al., 2014). The resulted 74 valuable recommendations will not be listed in this report, referring to the original work.

Guàrdia, Maina and Sangrà (2013) elaborated further design principles, by pursuing a pedagogical approach from a learner's perspective.

Additional best-practices, which are not explicitly mentioned in this chapter, can also be found in the previous chapters of this report.

VII. RECOMMENDATIONS AND FINDINGS

Most of the learning settings from past up to present were facing attrition issues. Several methods got deployed for their reduction. Ever since, one very important aspect was that the social interactions between students have a very significant, positive impact on the learning behaviour and the reduction of attrition. This means, that actions should be taken to improve the ways students communicate and interact with each other. MOOC courses should foster on learning in groups, where the participants can push each other onwards. In this way, better performing students can also help worse performing ones in understanding specific content of the course, which consequently will reduce the dropout rate. Various group activities naturally promote learning in groups. The elaborated literature has shown that it is not important to deploy many tools for the same purpose in parallel, it is more important to find unified, simplified tools which provide a good and intuitive way of usage, by considering usability aspects. Discussion forums have proven being a well-accepted means for sharing information and ideas among the participants and also for getting in touch with MOOC instructors. The presented literature has shown that it is recommended to differentiate so-called healthy and unhealthy attrition among participants, which can be further subdivided. These distinctions can be useful for deriving early indicators of potential dropout candidates, where some counter-measures could be initiated. The level of difficulty, the workload and the duration of a MOOC have negative impact on completing a course. A self-paced time-table can be recommended if the course has a higher workload in order to improve its retention rate. In case of a more difficult course students should be allowed to plan their own time-schedule, otherwise a fixed time-table set by an instructor gets recommended. It was found out that final exams and peer assessments have a positive influence on finishing a MOOC. Furthermore, it was mentioned in the presented literature that automated grading significantly increases the learners' satisfaction, which negatively applies to other factors like difficulty, workload and duration. Finding ways to motivate and engage students will lead to more completers. Depending on the MOOC's discipline also wiki-systems can be useful. Interlinking learning content with different, selected and useful internal and external resources can be a good way of improving the learning process and to acquire related knowledge.

Chapter “V. USER MODELS AND USER PREDICTIONS“ resulted in further recommendations: The aim to classify students into either completer-non completer or in relation to attrition and retention, is harder the fewer weeks (or in general time) get considered in the analysis. On the other hand, the earlier it is possible to identify non completers or unhealthy attrition the more time is available for a tutor or professor to take actions in order to put those students back on their right tracks. One important point is therefore the correct assignment of students according to healthy and unhealthy attrition, because the mitigation of the dropout should only focus on the latter group, rather than on the first one.

When only information about the first few weeks (or in general a short amount of time) is considered, it becomes necessary to build up more significant features. In order to do

so, it would make sense to have a graded assignment or an entry test already within the first week of a MOOC. Information as for example “Score”, “Assignment Duration”, “Number of wrong/right answers” would constitute important features that could help to obtain better predictions. The results of the first task are a strong indicator of students' behaviour within the rest of the MOOC's duration. A student failing or getting a lower score in the first task is in fact more likely to drop out or to fail the MOOC in the end.
Encouraging social interactions through forum activities and peer evaluations is also a potentially valid way to mitigate dropout. Students who constantly engage in the MOOC forum, create discussions, reply to other students' questions, show high interest in the MOOC and therefore have higher motivations to succeed. Trying to engage a larger number of students to participate in a forum is advisable, but it is not a trivial task and most of the time it is not easily achievable. However, a highly active forum, even if only animated by a few students, could encourage other students as well, who do not actively participate, to at least spending time reading existing discussions and maybe finding answers to their concerns. Although in this way the forum participation would not be improved, it may however bring some improvements and more knowledge to all students.

Social interactions in the form of group activities represent another interesting solution. Ideally, it would be of great benefit to create balanced groups, in which there is a mix of completers and non completers. In this scenario worse performing students would put more effort into solving an assignment, because of the help and the stimulating presence of the good performing ones. Group activities would also help to create bonds among students which could be helpful during the rest of the MOOC. For example, a student who does not actively use the forum could still directly contact a member of the group where the student was part of in case of needing help or having a question regarding a particular topic.

In chapter “VI. BEST PRACTISES“ it was mentioned that it is important to identify the desired audience for a MOOC and to design the content accordingly so that their needs get met. This ranges from the course's content up to the used language and from the teaching strategies to the assessments. The according pre-knowledge of the target group should influence the design of the course as well. The prerequisites for a course should be clearly stated so that potential registrants already know them beforehand, which will reduce the dropout rate. It gets recommended that the number of desired learning outcomes should be appropriate to the length of the course, which further implies that a course shall not be overloaded with too much content. In general, a course's timing and a course's length represent vital ingredients of a MOOC. Reducing the length of a MOOC may positively affect the retention rate. Most courses have a length of four to eight weeks, which can be seen as good reference values. In order to provide a course for a massive crowd of participants, group work and collaborations should be focused instead of many individual tests and written exercises, otherwise the course will not scale any more and the assessments will become impractical. Furthermore, automated grading should be promoted, so that the students receive early feedback and always know their current performance, which will improve their satisfaction. The diversity of the learners is an aspect which should also be considered in detail. It is recommended to introduce means which allow inexperienced learner to overcome and catch up on their deficits.

Examples would be improving aspects of usability and providing introductory lessons and tutorials. Well-prepared, short video clips have also proven to be a very useful tool for learning. Furthermore, participants should always know beforehand what they will be confronted next in the ongoing course so that they are mentally prepared. This can be again achieved with e.g. short video clips or some text, which outlines a week's content. Supplementary material and external resources can be good means for understanding difficult content. Discussion forums can be used by the participants to further improve the understanding of the course's content, either by active participation or by passively consuming its content. Textbooks are recommended in the course if they are freely available. Handing out a certificate at the end of the course can also affect retention. The completion rate of a course also depends on the topics which get taught.

The conducted survey in section “Survey with MOOC Maker Partners“ has shown that interactive, gamified tasks and engaging assignments or activities are good means for keeping the participants attracted to the course. Another recommendation was that learning should be enhanced by the creation of artefacts, like texts, videos, presentations, slidecasts, mind maps, etc., which should be freely accessible online. These artefacts shall demonstrate the learners' knowledge and competencies regarding the studied material. Their publicly available outcomes can also inspire and motivate other participants. Knowing that some outcomes of assignments will be visible to other students can promote better performance as well. So-called “bootcamp” modules can be introduced for the first or the first two weeks, which shall give the participants the possibility to get acquainted to the whole, new environment of a MOOC, covering the spaces, tools, services and the different processes of work and communication. Further activities to support retention are using self-regulated learning approaches and embedding many interactive activities to further engage the learners. Carefully designing and structuring the content are additional recommendations. Tutors and facilitators play a very important role for achieving a lower dropout rate. In some courses students reported less support as one reason for getting unmotivated over time, which can be seen as one pre-stage of non-completers. The survey also resulted in the recommendation to deploy special pedagogical design elements like gamification and others to further improve the retention rates.

Additional recommendations can also be found in the previous chapters.

Selected Strategies to reduce Attrition

The Tables 12 to 14 are inspired by the literature of the subsection “Retention, Attrition and Persistence“ of chapter “IV. ATTRITION AND RETENTION IN ONLINE-LEARNING AND MOOCs”, where Berge and Huang (2004) defined a classification with relevant variables for attrition and retention, which can be seen in Table Error: Reference source not found. Additionally, some criteria from Park (2007) get corroborated, which can be seen in Figure Error: Reference source not found. The Tables 12 to 14 show selected measures and strategies to counteract high attrition rates in MOOCs, by grouping them based on some aspects of both classification schemes. Table 12 lists selected “personal variables”, Table 13 “institutional variables” and Table 14 “academic variables”. “Personal variables” shall reflect strategies of a learners' perspective, which encompass demographic aspects (like age, gender, ethnicity, economic status and parental

expectation), individual aspects (like academic skills and abilities, time management, motivation, goals and commitment), prior educational experiences (like academic achievements and prior schooling experiences) and external factors (like scheduling conflicts, family issues, financial problems, managerial support and personal issues). “Institutional variables” comprise bureaucratic aspects, like mission, policy, institutional funding and budgeting, personnel, planning, advertising, describing and the process of creating MOOCs. “Academic variables” subsume educational design aspects and teaching approaches, the structural design and internal factors (like social aspects, social and academic integration and usability issues). Some best practices were considered being rather out of the scope of this work and only got slightly discussed in this report but in more detail in the referenced literature of this work. One such example would be in the field of “institutional variables”, discussing recommendations for e.g. funding, budgeting, promoting, planning MOOCs and designing a course in teams, where Richter (2013) and others present diverse strategies. For each mentioned strategy a corresponding rationale is added in order to corroborate its underlying idea.

Personal Variables

Strategy	Rationale
Provide an appealing and gamified course, where the participants can collect badges or something similar and obtain a certificate after completion.	Raising and obtaining students' motivation is an important factor for successful MOOCs. In case the students are and stay motivated they will continue the course with high probability.
Introduce appropriate ways of motivating and engaging students. Think of adding engaging and interactive assignments and activities. Additional alternation can be provided by e.g. small quizzes.	Such assignments will keep the participants attracted to the course and these mentioned measures will lead to more completers.
Differentiate the participants according to healthy and unhealthy attrition and derive early indicators for potential dropout candidates.	By deriving early indicators for potential dropout candidates, the underlying reasons can be extracted and some counter-measures can be accomplished.
Provide supplementary material and interlink learning content with different, selected and useful internal and external resources.	These additional sources of information can improve the learning process. Worse performing students can use the resources in order to catch up on particular topics with personal deficits. Students can improve in those fields where they still have some backlog and difficult content can be reviewed in this way.
Deploy self-paced course modules.	Self-paced modules can give some participants the possibility to accomplish the course and some other duties, like job commitments, as well. Thus, scheduling conflicts can be mitigated.

Strategy	Rationale
“Bootcamp” modules can be introduced for the first or the first two weeks, depending on the course's duration.	These modules shall give the participants the possibility to get acquainted to a new learning environment of a MOOC, which covers e.g. the tools, the services and the various working processes.
Consider the natural diversity of the learners and think of introducing means which allow inexperienced learners to close up to a common knowledge base. Some measures comprise improving aspects of usability, providing introductory lessons and tutorials.	Reaching a common knowledge base at the beginning of a course noticeably reduces difficulties and struggles in the ongoing course, which might otherwise lead to a higher dropout rate.

Table 12: *Selected strategies to reduce attrition, concerning personal variables*

Institutional Variables

Strategy	Rationale
Clearly identify the desired audience for a MOOC and design the content accordingly.	Clearly identifying the targeted audience enables the MOOC designers to plan the course's content so that the audience's needs get perfectly met. This ranges from the course's content up to the used language and from the teaching strategies to the assignments and assessments. The target group's pre-knowledge should also influence its design.
Plan and produce shorter MOOCs. Commonly, a course which lasts somewhere between four and eight weeks is a good guidance level.	Less research has been conducted in determining the ideal length of a MOOC, but most observed courses have a length between four and eight weeks. Thereby, participants also have a closer goal to reach. It could be observed that the duration of a MOOC has a direct, negative impact on completing a course. Furthermore, the course's timing represents a vital ingredient of a successful MOOC.
The prerequisites of an online course should be clearly stated and visible for potential registrants, before the need of registration. The same applies to the course's content.	The potential registrants should be aware of the course's content, they should know what will be covered and what will be expected from them by listing the prerequisites. In doing so, fewer students will just register for curiosity, which makes a student's decision for registering more reasonable. In case students know what they are approaching, they start thinking about it in more detail and will stick to their final decision more strongly.

Strategy	Rationale
The number of desired learning outcomes should be appropriate to the length of the course.	This strategy makes sure, focussing on the main aspects concerning the learning outcomes and shall further guarantee that the course is not overloaded with too much content.
Deploy tutors and facilitators	Tutors and facilitators play an important role for achieving a lower dropout rate and can keep the number of unmotivated students low. Furthermore, they act as mediators between the teaching personnel and the registered students.
Tutors and facilitators should send out notifications about pending tasks, as well as current news on a regular basis.	In this way the participants keep informed and don't miss a deadline by accident.
Make sure to plan sufficient time and resources for creating a MOOC. In the process of creating MOOCs the teachers play an important role and teaching assistants should also be involved. Furthermore, pedagogical design experts and technological infrastructure experts are recommended to support the MOOC developers. Another team can be responsible for creating the video material, etc. Additional support can be supplied by special units in e.g. audio-visual production, IT and management.	Creating MOOCs is a time consuming process, where MOOC creators and the involved teams often face problems in meeting the defined deadlines for delivering the course materials, etc. Thus, it is important to build teams with specific responsibilities and to define a reasonable time schedule, ideally with some backup time. Correct planning will finally pay off by obtaining an accordingly good MOOC.
Plan regular meetings with the teams involved in the process of creating a MOOC.	In the meetings the advances can be seen, problems can be faced and potential improvements of the course can be suggested and discussed.

Table 13: *Selected strategies to reduce attrition, concerning institutional variables*

Academic Variables

Strategy	Rationale
Improve the way students communicate and interact with each other and consider usability aspects.	If the communication and interaction methods get improved, the participants are more likely to use them in case of questions, problems, group exercises, which then reduces the feeling of disconnect and fosters collecting positive impressions.
Foster on learning in groups	Social interaction has a significant, positive impact on the learning behaviour, because students can positively push each other onwards. Better performing students can help worse performing ones in understanding the content, which will consequently reduce the dropout rate.
Create group activities and collaborations	Various group activities naturally promote learning in groups. They enable participants getting to know each other and allow building up connections. Balanced groups with a mix of better and worse performing students would be ideal. Worse performing students would get stimulated by the presence of better performing ones and might invest more time in solving the assignments. Furthermore, group work and collaboration are good means for preparing a course for a massive crowd of participants, which guarantees practical assessments and is one key factor for a well-scaling course.
Use discussion forums among students and instructors	Discussion forums are a well-accepted means for sharing information among participants as well as for getting in touch with MOOC instructors and tutors. Commonly, concerns of participants are relevant for several students, which makes contacting instructors via a forum more appropriate. Forum participation can also encourage others in doing so, or at least increase the amount of students who spend time reading existing discussions and probably finding answers to their concerns. Forums are useful for students in order to further improve the understanding of the course's content.

Strategy	Rationale
Don't deploy different tools for the same purpose and only use the most appropriate ones. Stick to unified, simplified tools, with a good and intuitive usage.	If several tools are used for the same purpose and students can choose the one which they prefer, the information flow could be inconsistent and isolated groups with non-shared knowledge could show up. Furthermore, the more tool are deployed the more maintaining is required. Additionally, some participants might not be sure about which tool to use.
Find a moderate level of difficulty and a balanced workload. Deploy self-paced time-tables in case of courses with a higher workload. Courses with a lower workload can have a fixed time-table. A recommendation is to plan a weekly workload between three to four hours, up to ten hours a week.	The level of difficulty and the workload have a negative impact on completing a course. Self-paced elements give the participants more freedom in planning and can improve retention rates in courses with a higher workload.
Enable downloading educational material.	This allows learning some content in an offline-phase.
Integrate automated grading techniques	Automated grading significantly increases the learners' satisfaction because the participants get rapid, early and continuous feedback and always know their current performance, which has a positive influence on finishing a MOOC. This is also applicable for courses with a high workload, a high level of difficulty or for courses having a longer duration.
Consider providing short video clips, which cover some of the main concepts	Video clips have proven being a useful means of learning because different sensory perceptions of a human get addressed.
Let the participants know beforehand what they will be confronted next in the ongoing course, which can be achieved with e.g. short video clips or some text, which outlines the week's content.	This strategy guarantees that the participants get the feeling of being informed about the schedule of a course, so that they are mentally prepared, which gives certainty.
Let the students create some artefacts like texts, videos, presentations, mind maps or something similar and make them accessible to all course participants.	These artefacts will reflect the learners' knowledge and competencies which were imparted in the course. Publishing the results will inspire and motivate other students and it can promote participants for producing better results.
Design and structure the course carefully	Well elaborated, structured and constructive learning material with a common thread will improve the students' understanding and learning process.

Strategy	Rationale
Consider integrating peer assessments	Peer assessments have shown a positive impact on finishing a MOOC. It is one means of ensuring a sustainable MOOC design, because grading the work of students can partially be outsource to the participants themselves. Moreover, the students perceive having more responsibility, which can have positive side effects for both parties.
Provide a student's progress page	Many major MOOC providers offer a personal progress page, which supports the learners in having an overview of the course.
Divide the course into sections, including activities where most of them can be completed within one week.	This is a recommendation of several MOOC providers.
Think of providing a wiki-system	A wiki-system allows students to submit additional resources for the course and improves the learning community.
Consider integrating social networks like Facebook and Twitter	Social networks are well accepted by many users and allow publishing updates. Students can stay up to date, state questions or initiate discussions.
Define the assessments from the very beginning of the course and clearly state the assessment policies, as well as how the final scores will be calculated.	In this case the participants exactly know the process of grading, otherwise confusions and uncertainties may arise.

Table 14: *Selected strategies to reduce attrition, concerning academic variables*

VIII. SUMMARY

MOOCs are popular, because they can reach a broad community of learners. Due to the open approach, a diverse group of learners can be brought together regardless of their social and cultural background. Furthermore, it allows a group of geographically scattered learners to collaborate and learn autonomously. On the other side, by using MOOCs students get faced with issues like isolation and disconnect, which could also be seen in distance learning approaches (Croft, Dalton, & Grant, 2010). In some cases, students fail to self-organize themselves, meaning they are not prepared to manage their own learning and they are facing problems in properly using the given learning tools, as well as completing the prepared learning activities. These issues can be seen contributing to the large dropout rate in MOOCs, where only a small portion of about 10% of the initial registrants successfully complete the course (Chamberlin & Parish, 2011; Daniel, 2012; McAuley, Stewart, Siemens, Cormier, & Commons, 2010). Additionally, the high dropout rate is also a major concern for those who invested time and effort but did not complete the course. There can be teachers and tutors involved, who support the learning groups, assess the assignments and provide feedback. In order to reduce the attrition rate, one suggestion is to enhance the support for the students. Concerning the perspective of Information Communications and Technology (ICT), it is recommended that universities should provide stable and reliable learning platforms and a corresponding network infrastructure which is capable of serving hundreds and thousands of students. On the other hand, the students must also possess suitable, technical means in order to access the tool, materials and activities (Gütl, Rizzardini, et al., 2014).

This summary illustrates some selected aspects concerning MOOCs, where also others exist. On the one hand, it shall be seen as a motivation for diving into the broad world of MOOCs which have lots of future potential. On the other hand, this overview also shows some bottlenecks, drawback and issues, which need to be eliminated over time. This report addresses and discusses the benefits of MOOCs, as well as issues which appear. The main goals are to raise awareness about the high dropout-rates in MOOCs and to come up with different models and solution to counteract this trend.

A literature survey reveals relevant content according to this field. The report starts with some background information, by defining common terms, presenting a selection of prominent learning- and teaching-approaches, as well as recommendations in the literature. In order to come up with sufficient solutions, the historical learning settings get discussed in detail. It shows that issues like attrition or dropout and retention are aspects which were concerns in most of the educational learning settings from the past up to the present. Subsequently, the literature survey approaches the present where MOOCs appear, which will then be discussed in detail and corresponding aspects of attrition and retention will be addressed, by considering the previous findings in the historical section as well, where numerous suggestions show up.

Practical results of experiments over a set of eleven MOOCs offered by University of Galileo get presented. Firstly, a general overview of the MOOCs and an analysis in terms of dropout are given. These are the starting point for classification experiments, which aim to identify classes of students according to their interaction with the platform. The students are classified with respect to their success within the MOOCs and then with respect to their goals and expectations towards the MOOCs, concerning attrition and retention, where the results for both experiments get presented. Furthermore, a discussion about feature selection and ranking gets raised.

Later on, best practices result out of different sources, namely from a conducted survey, the elaborated literature as well as from a performed meta analysis. The conduced survey with MOOC creators further enhances the general insights. It was conducted with MOOC creators of the MOOC Maker Consortium. The performed meta analysis considers further literature, with special focus on sophisticated and profound best practices and recommendations.

Finally, some selected recommendations get presented, which result from all preceding chapters in order to improve MOOCs and counteract aspects of attrition and dropout.

ACKNOWLEDGEMENTS

This deliverable has been co-funded by the Erasmus+ Programme of the European Union, project MOOC-Maker (561533-EPP-1-2015-1-ES-EPPKA2-CBHE-JP).

LIST OF FIGURES

LIST OF TABLES

REFERENCES

Adamopoulos, P. (2013). What Makes a Great MOOC? An Interdisciplinary Analysis of Student Retention in Online Courses. *Thirty Fourth International Conference on Information Systems*, 1–21.

Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Gutiérrez-Rojas, I., Leony, D., & Parada, G. H. A. (2014). Designing Your First MOOC from Scratch: Recommendations After Teaching "Digital Education of the Future." *eLearning Papers, From the Field*, *37*(3), 1–7.

Angelino, L. M., Williams, F. K., & Natvig, D. (2007). Strategies to Engage Online Students and Reduce Attrition Rates. *The Journal of Educators Online*, *4*(2), 1–14. Retrieved from http://eric.ed.gov/?id=EJ907749

Avakian, A. N., & others. (1982). Race and Sex Differences in Student Retention at an Urban University. *College and University*, *57*(2), 160–165.

Balakrishnan, G., & Coetzee, D. (2013). Predicting student retention in massive open online courses using hidden markov models. *Electrical Engineering and Computer Sciences University of California at Berkeley*.

Bean, J. P. (1985). Interaction Effects Based on Class Level in an Explanatory Model of College Student Dropout Syndrome. *American Educational Research Journal*, *22*(1), 35–64. http://doi.org/10.3102/00028312022001035

Bean, J. P., & Metzner, B. S. (1985). A Conceptual Model of Nontraditional Student Attrition Undergraduate. *Review of Educational Research*, *55*(4), 485–540. http://doi.org/10.3102/00346543055004485

Bedford, M. H., & Durkee, P. E. (1989). Retention: Some more ideas. *NASPA Journal*, *27*(2), 168–171.

Berge, Z. L., & Huang, Y.-P. (2004). A Model for Sustainable Student Retention: A Holistic Perspective on the Student Dropout Problem with Special Attention to e-Learning. *Deosnews*, *13*(5), 26. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/citations?doi=10.1.1.129.1495

Bersin, J. (2013). The MOOC marketplace takes off. *Forbes, November*, *30*.

Boudreau, C. A., & Kromrey, J. D. (1994). A longitudinal study of the retention and academic performance of participants in freshmen orientation course. *Journal of College Student Development*.

Bozkurt, A., Akgun-Ozbek, E., Yilmazel, S., Erdogdu, E., Ucar, H., Guler, E., … Aydin, C. H. (2015). Trends in distance education research: A content analysis of journals

2009-2013. *International Review of Research in Open and Distance Learning*, *16*(1), 330–363. Retrieved from http://www.irrodl.org/index.php/irrodl/article/viewFile/1953/3261

Brinton, C., Chiang, M., Jain, S., & Lam, H. (2013). Learning about social learning in moocs: From statistical analysis to generative model. *arXiv*, 1–11. Retrieved from http://arxiv.org/abs/1312.2159

Brooke, J. (1996). SUS - A quick and dirty usability scale. *Usability Evaluation in Industry*, *189*(194), 4–7. Retrieved from http://hell.meiert.org/core/pdf/sus.pdf

Carey, K. (2012). A future full of badges. *The Chronicle of Higher Education*, *58*(32).

Carr, S. (2000). As distance education comes of age, the challenge is keeping the students. *Chronicle of Higher Education*, *46*(23).

Chafkin, M. (2013). Udacity's Sebastian Thrun, godfather of free online education, changes course. *Fast Company*, Online. Retrieved from http://mondrian.die.udec.cl/~mmedina/Desvarios/Files/Sebastian-Thrun-changes-course.pdf

Chamberlin, L., & Parish, T. (2011). MOOCs: Massive Open Online Courses or Massive and Often Obtuse Courses? *eLearn*, *2011*(8), 1. http://doi.org/10.1145/2016016.2016017

Chyung, S. Y. (2004). Hoping to reduce attrition? Follow the SIEME model and investigate motivation-hygiene factors. In *Distance Teaching & Learning conference at Madison, WI*.

Clow, D. (2013). MOOCs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13* (p. 185). New York, New York, USA: ACM Press. http://doi.org/10.1145/2460296.2460332

Cook, D. A. (2007). Web-based learning: Pros, cons and controversies. *Clinical Medicine, Journal of the Royal College of Physicians of London*. http://doi.org/10.7861/clinmedicine.7-1-37

Cormier, D., & Siemens, G. (2010). The Open Course: Through the Open Door--Open Courses as Research, Learning, and Engagement. *EDUCAUSE Review*, *45*(4), 30–32,. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ951238

Croft, N., Dalton, A., & Grant, M. (2010). Overcoming isolation in distance learning: Building a learning community through time and space. *Journal for Education in the Built Environment*, *5*(1), 27–64. http://doi.org/10.11120/jebe.2010.05010027

Daniel, J. (2012). Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. *Journal of Interactive Media in Education*, *2012*(3), 18. http://doi.org/10.5334/2012-18

Downes, S. (2005). E-learning 2.0. *ACM eLearn Magazine*, (October 2005), 1–7. http://doi.org/10.1145/1104966.1104968

Downes, S. (2008). Places to go: Connectivism & connective knowledge. *Innovate: Journal of Online Education*, *5*(1), 6. Retrieved from http://nsuworks.nova.edu/cgi/viewcontent.cgi?article=1037&context=innovate

Downes, S. (2012). *Connectivism and Connective Knowledge: essays on meaning and learning networks*. *National Research Council Canada*.

EDUCAUSE. (2011). THINGS YOU SHOULD KNOW ABOUT …™ MOOCs. *EDUCAUSE Learning Initiative*, 1–2. Retrieved from http://net.educause.edu/ir/library/pdf/ELI7078.pdf

Fowler, G. A. (2013). An Early Report Card on Massive Open Online Courses. Retrieved from http://online.wsj.com/news/articles/SB10001424052702303759604579093400834738972

Fox, A. (2013). From MOOCs to SPOCs. *Communications of the ACM*, *56*(12), 38–40. http://doi.org/10.1145/2535918

Fox, A., & Patterson, D. (2012). Crossing the software education chasm. *Communications of the ACM*, *55*(5), 44. http://doi.org/10.1145/2160718.2160732

Gamage, D., Fernando, S., & Perera, I. (2015). Factors leading to an effective MOOC from participiants perspective. In *2015 8th International Conference on Ubi-Media Computing (UMEDIA)* (pp. 230–235). IEEE. http://doi.org/10.1109/UMEDIA.2015.7297460

Gené, O. B., Núñez, M. M., & Blanco, Á. F. (2014). Gamification in MOOC: Challenges, Opportunities and Proposals for Advancing MOOC Model. In *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 215–220). http://doi.org/10.1145/2669711.2669902

Glass, J. C., & Garrett, M. S. (1995). Student Participation in a College Orientation Course, Retention, and Grade Point Average. *Community College Journal of Research and Practice*, *19*(February 2015), 117–132. http://doi.org/10.1080/1066892950190203

Gleason, B. J. (2004). Retention issues in online programs: A review of the literature. In *Second AIMS International Conference on Management* (pp. 28–31).

Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC Design Principles. A Pedagogical Approach from the Learner's Perspective. *eLearning Papers*, *33*(May), 1–6.

Gütl, C., Chang, V., Hernández Rizzardini, R., & Morales, M. (2014). Must we be concerned with the Massive Drop-outs in MOOC? - An Attrition Analysis of Open Courses. In *Proceedings of International Conference of Interactive Collaborative Learning ICL*.

Gütl, C., Rizzardini, R. H., Chang, V., & Morales, M. (2014). Attrition in MOOC: Lessons Learned from Drop-Out Students. In *Learning Technology for Education in Cloud-MOOC and Big Data: Third International Workshop* (Vol. 446, pp. 37–48). Santiago: Springer. http://doi.org/10.1007/978-3-319-10671-7_4

Hallmon, D. (2014). What are currently the best practices in MOOC instruction, and what does the research suggest can be done to enhance those practices in the future? Retrieved May 29, 2016, from https://www.quora.com/What-are-currently-the-best-practices-in-MOOC-instruction-and-what-does-the-research-suggest-can-be-done-to-enhance-those-practices-in-the-future

Hara, N., & Kling, R. (2000). Students' distress with a web-based distance education course: An ethnographic study of participants' experiences. *Information Communication and Society*, *3*(4), 557–579.

Harris, R., & Nikitenko, G. (2014). Comparing online with brick and mortar course learning outcomes An analysis of quantitative methods curriculum in public administration. *Teaching Public Administration*, *32*(1), 95–107. http://doi.org/10.1177/0144739414523284

Hernández, R., Gütl, C., & Amado-Salvatierra, H. R. (2014). Cloud Learning Activities Orchestration for MOOC Environments. In *Learning Technology for Education in Cloud. MOOC and Big Data* (Vol. 446 CCIS, pp. 25–36). http://doi.org/10.1007/978-3-319-10671-7_3

Hill, P. (2012). Online Educational Delivery Models: A Descriptive View. *EDUCAUSE Review*, *47*, 84–86,. Retrieved from https://net.educause.edu/ir/library/pdf/ERM1263.pdf

Holdaway, X. (2014). Major Players in the MOOC Universe. The Chronicle of Higher Education. Retrieved May 7, 2016, from http://chronicle.com/article/The-Major-Players-in-the-MOOC/138817/

Hyman, R. E. (1995). Creating Campus Partnerships for Student Success. *College and University*, *71*(2), 2–8.

Jaggars, S. S. (2012). Beyond flexibility: Why students choose online courses in community college. In *American Educational Research Association Annual Meeting, Vancouver, Canada*.

Joosten, T. (2013). Ten Questions for MOOC Design. Retrieved May 29, 2016, from http://professorjoosten.blogspot.co.at/2013/04/ten-questions-for-mooc-design.html

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses Massive Open Online Courses. *International Review of Research in Open and Distance Learning*, *15*(1).

Jordan, K. (2015). Massive Open Online Course Completion Rates Revisited : Assessment , Length and Attrition. *International Review of Research in Open and Distributed Learning*, *16*(3), 341–358. http://doi.org/10.13140/RG.2.1.2119.6963

Kay, R. H., & Loverock, S. (2008). Assessing emotions related to learning new software: The computer emotion scale. *Computers in Human Behavior*, *24*(4), 1605–1623. http://doi.org/10.1016/j.chb.2007.06.002

Keegan, D. (1996). *Foundations of distance education*. Psychology Press.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13* (p. 170). New York, New York, USA: ACM Press. http://doi.org/10.1145/2460296.2460330

Knowles, M. S. (1975). Self-directed learning: a guide for learners and teachers. *Selfdirected Learning A Guide for Learners and Teachers*.

Kolowich, S. (2013). Coursera Takes a Nuanced View of MOOC Dropout Rates. *Chronicle of Higher Education*, *59*, A22–A22.

Lee, Y., & Choi, J. (2013). A structural equation model of predictors of online learning retention. *The Internet and Higher Education*, *16*(1), 36–42. http://doi.org/10.1016/j.iheduc.2012.01.005

Lee, Y., Choi, J., & Kim, T. (2013). Discriminating factors between completers of and dropouts from online learning courses. *British Journal of Educational Technology*, *44*(2), 328–337. http://doi.org/10.1111/j.1467-8535.2012.01306.x

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education*, *48*(2), 185–204. http://doi.org/10.1016/j.compedu.2004.12.004

Lewin, T. (2013). Universities abroad join partnerships on the web. *The New York Times*, *20*(2013), 2–21.

Liu, M., Kang, J., Cao, M., Lim, M., Ko, Y., & Weiss, A. S. (2013). Understanding MOOCs as an Emerging Online Learning Tool: Perspectives From the Students. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 2008–2015). Retrieved from https://www.learntechlib.org/p/115173/

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs : A Systematic Study of the Published Literature 2008-2012. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1455/2531

Lowell Bishop, J., & Verleger, M. (2013). The Flipped Classroom : A Survey of the Research. *Proccedings of the Annual Conference of the American Society for Engineering Education*, 6219. http://doi.org/10.1109/FIE.2013.6684807

Mackness, J., Mak, S. F. J., & Williams, R. (2010). The Ideals and Reality of Participating in a MOOC. In *Learning* (Vol. 10, pp. 266–274). Retrieved from http://www.lancs.ac.uk/fss/organisations/netlc/past/nlc2010/abstracts/PDFs/Mackness.pdf

Mak, S. F. J., Williams, R., & Mackness, J. (2010). Blogs and forums as communication and learning tools in a MOOC. *Proceedings of the 7th International Conference on Networked Learning 2010*, 275–285.

Martinez, M. (2003). High Attrition Rates in E-learning: Challenges, Predictors and Solutions. *The ELearning Developers Journal*, (July 14), 1–9. Retrieved from http://www.elearningguild.com/pdf/2/071403MGT-L.pdf

Masters, K. (2011). A Brief Guide To Understanding MOOCs. *The Internet Journal of Medical Education*, *1*(2), 1–5. http://doi.org/10.5580/1f21

McAuley, A., Stewart, B., Siemens, G., Cormier, D., & Commons, C. (2010). The MOOC model for digital practice. *Massive Open Online Courses: Digital Ways of Knowing and Learning*, 1–64. Retrieved from www.elearnspace.org/Articles/MOOC_Final.pdf

McMahon, M. (2013). A Study of the Causes of Attrition Among Adult on a Fully Online Training Course. *Irish Journal of Academic Practice*, *2*(1), 1–26. Retrieved from http://arrow.dit.ie/cgi/viewcontent.cgi?article=1017&context=ijap

McNeely, J. (1938). *College student mortality*. Washington: United States Government Printing Office.

Meister, J. (2002). Pillars of e-learning success. *New York: Corporate University Exchange*.

Moody, J. (2004). Distance Education. Why are attrition rates so high? *Quarterly Review of Distance Education*, *5*(3), 205–210. Retrieved from http://eric.ed.gov/?id=EJ874960

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education*, *14*(2), 129–135. http://doi.org/10.1016/j.iheduc.2010.10.001

Moore, R., & Miller, I. (1996). How the use of multimedia affects students retention and learning. *Journal of College Science Teaching*, *25*(4), 289–293.

Moskal, P., Dziuban, C., & Hartman, J. (2013). Blended learning: A dangerous idea? *Internet and Higher Education*. http://doi.org/10.1016/j.iheduc.2012.12.001

Nash, R. D. (2005). Course completion rates among distance learners: Identifying possible methods to improve retention. *Online Journal of Distance Learning Administration*, *8*(4), 1–26.

Oecd. (2007). Giving Knowledge for Free - The emergence of open educaitonal resources - Summary. *Knowledge Creation Diffusion Utilization*, 153. http://doi.org/10.1787/9789264032125-en

Pane, J. F., Corbett, A. T., & John, B. E. (1996). Assessing dynamics in computer-based instruction. *Proceedings of CHI 96 Conference on Human Factors in Computing Systems*, 197–204. http://doi.org/10.1145/238386.238482

Pappano, L. (2012). The Year of the MOOC. *The New York Times*, 1–7. Retrieved from http://www.edinaschools.org/cms/lib07/MN01909547/Centricity/Domain/272/The Year of the MOOC NY Times.pdf

Park, J. H. (2007). Factors Related to Learner Dropout in Online Learning. *Proceedings of the 2007 Academy of Human Resource Development Annual Conference*, (2005), 1–8. Retrieved from http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED504556

Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Educational Technology and Society*, *12*(4), 207–217. Retrieved from http://www.ifets.info/journals/12_4/18.pdf

Pisutova, K. (2012). Open education. In *2012 IEEE 10th International Conference on Emerging eLearning Technologies and Applications (ICETA)* (pp. 297–300). IEEE. http://doi.org/10.1109/ICETA.2012.6418317

Pouezevara, S. R., & Horn, L. J. (2016). MOOCs and Online Education: Exploring the Potential for International Educational Development. *RTI Press Occasional Paper*.

Richter, S. (2013). Tips for Designing a Massive Open Online Course (MOOC). Retrieved May 29, 2016, from http://facdevblog.niu.edu/tips-for-designing-a-massive-open-online-course-mooc

Rizzardini, R. H., Amado-Salvatierra, H., & Gütl, C. (2013). Cloud-based learning environments: Investigating learning activities experiences from motivation, usability and emotional perspective. In *CSEDU 2013 - Proceedings of the 5th International Conference on Computer Supported Education* (pp. 709–716). Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84887148878&partnerID=tZOtx3y1

Rizzardini, R. H., Gütl, C., Chang, V., & Morales, M. (2014). MOOC in Latin America: Implementation and Lessons Learned. In L. Uden, Y.-H. Tao, H.-C. Yang, & I.-H. Ting (Eds.), *The 2nd International Workshop on Learning Technology for Education in Cloud* (pp. 147–158). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-007-7308-0_16

Rosen, R. J. (2012). Overblown-claims-of-failure watch: How not to gauge the success of online courses. *The Atlantic*.

Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. *Internet and Higher Education*. http://doi.org/10.1016/S1096-7516(02)00158-6

Russell, D. M., Klemmer, S., Fox, A., Latulipe, C., Duneier, M., & Losh, E. (2013). Will massive online open courses (moocs) change education? *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*, 2395. http://doi.org/10.1145/2468356.2468783

Saks, K., & Leijen, Ä. (2014). Distinguishing Self-directed and Self-regulated Learning and Measuring them in the E-learning Context. *Procedia - Social and Behavioral Sciences*, *112*, 190–198. http://doi.org/10.1016/j.sbspro.2014.01.1155

Sanchez-Gordon, S., & Calle-Jimenez, T. (2015). Relevance of MOOCs for training of public sector employees. In *2015 International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1–5). IEEE. http://doi.org/10.1109/ITHET.2015.7218016

Seel, N. M. (Ed.). (2012). *Encyclopedia of the Sciences of Learning*. Boston, MA: Springer US. http://doi.org/10.1007/978-1-4419-1428-6

Serwatka, J. A. (2005). Improving retention in distance learning classes. *International Journal of Instructional Technology and Distance Learning*, *2*(1), 59–64. Retrieved from http://www.itdl.org/journal/jan_05/article06.htm

Spyropoulou, N., Pierrakeas, C., & Kameas, A. (2014). CREATING MOOC GUIDELINES BASED ON BEST PRACTICES. In *EDULEARN14 Proceedings* (pp. 6981–6990). IATED. Retrieved from http://espa-ye5.eap.gr/sites/default/files/EDULEARN14_paper draft.pdf

Terry, N. (2001). Assessing enrollment and attrition rates for the online MBA. *THE Journal (Technological Horizons In Education)*, *28*(7), 64. Retrieved from https://thejournal.com/articles/2001/02/01/assessing-enrollment-and-attrition-rates-for-the-online-mba.aspx

Tina Seelig's Profile | Stanford Profiles. (n.d.). Retrieved May 30, 2016, from https://profiles.stanford.edu/tina-seelig

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*(1), 89–125. http://doi.org/10.3102/00346543045001089

Tinto, V. (1982). Limits of Theory and Practice in Student Attrition. *The Journal of Higher Education*, *53*(6), 687–700. http://doi.org/10.2307/1981525

Tinto, V. (1987). Rethinking the Causes and Cures of Student Attrition. *The University Chicago and London*.

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition. University of Chicago Press* (2nd Ed.). ERIC.

Tseng, S. C., & Tsai, C. C. (2010). Taiwan college students' self-efficacy and motivation of learning in online peer assessment environments. *Internet and Higher Education*, *13*(3), 164–169. http://doi.org/10.1016/j.iheduc.2010.01.001

Ulrich, C., & Nedelcu, A. (2015). MOOCs in Our University: Hopes and Worries. *Procedia - Social and Behavioral Sciences*, *180*, 1541–1547. http://doi.org/10.1016/j.sbspro.2015.02.304

Watson, J. (2008). Blended Learning: The Convergence of Online and Face-to-Face Education. *North American Council for Online Learning*, *572*, 16. http://doi.org/10.1016/j.aca.2006.05.012

Welsh, E. T., Wanberg, C. R., Brown, K. G., & Simmering, M. J. (2003). E-learning: emerging uses, empirical results and future directions. *International Journal of Training and Development*, *7*(4), 245–258. http://doi.org/10.1046/j.1360-3736.2003.00184.x

Willging, P. A., & Johnson, S. D. (2004). Factors That Influence Students Decisions To Dropout Of Online Courses. *Journal of Asynchronous Learning Networks*, *8*(4), 105–118. Retrieved from http://www.adesignmedia.com/OnlineResearch/factors-dropoutv8n4_willging(important).pdf

Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulation of learning within computer-based learning environments: A critical analysis. *Educational Psychology Review*, *20*(4), 429–444. http://doi.org/10.1007/s10648-008-9080-9

Xu, D., & Jaggars, S. S. (2013). Adaptability to Online Learning: Differences across Types of Students and Academic Subject Areas. CCRC Working Paper No. 54. *Community College Research Center, Columbia University*.

Yang, D., Sinha, T., Adamson, D., & Rose, C. P. (2013). Turn on, Tune in, Drop out: Anticipating student dropouts in Massive Open Online Courses. In *Proceedings of the 2013 NIPS Workshop on Data-Driven Education Workshop* (pp. 1–8).

Yousef, A. M. F., Chatti, M. A., Schroeder, U., & Wosnitza, M. (2014). What drives a successful MOOC? An empirical examination of criteria to assure design quality of MOOCs. In *Proceedings - IEEE 14th International Conference on Advanced Learning Technologies, ICALT 2014* (pp. 44–48). http://doi.org/10.1109/ICALT.2014.23

APPENDIX I

In the following, the corresponding primary questions of the conduced survey described in section “Survey with MOOC Maker Partners“ are listed:

1. General Questions

- What is your university ranking according to https://www.timeshighereducation.com/world-university-rankings/2016/world-ranking?
- What are your lab's or institution's estimated experiences in creating MOOCs?
- Is your institution or lab currently actively offering/teaching MOOCs?
- Is your institution or lab currently actively offering/teaching online courses only for residential students?

2. General Questions about the Lab's or Institution's Experience

- For how many years does your lab or institution create and offer MOOCs?
- How many different MOOCs has your lab or institution created?
- How many MOOCs in total has your lab or institution created? Please count the total number of MOOCs offered (taking into consideration MOOCs offered multiple times).
- Which platforms have been used for the MOOCs? (like edX, Coursera, Udacity, ...)
- Which types of MOOCs have been offered?

3. Information about Lab's or Institution's offered MOOCs

- What are the types of learning activities for the offered MOOCs (e.g. video content, quizzes, etc.)
- What are the durations in weeks of your provided/created MOOCs so far?
- Can anybody register to these courses?
- In which languages are the MOOCs offered? (Please use the 2-letter ISO 639-1 language codes from https://en.wikipedia.org/wiki/List_of_ISO_639-1_codes, examples: en, de, es, ...)
- What are the fields/subjects of the offered MOOCs (e.g. Mathematics, Chemistry, Physics, Biology, Life Science, ...)
- How sophisticated are these courses?

- Do you have prerequisites for joining these courses?
- Is there additional course material available?
- Is there a final exam?
- Is there group work involved?
- Are there self-paced elements in the courses, meaning participants can (partially) define their own time-schedule?
- Is there a general discussion forum available, which students could use?
- Can the participants contact an educator/instructor in case of questions or problems?
- Do students get a certificate at the end?
- Can the students earn virtual badges or something similar?
- Do you use any other motivational elements for the participants in these courses? If you do, please list them here:
- Is there a fee to pay in order to participate in the courses?
- How do you plan and create your MOOCs (what team members)? How much effort is involved in this (time resources, financial resources)?
- Are the courses based on the content and structure of a book?
- Is the number of participants limited in the courses?
- Are students involved in grading some work of other students?
- Are the courses mandatory for some participants?

4. Issues and Recommendations

- Which problems/issues did you face?
- Which improvements could you make?
- What should/could be improved in future?
- Are there any issues which should be faced?
- What are your positive experiences?
- Are there any negative experiences?
- Please describe your best practice in creating MOOCs.
  You can e.g. consider
  - the team being involved,
  - the technology of the platforms,
  - the learning activities,

- the MOOC duration and the effort for students per weeks,
- as well as other organizational issues.

- Do you have any special activities to support retentions and overcome high dropout? If so, please describe them here: