UNIVERSITY OF GOTHENBURG

DEPARTMENT OF EDUCATION, COMMUNICATION AND LEARNING

THE ROLE OF VIDEOS FOR STUDENT ACTIVITY IN ONLINE LEARNING ENVIRONMENTS

A comparative analysis of student engagement in Massive Open Online Courses (MOOCs)

Franziska Müller

Thesis:	30 higher education credits
Program and/or course:	International Master's Programme in IT & Learning
Level:	Second Cycle
Semester/year:	Spring term 2016
Supervisor:	Christian Stöhr
Examiner:	Berner Lindström
Report no:	VT16-2920-003-PDA699

Abstract

Thesis:	30 higher education credits
Program and/or course:	International Master’s Programme in IT & Learning
Level:	Second Cycle
Semester/year:	Spring term 20xx
Supervisor:	Christian Stöhr
Examiner:	Berner Lindström
Report No:	VT16-2920-003-PDA699
Keywords:	MOOC, massive open online course, video, student engagement

---

Purpose: This thesis aims at discovering basic learner engagement constructs in three massive open online courses (MOOCs) offered by two Swedish higher education institutions by focussing on learner - content interaction. More specifically it concentrates on learner engagement with lecture videos and different characteristics which influence this engagement. It furthermore analyses MOOC learner demographics and integrates the results into recent work in the field. In sum, the thesis gives first insights and guidance for future research with the support of the course data as well as implication for MOOC design and development.

Theory: Self-Determination Theory of Motivation

Method: Explorative Data Analysis, Descriptive and basic inferential statistics

Results: For video lenght and video completion rate there is a moderate negative relationship in all three courses. For other characteristics results are mixed. The demographic data confirms findings of current MOOC research, namely low course completion rates, declining active student numbers, high number of students from countries such as the United States and India, high educational background and IT affinity, but also predominant use of video lectures compared to other course resources. Some unexpected results with respect to the demographic data were identified for the student population receiving course certificates.

Foreword

Finishing and handing in this thesis meant much more to me than submitting a final project for a degree. I want to express my deepest gratitude to those who supported me on this challenging way. Without you, this would not have been possible.

Table of content

1. Introduction

Numerous technological advancements promised to disrupt education, with almost as many not fulfilling those promises. Heated debates were conducted on directions and influences - does technology steer education or does education steer technological developments? After emerging concepts such as distance education and open educational resources, especially massive open online courses (MOOCs) have been the center of attention in media and research. These massive open online courses have been initiated by US-american elite universities and at best, they finally promised to make high quality education available for everyone everywhere for free. These courses attracted masses of learners, generating huge amounts of data. Whereas for some, these data confirmed prejudices, others hold on to the potential of shedding light into the nature of learning in those environments with potential implications for offline learning as well. Terrabytes of data of learner traces have been analysed, interpreted and published. Four years after the year of the MOOC the disrupters of education are still alive and kicking. First and foremost, these course where one thing: massive. Critics have tackled the openness, the for-free-ness, the benefits for underprivileged learners. The excessive enthusiasm has turned to constructive criticism. Whereas the interest in MOOC research has slightly abated in the USA, European higher education institutions show ongoing interest in the development and research of MOOCs. Learner data has been enriched by numerous concepts and models of learning success with learner-centered approaches being on the rise. Demographic data has been clustered not only to describe courses post-hoc, but - with the support of multi-dimensional models - to predict learner behavior and adaptively support the individual learning progress. Learning is acknowledged as a complex concept where different theoretical explanations gear into each other to account for the learning process, knowledge creation and knowledge transfer. A MOOC does not per se represent a better learning environment. It is however an emerging mean to make knowledge available and education borderless. In a society where education might be one of the most valuable goods existing, understanding open online learning environments and improving them is a fruitful motivation for a research project. Also, the dominant research by US-institutions has been extended by contributions from all over the world; different perspectives which enrich the landscape with student demographics, motivations and intentions for taking MOOCs. Interestingly, some observations are confirmed spanning over course subjects, course set-ups and pedagogical approaches whereas for others there is no common understanding in sight. MOOC provider platforms mature and strengthen their business models. Data analysis is an established part of the research, emphasizing the importance of learner engagement for learning success. Whereas the disruption of higher education might not have occurred as expected, the interdisciplinary research into MOOCs still continues to produce interesting results with potential impact on learning in MOOCs and beyond.

1.1. Massive Open Online Courses (MOOCs)

Massive open online courses have been described as everything on a scale from revolutionizing the concept of learning and being overestimated in their value for a learning society. This discussion can be be clustered from several perspectives and different levels as well as analysed through different lenses. In addition, to the term behind each letter in the acronym authors have provided strength and weaknesses, challenges and opportunities. This chapter presents a critical evaluation of MOOCs and concludes why despite an ongoing debate it is still worth to contribute with research into massive open online courses. The term learner-centeredness used in this thesis indicates two-fold. Firstly it refers to the learner as being empowered to learn self-directedly and secondly, it refers to data analysis which focuses on individual learning needs and patterns to guide research from this perspective as well.

The term itself originated in connection to a course created by Siemens and Downes in 2008 ('Connectivism and Connective Knowledge', Cormier & Siemens, 2010). MOOCs are defined as scalable, accessible online learning environments which come in course-like structure (UKÄ, 2016a). The term *massive* describes the component of scalability, which does not necessarily refer to the actual number of learners. Instead, it represents the possibility of enrolling more learners and running the course several times without increasing course resources respectively. It is the unpredictable scalability component, which distinguishes MOOCs from free online courses with the latter not being a new phenomenon (see Johnstone, 2005 as cited in Seaton, Bergner, Chuang, Mitros, & Pritchard, 2014). *Open* stands for accessibility of the learning content, which is not restricted by any monetary, competence-based or platform-related components. *Online* refers to the ability to find the entire course material online and *course* describes the idea that learning is an entity matching traditional notions of university credit courses. Successfully completing a MOOC does not lead towards higher education credits automatically. It is the scalability and the structure as a course which distinguish MOOCs from for example open education resources (OER) and digital learning materials (DLMs). Referring to Jansen and Schuwer (2015), this definition is also in line with various research projects on MOOCs funded by the European Union. The discussion on the definition of MOOCs is ongoing which each of the four concepts being part of it. One major critique is that MOOCs are no longer free (Dodd, 2014; McGuire, 2014) as acknowledged for example by the definition of Selwyn, Bulfin, and Pangrazio (2015).

MOOCs are a global phenomenon, having aroused interest for development and research of those courses first in the USA and expanding quickly to the rest of the world (Grossman, 2013). De Freitas, Morgan and Gibson (2015) identify business models and open access for education as powerful impact factors. Technology is the enabler for the emerging MOOC trend, with broadband, portable devices and life through social media as examples. According to De Freitas et al. (2015) the "Americanization of learning" (p. 461) has not been addressed in MOOC literature. 2012 was phrased as "the year of the MOOC" by Pappano (2012) and illustrated the outreach beyond higher education. Hyman (2012) went so far as calling 2012 the year of disruptive education and Vardi (2012) asked the question if MOOCs will destroy academia. Several books were published following the hype (cf. Kim, 2014; Porter, 2015; Hollands & Tirthali, 2015; Pethuraja, 2015; Mendoza-Gonzalez, 2016). 2013 and 2014 have been the years of critical evaluation of MOOCs with frequent points of criticism such as high dropout rates or low completion rates respectively, intense resources needed for development or the one-size-fits-all approach (Zemsky, 2014). Demographic data delivered insights into learners' backgrounds which in most courses were represented as well educated, IT-versile individuals (Fischer, 2014; Hill, 2013; Selingo, 2014; Hansen & Reich, 2015). Others have identified teachers as an important MOOC learner population (Seaton, Coleman, Daries, & Chuang, 2015). Moreover, the development of different payment models evolved, thus questioning the notion of MOOCs being offered for free. Kalz and Specht (2014) identified a lack of addressing learning design when it comes to tackle the challenges of numerous diverse participants of MOOCs and low completion rates. Whereas some authors contribute from specific lenses, others try to cover all layers of the multifaceted MOOC environment (Haggard, 2013).

Fischer (2014) offered an overall perspective from the learning science and identifies a need to balance the two extremes of a MOOC hype and a MOOC underestimation. He also pinpoints a lack of research focussing on the learning sciences, with most research targeting economic and technological perspectives and leaving aside qualitative and quantitative data. From his standpoint, MOOCs are only one puzzle piece in a learning landscape and in an early stage of development. Fischer framed main issues which are represented by learner data, where three appear specifically appealing. On one hand, MOOCs seem more appropriate for courses where answers are known. This is a statement tightly connected to the development of educational technology and online education as well as related theoretical underpinnings. One the other hand, experimentation might be oppressed by high production costs. Finally, MOOCs tend to attract a certain student type. Eisenberg and Fischer (2014) addressed learning success in the light of high dropout rates which are characteristic for MOOCs. Dropout rates direct to the delta between initially enrolled and finally graduating learners and approximate around 90% as consolidated by Khalil and Ebner (2013). Jordan (2014) identified an average of 6.5% completion rates, her ongoing research project pointing towards 15% (Jordan, 2015). So called success rates raise two main questions. As Eisenberg and Fischer (2014) accentuated, it is questionable if dropout rates are a meaningful indicator for learning success. Rather a combination of identifying meaningful learning activities for the individual and measuring the perception of the fulfillment of these activities could account as a success factor. Also, dropout rates seem high on the relative level whereas in absolute numbers a 90% dropout rate for a course with 25.000 learners - an average for enrolled MOOC learners identified by Jordan (2015) - would imply that 2.500 learners still finish the course successfully.

In sum, the challenges concerning MOOC development and research as presented in this chapter have to be weighted and considered with respect to the course of this thesis. MOOCs are hence not presented as the panacea for higher education but rather as a promising contribution to a diverse landscape of learning and especially to research work on learning within this landscape. MOOC development must not be equated with development of learning. MOOCs are not analysed despite but because of the recent debate, where their adoption is beneficial when going beyond building up reputation and attracting new students to promoting exchange and collaboration between developers and researchers (Gaebel, 2014).

1.2. Making sense of increasing amounts of data available[1]

Simultaneously with technological advance opening up new learning environments, the availability of data and methods for analysing them resulted in emerging themes such as data driven research and big data in general as well as learning analytics and educational data mining specifically in the field of education. Especially for MOOCs, the potential promise of benefits for research due to the waste amount of data collected motivated the research community to dig deeper into these concepts. Learner-centeredness, a recurring theme directed to data analytics in the learning sciences, is an aspect which guides the development and unites different strands. Originally, this development is based on corporate roots, with consulting companies setting the pace for publishing reports on big data as a business trend and analysing implications for diverse branches (Manyika et al., 2011; Accenture, 2015). Big data loosely refers to "data sets so large and complex that they become awkward to work with using standard statistical software" (Snijders, Matzat, & Reips, 2012, p. 1). Over time, diverse definition have been set-up, the three (Laney, 2001) or four V's ("The Four V's of Big Data", 2013) being used commonly to describe the characteristic components Volume, Variety, Velocity and Veracity (Mauro, Greco, & Grimaldi, 2016).

Manyika et al. (2010) describe educational services as having above average big data value potential, whereas the overall ease of untapping this value is rated lower as for other sectors (p. 10). From a critical perspective, big data can not be put on a level with big information. The data alone does not provide insights, rather it is the methods used and the interpretation of the analyst as well as the context both are placed. This plays a role for meaningful conclusions as well. They consider more than numbers but a broad interplay of additional contextual information. It has also been argued that the bottleneck for meaningful information based on coherent data analysis is not the gathering of data. Rather it is the use of already existing data sets in a more meaningful way. Privacy and ethical issues have been under consideration with decisions to be made about which data to collect, how to inform the affected user and how to deal with sensitive information (Morse, 2015; Halevy, Norvig, & Pereira, 2009). Big data is the starting point of an interesting

[1] Parts of this chapter have been published earlier as a public online slide share https://prezi.com/njjkzziksbwa/treasure-hunt-with-learning-analytics/ and a private blog https://goo.gl/OQXcCnL.

technological development which has agitated the research community in general and in particular the learning sciences. Kitchin (2014) goes as far as to describe the exploratory science enabled by big data as a fourth paradigm of science. Hey, Tansley, and Tolle (2009) label the fourth paradigm the data-intensive scientific discovery. This marks the beginning of a discussion on epistemology, quantitative and qualitative methodological research approaches and the capacity of the respective research fields. Big data in education has moved towards the fields of educational data mining and learning analytics with the clear focus on bolstering the learner in online learning environments (Romero & Ventura, 2010; Siemens, 2013; Baker & Siemens, 2014). This development responded also to emerging doubt if data would be the last word on the subject following the ongoing debate in educational research on the dualism between qualitative and quantitative research methods (Perry & Nichols, 2014; Pring, 2000).

Data from MOOCs enabled not only instructors to learn more about learners' engagement but also researchers to analyse and evaluate huge data sets from thousands of participants. The field of learning analytics emphasizes its interdisciplinarity and the meaningful connections of different technical, pedagogical and social perspectives. Suthers and Rosen (2011) identify the main challenge of bringing together fragmented digital traces of users, with data not capturing everything, deciding on which data matters, how to bring data together in a meaningful way and create multi-dimensional models, and finally the questions of privacy and ethics. Siemens and Long (2011) highlight how quantity affects ways and methods used to approach data as well as make sense of it. Knight, Buckingham Shum, and Littleton (2014) identify learning analytics as implicitly or explicitly promoting particular assessment regimes in the epistemology, assessment and pedagogy triad. Suthers and Verbert (2013) define learning analytics as "the middle space" (p. 1) between learning and analytics. In their paper they elaborate three main themes for future research in the field of learning analytics: "the middle space" (p. 1) which focusses on the intersection between learning and analytics (and avoids to prefer one), "productive multivocality" (p. 2) which emphasizes the challenge of unifying a multifaceted research field by focusing on analyzing a common data ground and "the old and the new" (p. 2) which enhances learning as a century-old idea that is continuously accompanied by new tools. Given the rich online learning landscape, clustering learning environments can be the first step of detecting characteristics, underlying epistemology-assessment-pedagogy beliefs and thus identifying the appropriate measures of learning analytics. Particularly focussing on MOOCs, two main clusters have emerged as illustrated in Figure 1.

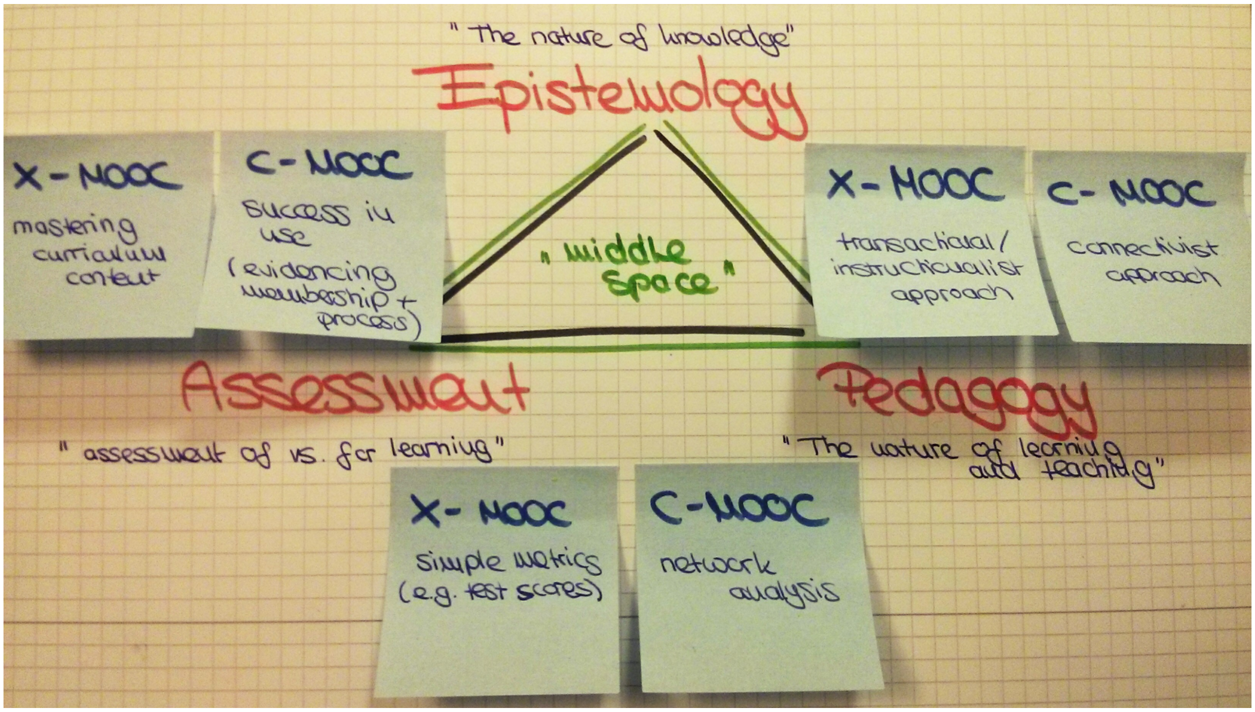

Figure 1. Epistemology–Assessment–Pedagogy triad based on Knight et al. (2014), p. 4.

Rodriguez (2012) classified MOOCs as either x-MOOCs ("AI-Stanford like courses", following the cognitive-behaviorist tradition) or c-MOOCs (following the connectivist tradition). The former is rooted in cognitive-behavioristic traditions with instruction focusing on individual learners whereas the latter is rooted in connectivist traditions with instruction focussing on social interaction between learners. The term "x-MOOCs" was not coined by Rodriguez, but Liyanagunawardena, Adams and Williams (2013) who established ties to Daniel (2012). For his classification Rogriguez used Anderson and Dron's (2011) "Three Generations of Distance Education Pedagogy" where they coin three pedagogy concepts in distance education. The implication of this classification is a varying view on teaching, social and cognitive presence in the online learning environment. This needs to be considered when analyzing the underlying epistemological concept and the assessment formats.

Besides common features, this relates especially to the role of course instructors, the definition of openness (access vs. openness to personalized learning), connectedness and guidance. Knowledge is either generative (c-MOOC) or declarative (x-MOOC). Without a coherent triad the best assessment strategy does not tackle the real learning taking place. Furthermore, the triad can be used to continuously challenge the assumptions of each corner. This is of importance, as for the description and success evaluation of MOOCs usually simplistic demographics are being used. Drop-out rates and final grades are considered to reflect the course quality and the learner success. More advanced attempts built on methods from the learning analytics fields, analysing single learner paths in form of click-stream data or multi-factor models to identify learners' success within the learning community suitable for a learning environment characterised by openness, scalability and self-directedness. They enable to analyse behavior in a learning environment and by this providing groundwork and potential for the improvement of learning environments, and individual learner's feedback and learning success (Shum & Ferguson, 2012).

In general, learning analytics are based on learning theories, more specifically, social learning analytics pinpoint learning elements that are significant in a participatory online culture. They acknowledge that learners are not learning alone but engaging in a social environment, where they can interact directly or their actions can be traced by others. The challenge of implementing these analytics is still present. For a data driven research approach, other fields building their results on analytics have to be considered and their methods need to be understood to interpret results and use them to develop research approaches further. Learning analytics and educational data mining can build on learner success and social networks (Gašević, Zouaq, & Jenzen, 2013; Grunspan, Wiggins, & Goodreau, 2014) based for example on Haythornthwaite (1996) or direct the attention towards learning design (Lockyer, Heathcote, & Dawson, 2013). In the words of Koedinger, D'Mello, McLaughlin, Pardos and Rosé (2015) as well as Singer and Bonvillian (2013), in an interdisciplinary field of educational data mining and learning analytics, research questions on how learning can be relevantly modeled, can be rewarding for students and anticipates with respect to the two revolutions in learning, increasingly affordable and accessible courses as well as attention on the learning science.

2. Aim of this thesis

This thesis aims at exploring learner engagement with video lectures as course components of three MOOCs of two Swedish universities. It is recognized that learning is not represented by single course component interaction. However, as improved learner interaction with respect to course components positively affects the learning process and the learning results, research into student engagement must take on a central position in the context of both - traditional and online education. The central research question guiding this thesis is how video characteristics influence and correlate with completion rates and completed views by learners. Further, quality of course resources are considered as of major importance for student engagement in MOOC environments. In addition, it is contemplated that a learner focus and the concentration of course resources primarily used by the learner can reveal insights into the learning process. Videos have been in the focus of prior research, however only a few focus on MOOCs as a learning environment. Furthermore, authors did not cover a European perspective on these course components nor did they use learner activity data for their analysis. (cf. Milligan, Margaryan, & Littlejohn, 2013; Coetzee, Fox, Hearst, & Hartmann, 2014; Ho et al., 2014). Data richness does not automatically represent meaningful information as the result of applied statistical methods. This thesis picks up assets and drawbacks of the research method used and weights them accordingly. The data analysis executed in this research is guided by strong grounding in the learning science, guiding hypothesis and iteration and adaptation of hypothesis to enable essential results and information on learners behaviour in a specific online learning environment. This work intends to contribute with research on how learner actually leave traces during their MOOC studies, thus how they behave and what results can contribute to the development of such online learning environments and the academic community embracing those learning environments to shed light into learning processes and learner's engagement. Three major aims will be covered by this thesis. Firstly, it will contribute to European MOOC research from a Swedish perspective, which has been identified as inadequately represented. Secondly, current research around video lectures as learning components in MOOCs firmly grounded in the learning sciences by applying a student-centered approach shall be expanded. Finally, results from the thesis will be used to develop design implications.

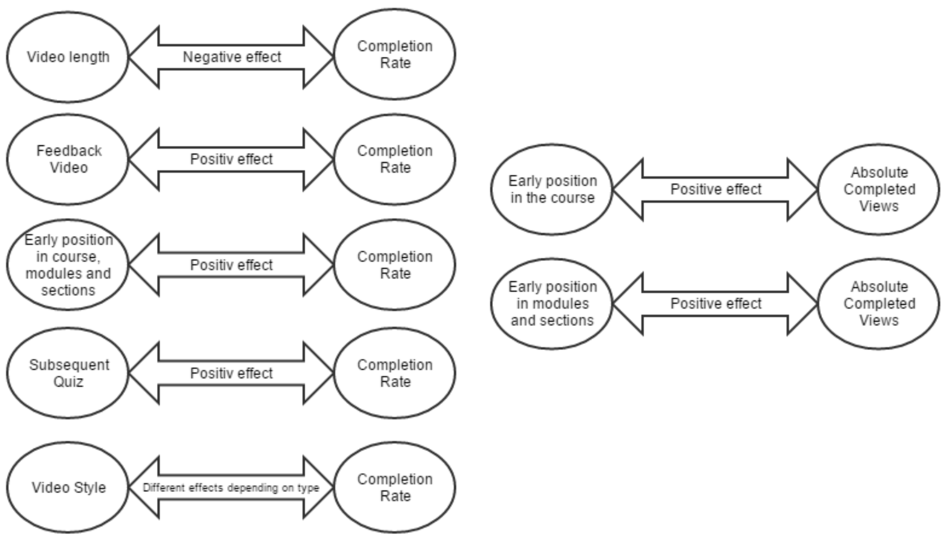

Figure 2. Hypothesis derived from the research question.

Based on the research question seven hypotheses have been formed as can be seen in Figure 2. A negative effect is expected for the relationship between video length and completion rate. Positive effects are expected for the relationships between completion rate and feedback video, early position in course, modules and sections as well as quizzes which follow the video. Deviating effects are expected when it comes to the relationship between completion rate and different production styles. In turn, positive relationships are expected between the variables absolute completed views and early position in the course as well as early position in the modules and sections respectively.

Chapter 1 functions as the introductory chapter to the main themes of the thesis: massive open online courses and learner-centered data-driven research approaches. This chapter gives an overview of the aim of this thesis, the research questions and the derived hypothesis. In chapter 3 relevant background information are provided. First of all, MOOCs are places in the learning landscape in the subchapter 3.1. Afterwards, edX as a MOOC platform provider is introduced in subchapter 3.2. After the importance of videos for the learning process in xMOOCs is explained in subchapter 3.3, the next subchapter 3.4 completes the background chapter with an overview of MOOCs in Swedish higher education. Chapter 4 summarises and classifies earlier research on engagement in MOOCs with a focus on video interaction. The theoretical model illustrated in chapter 5 is followed by the representation of the research method in chapter 6. Besides the methodology for the literature review, this chapter explains in detail the research approach, the data structure and the three MOOCs analysed as the data basis. Results are subsequently represented in chapter 7. After the discussion of the results and limitations in chapter 8, future research and implications for design are discussed in chapter 9. Finally, the summary chapter 9 recapitulates the work.

3. Background

3.1. Placing MOOCs in the learning landscape[2]

MOOCs are both - outcome of technological and pedagogical advances as well as starting point for the analysis of large data sets representing interaction with online learning environments. Where learning and technology meet, an understanding of both concepts and their interrelation is essential for the foundation for research and further discussion. As a recent trend, Liyanagunawardena et al. (2013) observe that learning technologies become tailormade and can adapt flexibly to different users. MOOCS are one example of emerging trends in education. This chapter builds upon the three perspectives on cognition and learning described by Greeno, Collins and Resnick (1996): the behavioristic, the cognitive and the situative. The field of learning sciences has changed over time and has since then be influenced by and was influencing factor for technological developments targeting the educational sector. One discipline grounded in learning sciences is education, aiming at facilitating the learning process. Changing cultural, social and technological circumstances call into question existing beliefs of the educational process (Kalantzis & Cope, 2012). Trends of globalisation and digitalisation challenge traditional beliefs of education, with educational settings opening up and lines between formal and informal learning fading. Interdisciplinary approaches enable the learning sciences to execute research with the support of other fields but also increases the complexity and efforts of collaboration. Besides many overlapping areas and terms being coined to describe these development, common ground is targeting and understanding the learner, with his/her needs, backgrounds and social learning environments.

Perry and Nichols (2014) emphasize how different theoretical perspectives can be used to explain a phenomenon in educational research. In the field of education and in the social

[2] Parts of this chapter have been published earlier on a private blog https://goo.gl/OQXcCn.

sciences, there is no overarching theory explaining everything but rather several theoretical models which can be applied to interpret observations and to identify impacts on student's learning and behaviour (Glaser et al., 2001). The context of learning is rich, technology is one important part of it.

Theoretical underpinnings within the science of learning have developed from more or less radical behaviorists which believe in observable behavior as the true scientific approach to research on learning, over the study of mental functioning triggered by the new field of cognitive science in the late 1950s to the importance of the social and cultural contexts of learning more recently. These perspectives have also shaped pedagogy, instruction and the design of offline and online learning environments.

From the behavioristic perspective which emerged during the 1930s, knowing is an observable connection between stimulus and response whereas learning is forming these connections through the process of (non-)reinforcement. The learning process starts with simple components of a skill which are combined or differentiated to acquire more complicated ones. This implies mainly extrinsic motivation which is needed for the learning process whereas effects depend on internal factors. Transfer of knowledge occurs when learned behaviours can be applied in different situations and depends on the amount of connections and similarity of stimuli (as opposed to the cognitive perspective) (Greeno et al., 1996). Glaser et al. (2001) point out that several components of recent cognitive theories describing knowledge and skill acquisition are further developed variants of the stimulus response associative theory. However, behavioristic positions can not account for underlying structures of mental events nor the copiousness of thought and language procession. Behaviourism resulted in Bloom's Taxonomy of Learning as well as sequential learning planning (Carlile & Jordan, 2005).

The cognitive view is based on theories of cognitivism, with accommodation and assimilation as key concepts to explain knowledge which is created based on own experience. Learning itself is an active process rather than the construction of knowledge being a product (Perry & Nichols, 2014). Internal representation are created while people learn and they are based on how knowledge is encoded, organised and retrieved (NRC, 1999). New information is integrated into existing frameworks of structured knowledge, enabling the learner to go beyond collecting facts and procedures to more complex tasks such as to interpret situations and solve problems. Cognitive theories adapt behaviouristic approaches by taking into account the nature of knowledge someone acquired (and not only how much knowledge someone acquired). Instruction took over the promotion of for example active listening and learning chunks (Carlile & Jordan, 2005).

The situative (or sociocultural) view emphasize the context of learning as an important component contributing to knowledge creation, where context refers to engagement in practice or community. This view developed further the cognitive perspective which nearly exclusively focusses on individual learning. The situative view on learning acknowledges learning as a collective activity. Knowledge creation is seen as mediated by cultural artifacts with the aim to participate in a particular community (Greeno et al., 1996).

The development from behaviorism over cognitivism to situative/pragmatistic sociohistoric views can be recognized in developments of educational technology as well. Learning theories can be related to paradigm shifts in instructional technology. Whereas Computer Assisted Instruction (CAI) can be classified as behavioristic (how well can software support the learner to achieve specific knowledge), Intelligent Tutoring Systems (ITS) belong to the cognitivistic perspective (how well does software mimic a real teacher). LogoasLatin can be arranged within the constructivistic tradition (how well can software support students in transferring knowledge) whereas Computer Supported Collaborative Learning (CSCL) is connected to situated learning (how well does the software support learners in engaging in knowledge communities) (Greeno et al., 1996; Koschmann, 1996). The major outcome of this development is that as views on knowledge, learning and transfer develop, the role of technology is shifting, too. Current developments in IT & Learning emphasize learner-centered learning environments and scaffolding. Learner centeredness describes the focus on the learner's psychological learning process or her/his participation in a sociocultural learning process (Hoadley & Van Haneghan, 2012). The generations of distance education exemplify the changing roles of cognitive, social and teaching presence (Anderson & Dron, 2011). Distance education has emerged from correspondence education, with learners and instructors physically separated but in constant exchange (Keegan, 1996; Holmberg, 2001; Peters, 2010). Another notion of this development is openness of education, with learning resources being available for everyone (also for reuse and modification) anytime anywhere. MOOCs are a further development of distance education with their routes in the movement of open education (Jansen & Schuwer, 2015).

MOOCs emerged in an US-american higher education context with universities asking for high tuition fees. Kovanović, Joksimović, Gaševic, Siemens, and Hatala (2015) identify a decreasing coverage of MOOCs in media in 2014 with increasing focus on the position of MOOCs in higher education, analytics and big data as well as adaption of this online learning environment in different parts of the world. Selwyn et al. (2015) describe the media discourse in the terms of general change as well as massification, marketization and monetization of higher education. Anders (2015) describes this development with the support of Gartner's hype cycle, where MOOCs reached their peak in 2012, the point of disillusionment in 2013/2014 and can now be placed in the "slope of enlightenment" which might partly lead to meeting high expectations through a combination of practical applications and long term impact. Research on MOOCs has mostly focused on case studies, how this learning format affects higher education, or how theories of education frame this research object (Liyanagunawardena et al., 2013). Most importantly, the notion of *the* MOOC has shifted to a more multifaceted understanding of different types of MOOCs, their notion of learning, their pedagogic approach and how they are used by learners and instructors.

Indeed, early critics of the MOOC movement are based on a specific MOOC type or MOOC platforms defining the MOOC outline (compare for example Rees, 2013; Kolowich, 2013b; Woolf, 2014; Rehfeldt, Jung, Aguirre, Nichols & Root, 2016). Even the common distinction between x-MOOCs and c-MOOCs has been challenged by several perspectives including hybrids and blended learning (Anders, 2015; De Freitas et al., 2015). x-MOOCs mainly contain videos plus quizzes or assignments, edX being a typical MOOC provider building upon this perspective (Conole, 2013). Both formats reach scalability by limiting synchronous learning activities and individual academic feedback. Peer learning is an important component within both but the role of instructor differs (hierarchical vs. distributed view on learning) (Universities UK, 2013). Arguing that such a two-extreme classification is too simplistic, Conole (2013) suggests a scheme based on 12 dimensions to evaluate diverse important characteristics of MOOCs from degree to openness, over amount of multimedia and communication towards type of learner pathway etc. Anders (2015) establishes a frame of a continuum which includes multiple theories and applications of MOOCS.

Learning Theories		
Cognitive-behaviorist	Social-constructivist	Connectivist
Individuals	Groups/Communities	Crowds/Networks
Prescriptive	Prescriptive/Emergent	Emergent
MOOC Applications		
x-MOOCs	Hybrids	c-MOOCs
Content-based	Community and Task-based	Network-based

Table 1. Learning theories and their applications in MOOCs based on Anders (2015).

The above table highlights the continuum established by Anders (2015) with endless possibilities for additional hybridization, as learners are potentially in the role of adapting the learning resources to their needs (Clark, 2013; Roberts, Waite, Lovegrove, & Mackness, 2013; Beaven, Hauck, Comas-Quinn, Lewis, & de los Arcos, 2014). In addition to this, Rubens, Kalz and Koper (2014) describe their online master-class as being located in the middle between x-MOOCs and c-MOOCs, a hybrid in Anders' terms. Another example for the classification of MOOCs is the proposed one-third model by De Freitas et al. (2015), where one third of the learning experience time is invested in the format of video and audio materials, of interactive material (e.g. quizzes, assignments) and social interaction respectively. These classification synthesis depict the extensified research in the MOOC area as well as an increasingly tight connection between learning sciences and educational technology.

The most common measures for describing MOOCs and their success can be clustered in general platform-, course- and learner-related indicators. Shad (2013/2014) lists for example: the number of MOOCs available, MOOC provider, subjects, languages, top searches and top courses. Learner demographics are mostly clustered in socioeconomic status, educational background, IT afinity, prior experience with MOOCs and online learning environments and prior subject knowledge. When it comes to interaction with MOOCs, it's the development of students enrolled, completion rate and activity level of students which are described. A comparison of these key indicators between different courses on meta level led to the main critique of massive open online courses. Hansen and Reich (2015) question that MOOCs fulfill their long expected tasks of providing education to everyone based on the fact that they attract mostly learners with high socioeconomic status and strong educational backgrounds. Eisenberg and Fischer (2014) outline that the limitation for learning is not the access to learning material but the motivation to learn. In addition, Fischer (2014) emphasizes that not everything is 'moocable' (p. 154) and stresses the fundamental challenge of establishing a symbiosis between on-campus courses and MOOCs. On the other hand, Kortemeyer (2013) identified three problems with open educational resources, which could be potentially solved by MOOCs: discoverability and quality control of learning resources as well as putting the learning resource into the right context.

In conclusion, MOOC media coverage and research has evolved from the pure description of the new open online learning environment towards focussing on the notion of learning and how to improve learning for the individual learner's perspective. The historical development has shown that MOOCs stem from the movement of open online education, with the aim of making knowledge accessible anytime anywhere. Whereas it has been questioned if MOOCs can fully accomplish this goal, research in this area can contribute to support this process in future. The MOOC discourse evolved from supported business models to platforms for supporting education. After the long awaited educational reform was not provided, research, policy makers and big data entered the picture to concentrate on contributing to an improved learning process for the individual. Therefore it is necessary not only to understand who is learning with MOOCs but also how learners learn and how they engage with the different learning components of such courses. Also, it is of interest, how MOOCs are adopted in different parts of the world apart from the US.

3.2. edX as a MOOC provider

Several MOOC platforms emerged and applied different revenue models which benefit from the willingness of higher education institutions to participate in the open education movement (Yuan & Powel, 2013). edX is one of the big three MOOC provider (Round, 2013; McGuire, 2014) and was founded as a non-profit organization by the Massachusetts Institute of Technology (MIT) and Harvard University in 2012 (Lin, 2012). Today, the edX consortium consists of 12 universities and course cover diverse subjects, fields and areas. In comparison to other platforms, edX develops courses slower with smaller numbers of universities (Universities UK, 2013).

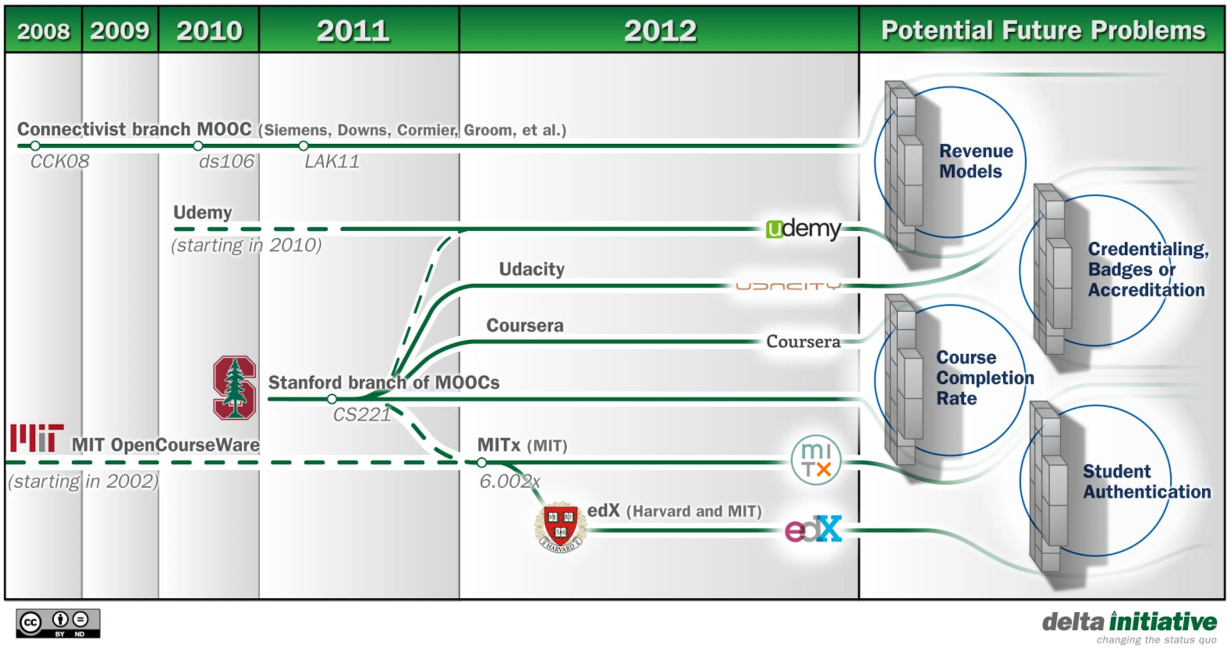

Figure 3. Emerging MOOC providers and potential future problems (Hill 2013).

Figure 3 describes the development of important MOOC providers and potential future problems. Besides being non-profit, edX is different from other providers in terms of being open source (Carr, 2013). Hashmi and Shih (2013) indicated that both founder universities invested around $30 million of financial resources to make this collaboration project possible. Kolowich (2013a) describes the business model which seems interesting given the fact that other platforms are for-profit organizations (e.g. Coursera and Udacity). Still, revenue models seem to be a crucial potential problem, thus resembling challenges of open educational resources in general after investment into those decreased (Kortemeyer, 2013). This reflects another perspective on high dropout rates of MOOCs: learners do not stay long enough to be willing to pay for certification. Whereas from a learning perspective completion rates might not mirror learning success, from the business model perspective it becomes obvious why dropout rates are in focus of discourse and critique. Proposals for revenue streams include paid certificates, the targeting of corporate training (Korn, 2014), blended approaches (Harris, 2013), and international collaboration/expansion (Meyer, 2013; cf. Mehaffy, 2012; Alstete, 2014). Meta-platforms or MOOC catalogues combine contents from different MOOC providers and create transparency for learners. The four potential future problems as presented in Figure 3 are on the one hand related to business models (with edX covering its costs by taking in a specific percentage of inflowing earnings from partnering universities) and on the other hand closely related to the learner perspective, although not necessarily related to learning quality itself. Accreditation, course completion rates and student authentication are of interest for learners but do not directly connect to how the platforms are used for learning or what the addressed problems say about the individual learning process. As indicated by for example Universities UK (2013), interests connected to MOOCs has moved from the underlying business model towards research on learning. Shah (2014) identifies increasing production quality of course content as well as the trend for institutions to choose Open edX for hosting MOOCs. Whereas in the beginning of the MOOC movement those resembled on-campus courses including for example clear start and end dates as well as deadlines, recently the anytime anywhere mindset has been put in the focus. Self-paced courses allow starting at any point of time and eventually assessments can be taken taking into consideration individual preferences regarding timing as well. Some MOOCs even run synchronous for the first time and then open up for a self-paced version. Learners will benefit from an even increasing course number and overlapping course content with higher content quality, whereas on course level the institutions will enter an intensifying struggle for learners' attention and retention.

The earlier distinguishment between x-MOOCs and c-MOOCs can be seen from the development in Figure 3 as well. Whereas the connectivist branch is concentrating on learner's networks emerging from massive online courses (Cormier & Siemens, 2010), the original setup of edX intents to support mainly the x-MOOC format and a behaviouristic perspective on learning. Course content is structured in modules and sections and individual assignments in the format of quizzes and assessments can be set-up. Collaboration via forums and wikis is possible and the same applies for peer-review and -assessment. Main components of the homepage are a course catalogue and general information about the platform. Course specific components are different content pages (based on edX or also external links), videos, discussion forums and wikis. Those components can be adapted based on the course and the preferences of instructors and MOOC developers. edX offers two different subscription offers, with and without support from the edX team. The platform aspect and how it is set-up needs to be kept in mind as it limits design implications (Alario-Hoyos, Pérez-Sanagustín, Cormier, & Kloos, 2014). Depending on the role of the user, different interfaces are available. Whereas the main homepage is intended to serve the learner and his/her interaction with the course content, edX studio is targeting the instructor role by providing a platform for actual course set-up and design. edX insights in contrast is an advanced source for visualizing learner data with respect to demographic data and learner engagement. This comes in form of an instructor's dashboard and which contains detailed information on learners, engagement and assignment results broken down to single video completion rates and assignment grades. Some of this data can be downloaded from the instructor perspective on the main course page as well. edX insights however, does not enable to track back certain activities on individual anonymised level nor does it allow the limitation to a certain period of time. This means that the dashboard is updated on a daily basis and data shown includes all learners up to the recent date. Whereas the level of granularity is useful during a running course, additional reports have to be downloaded to get condensed information. This is why edX offers data sets for downloading and analysing them further. In conclusion, when it comes to data analysis, there are mainly two options available. Using those interfaces intended to support synchronous analytics during the run of a course in form of a data dashboard (thus visualization of this data is already provided) and those intended to provide full master data for individual data analysis.

3.3. The importance of videos for the learning process

In general, massive open online courses consists of various learning objects, ranging from tools to display content, to trigger an active learning process or to promote collaboration between learners. Grading support can also be an important component of a MOOC (Kulkarni et al., 2015). Grünewald, Meinel, Totschnig, and Willems (2013) emphasize the importance of designing MOOC learning environments in a way that they support multiple ways of learning preferences. Besides different forms of participation, they also identify (lecture) videos as most helpful in their survey. Those lectures can come in different ways, from resembling traditional classroom lectures to short teasers for more elaborated content in other formats. The decisions how to design a course are said to lay primarily with the course instructor or designer, however the MOOC platform has tremendous design implications. For the edX platform its videos which have been described as “the meat of edX courses” (Roos, 2014) emphasizing centrality of these learning objects. In line with Roos, Bali (2014) describes videos as crucial components of xMOOCs, whereas critics point out that in this way a learning 1.0 object is still delivered in a web 2.0 (Butin, 2012). The arising question then is how interaction from the central components of such learning environments reflects and represents learning.

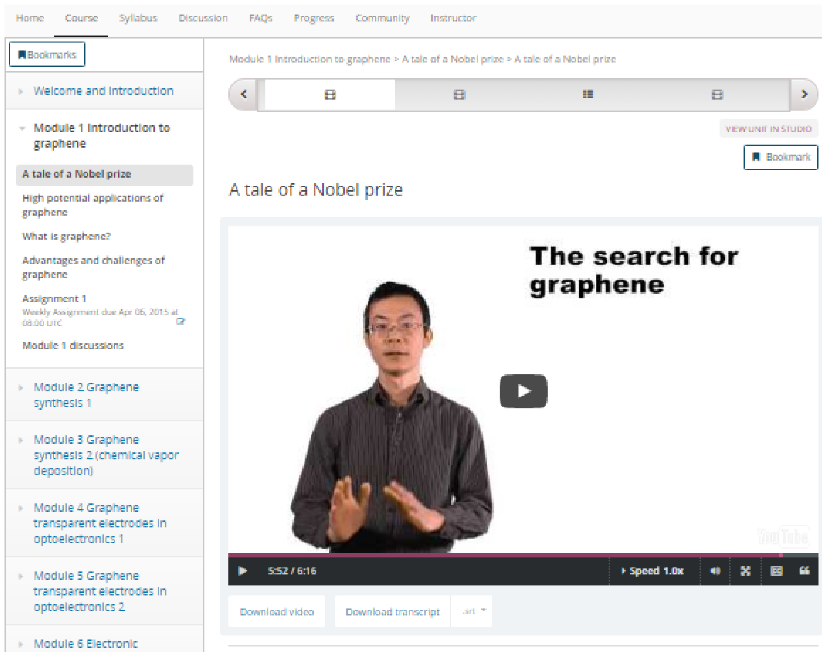

Figure 4. The typical hierarchical structure of an edX course.

Guo and Reinecke (2014) point out three hierarchical organized main areas of an edX course, which are visualised in Figure 4. The weeks of the course are shown on the left side of the web interface, whereas each week can have several subpages. In the above example, Module 1 corresponds to a week and has several subpages connected to it (A tale of a nobel prize being one of these subpages). The main frame in the middle of the page is a so called learning sequence consisting of videos, quizzes or assessments. Releasing and presenting content in chronological order at the platform does not hinder the learner from independently navigating between the different modules, subpages and sequences.

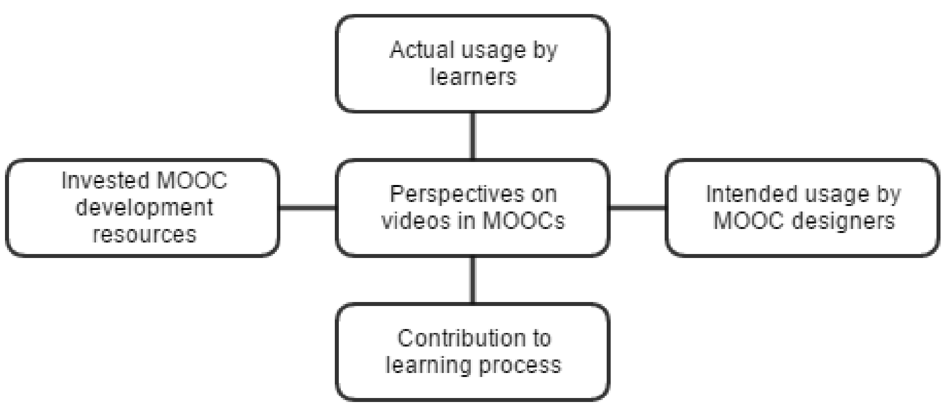

Figure 5: Perspectives on videos in MOOCs

Figure 5 depicts four different perspectives on videos in MOOCs, their separation not necessarily indicating that they can only be regarded separately. The limitations and reinforcements of a provider platform and its implications for design is only one perspective on videos as learning objectives. In addition, there is the perspective on intended learner interaction (as planned by the instructor), the actual learning interaction (as can be observed by generated learner data) and the resources needed to produce these learning objects in a certain quality and style (as e.g. time and money invested for production).

In the light of this thesis, those perspectives are of particular interest, where the individual learner process is affected. As Bali (2014) puts it "[...] the benefit of MOOCs lies not in the way they are designed, nor in what the instructor "assigns" participants, but rather in the spaces for engagement made possible by the course" (p. 52). To some extent this is supported by findings from Guo and Reinecke (2014) who suggest that besides the linearity of MOOC design and curricula, learners move primarily non-linear in these environments. This sheds light into the clash between intended design by instructors and actual usage by learners. Independently of the MOOC platform, Hew and Cheung (2014) identified instructors designing courses comparable to traditional higher education courses with videos being an established part of the content and often used as an opener for a module or week. Even though the interaction with the video is limited and mainly of receptive nature, learners reported on the advantages on pausing a video lecture and taking notes (Frank, 2012). Those videos enable knowledge transfer in an adaptive way (compared to traditional lectures for example) as they can support individualisation by replaying, stopping, pausing, watching later, referring to transcripts or downloading. According to Clark and Mayer (2011) it is this pacing and navigation control which lead to better learning results. McAuley, Steward, Siemens, and Cormier (2010) argue that learner engagement is an important requirement for learning in MOOCs. Further, videos can foster engagement as well as the interaction with videos can represent learners' engagement and thus the learning itself to a certain extend. It is crucial to consider the difference between intended use, actual use and implications based on the possible difference between those. Learners tend to consume content rather passively than actively engage in content production (Kop, 2011). Rai and Chunrao (2016) go as far as seeing a specific learner type - enthusiastic and self motivated - as the requirement for the success in MOOCs. In contrast to other statements, Glance, Forsey and Riley (2013) indicate that MOOCs might have a sound pedagogical approach. However, it's the context which might weaken this approach. In the learning context of MOOCs it's furthermore a fundamental challenge to contextualise interaction data and combine it with learner profiles and demographic data. The focus shall thus be how videos in MOOCs are used by learners and what interaction data can reveal about the learning process. At the same time, insights from learner backgrounds can provide meaningful additional information to shed light into learning with MOOCs and implication for design.

3.4. MOOCs in Swedish higher education[3]

It is noteworthy to mention that while there appears to be a recent downward trend in attention towards MOOCs in the United States, MOOCs have potential for growth in European Higher Education (Sursock, 2015). This is of importance, as the context of higher education has framed the idea of MOOCs. In countries with entry barriers when it comes to higher education from prestigious universities (especially due to high tuition fees), the development of scalable online solutions seems to tackle a situation which is not comparable to most European countries. Seeing the higher education system as the context, it is interesting to analyse MOOCs within Swedish higher education, where there are no tuition fees for locals and freestanding courses are an inherent part of the educational landscape. In addition, recent efforts on adaptation of the legislation to include MOOCs as specific components of the educational system exemplify the importance of the policy perspective in Europe as identified by Kalz et al. (2015). Fischer (2014) and Kalz et al. (2015) point out that challenges and opportunities of MOOCs have been determined but not yet avoided and exploited successfully. Here the in-depth analysis of the context of this study contributes to not only enriching research by a Swedish perspective but also bridging the gap between presenting research results and presenting contextual information. Collaboration and feedback to MOOC providers has been considered by Kalz et al. (2015), and this thesis strengthens this perspective by encouraging not only inter-collaboration with different institutions but also intra-collaboration within one university and different areas of expertise to fetch and analyse the online data traces.

To contextualise the research results, this chapter describes in brief the role of MOOCs in the Swedish higher education system. Higher education institutions (HEI) can have different intentions when discussion, developing and implementing MOOCs. Outreach and improvement of access to education (to less educated learners), branding and cost reduction are only a few motivations to name (Hollands & Tirthali, 2014b). In his review of 100 studies on MOOCs at universities, Karsenti (2013) identifies 12 impacts covering topics like student recruiting, test online teaching methods, and prepare learners for an information society. Some authors specifically address these motivations and those being involved in addressing MOOCs as an emerging trend in the educational landscape (Voss, 2013; Kelly, 2014). Yuan, Powell and Olivier (2014) discusses possible strategic choices based on the purpose behind the development of a MOOC or other online programmes. Jansen and Schuwer's (2015) four main clusters based on Allen and Seaman (2014/2015), Hollands and Tirthali (2014b) and Yuan et al. (2014) include MOOCs based on financial objectives, MOOCs based on reputational objectives, MOOCs based on innovation objectives and MOOCs based on societal and learner demands. The same authors also state, that european HEI are mainly motivated based on the last cluster and that the macro

[3] Parts of this chapter have been accepted for submission for the 2nd MOOCs in Scandinavia Conference 2016 (Campagnaro & Müller, 2016).

drivers in Europe are not new methods in big business neither reducing costs in higher education in comparison to US-american HEI.

Transformation is predicted from sector outsiders, such as management consulting companies (Bokor, 2012). According to Ng'ambi and Bozalek (2015) it is questionable if MOOCs will revisit the role for institutions in higher education. They emphasize understanding of how learning in MOOCs happens and what kind of learning they promote. Simultaneously, the authors criticise the lack of well-grounded effectiveness studies as well as pedagogical frameworks to direct institutions. Steffens (2015) discusses MOOCs as the most recent of developments within lifelong learning. Other authors draw parallels to the open educational resources (OER) movement and the overall generated impact on higher education beyond open universities (Lane, 2013). Dillahunt, Wang, and Teasley (2014) found those in need of education benefitting from the MOOC opportunities. Burd, Smith, and Reisman (2015) discover different approaches for institutions when it comes to business models in relation to MOOCs within higher education. Oriented towards the future, Lombardi (2013) asks five question which aim at having an overall implementation strategy, consultancy with stakeholders, available resources, a revenue model and a research approach in place. Hollands and Tirthali (2014a) take a deeper look at costs and costs comparison of MOOCs, Online and Hybrid Learning. Whereas Jansen and Schuwer (2015) admit that literature on MOOCs is still developing, Gaebel (2014) asks the interesting question “if institutions that deliver MOOCs do not award credits, does that not indicate that they do not consider MOOC as teaching?” (p. 27). Quality concerns with respect to MOOCs have been partly addressed by specifically developed MOOC quality labels (Rosewell & Jansen, 2014).

Although the rapid expansion of MOOCs can be contested as a worldwide phenomenon, there are specific aspects which apply for the European context (Gaebel, 2014). The development of MOOCs and research in the field has been dominated by US-american higher education institutions. In 2013, European institutions started to invest into the development of MOOCs (Jansen & Schuwer, 2015). In international comparison, the Swedish higher education system can be defined as open and flexible (UKÄ, 2016a). Sweden performs well with respect to tertiary education attainment rates and adults participating in lifelong learning. In addition, education expenditures per student are one of the highest in comparison to other EU countries (European Commission, 2015). Universities offer freestanding courses, which can, but do not necessarily have to, lead to a degree. It is the freestanding distance courses which resemble MOOCs most closely. Differences appear with respect to entry requirements and legal status of the learner. Besides the entry requirements, it is the applicant-seat ratio of more than three applicants per seat which limits the accessibility of these freestanding courses. When being accepted to a freestanding distance course the applicant becomes an official university student. Distance education has been an established component of the ordinary Swedish higher education system since 2002. In 2014 21% of the registered students (73.400 out of 344.000) were students of distance courses and programmes (including hybrid courses) (UKÄ, 2016a, p. 22). After a decline in the number of freestanding distance course students after 2010, this number increased again slightly in 2014. Freestanding distance courses are mostly offered in the fields of humanities, theology, juridic and social sciences. In comparison, Swedish MOOCs noted 64.000 students in 2014 and about 150.000 students in 2015. Up till now, 14 MOOCs have been given in Sweden (European Commission, 2016), four of them in Applied Sciences, three in Business, two each in Sciences and Sciences and Technology, and one each in Humanities, Mathematics and Statistics, as well as Social Sciences. MOOCs are described "as a new form of studying in Sweden" (UKÄ, 2015, p. 28). The importance of MOOCs for Swedish higher education received a significant push when the Swedish Higher Education Authority (UKÄ) released its report commissioned by the Swedish government in February 2016. Where the legal and financial situation of developing and implementing MOOCs was a factor of particular insecurity, UKÄ proposed a regulation to improve this situation, handing it over to the Swedish government (UKÄ, 2016a). Currently, this regulation proposal is under review with all affected higher education institutions. Consequently, the dialogue surrounding MOOCs in Sweden has bolstered the opportunities to strengthen collaboration and research in the area of flexible, open, online higher education in 2016. In Sweden, authors have already published on MOOCs and suggested implementing them as a kind of learning center thus following a blended approach (Norberg, Händel, & Ödling, 2015). However, up to now, not much is known and published about those MOOCs offered by Swedish HEI and the learners attending them. The discussion around the regulation proposal by UKÄ has shown that there is interest for such reports and especially for an inter-institutional exchange on the topic.

4. Earlier research on video engagement in MOOCs

The methodology applied for the literature review is described in more detail in the methodology chapter 6. This thesis was inspired by the article written by Guo, Kim & Rubin (2014) analysing student engagement in comparison to video style and length. One of their seven video production recommendation based on the analysis of four edX courses from different universities and fields is that shorter videos are much more engaging than others. They did analyse courses from American universities (MIT, Harvard and UC Berkley) and the fields computer science, statistics, artificial intelligence and chemistry. Broadening this analysis by expanding fields and universities is one recommendation by the authors for higher external validity. Other future directions could be interactivity and selectivity regarding videos as well as answering the question why students pause videos in connection with a lab experiment. All in all, 33 relevant articles have been identified. Those can be clustered in articles analysing engagement with a quantitative approach (based on learner data from the respective platform) or a qualitative approach (based on survey data from either learners or instructors). The data collected varied with respect to volume and depth of analysis of video engagement and description of video patterns. Also theoretical underpinnings varied. Courses mainly focus on computer science and STEM topics, with other topics being covered in forms of articles (cf. Burch & Harris, 2014).

Hughes and Dobbins (2015) point out that several studies have investigated learner engagement and performance in MOOCs. Videos mostly play a role in a more complex constructs of engagement, often paired with learner performance (cf. Belakrishnan & Coetzee, 2013). Four large-scale data analyses with results on video engagement in MOOCs were Guo et al. (2014) as already described, and in addition Breslow et al. (2013), Seaton et al. (2014) as well as Kizilcec, Piech and Schneider (2013). Breslow et al. (2013) as well as Seaton et al. (2014) conclude based on their interaction data analysis that learners in MOOCs spend most of their time watching videos. Breslow et al. (2013) also point out that time is "the principal cost function for students" (p. 16) which motivates the analysis of how they allocate this resource for learning purposes. Depending on which problems are supposed to be solved by learners (homework, midterm and final exams) the time spend can vary during the length of the course. In line with this, Seaton et al. (2014) emphasize that they do not cluster students based on demographics but based on how they interact with assessment items and how much time the spend in the course. The authors surmise that learners (here: those seven percent earning a certificate) spent most of the weekly invested time in interaction with lecture videos. They acknowledge the limitation of the population of certificate owners and suggest to analyse other populations such as those mastering the certificate without interacting with the content. Kizilcec et al. (2013) found four subpopulations of learners (completers, auditors, disengagers and sampler) with auditors being a large subpopulation of the students interacting mainly with videos whereas assessment, discussion and other interaction components are being skipped. For this work, three computer-science courses were analysed, leaving aside characteristics of individual courses, lecture videos and learners. Evans, Baker and Dee (2016) analysed data from 44 MOOCs running on the Coursera platform. They developed a regression model to predict learner engagement and persistence. When it comes to engagement with videos, the title and the position of the video seem to be a good predictor for engagement, whereas video length seems unrelated.

Other smaller analysis have been identified as well. Pursel, Zhang, Jablokow, Choi and Velegol (2016) found watching videos being positively associated with rates of completion and conclude that video views are a strong predictor for course completion. Their results indicate that pure intention to watch all videos (as indicated in the pre-course survey) is related to completing the course as well. Impey, Wenger and Austin (2015) describe engagement with video lectures broadly, by saying that the engagement and completion of those is initially higher at the beginning of a course and then rapidly declining. They analyze data from one Astronomy course with 24.000 enrolled students on the Udemy platform.

Connected to the style and segmentation of video lectures, two studies have been identified. Kizilcec, Bailenson and Gomez (2015) found that while most of the learners enjoyed videos in talking head format, they decided to watch those not being in this format or switched between the formats. Argumenting based on cognitive load, the authors suggest that the instructor is mostly an additional component to process when taking in new information (although the appearance of the instructor might support focussing on the video and the feeling of connectedness). Finally, presentation style did not have an influence on learning outcomes, and a strategic presentation style combining talking head and absence of talking head was inferior in particular for those learners preferring visual clues for learning. Doolittle, Bryant and Chittum (2015) focus on segmentation in video lectures, which was perceived as supporting the learning process. However when reaching a higher level, this segmentation was perceived as annoying. Even though the environment of the learners was not a large scale open course, these findings indicate the important perception of a *too much* in chunking videos and raise the question where exactly this *too much* begins.

Some research articles focus on videos, however, were based on surveys among the learners and not on interaction data from the course platform. Collins, Weber and Zambrano (2014) discuss videos as important course components in their paper but did not investigate further learners' engagement with those. Brinton et al. (2015) propose an adaptive educational system and their results indicate that learners prefer multimodal resources, with videos being favored. O'Sullivan (2016) constitutes that video components are within the more requested resources by the learner. Stephens and Jones (2014) findings suggest, that the limiting factor for learners is time (in line with Breslow et al., 2013) and that they prefer shorter videos.

Belanger and Thornton (2013) report that in their course videos the number of views declined during over time and also that the number of learners watching at least one video was almost twice as high as the number of learners taking any quiz in the course. However, this report has not been published in a peer-reviewed journal. These findings are in line with Ho et al. (2014) identifying playing a video as the main activity for the top five courses in their analysis. Yet another perspective is offered by Adams, Yin, Madriz and Mullen (2014) who identified x-MOOC learners and based on qualitative research found out that the engagement with videos can result in a "tutorial relationship" (p. 213) and thus being opposed to those results suggesting that videos generate isolation and loneliness among learners.

Other case studies and personal reports like Rice (2013) focus on describing difficulties with taking MOOCs while only touching upon the video components. Rice for example highlights the missing interaction offered by lecturers with the videos in MOOCs. In line with this, Evans (2012), Allon (2012) and Frank (2012) give recommendations for the design of video lectures in MOOCs. Evans and Allon claim short video segments to be useful to engage students. Frank (2012) reports that pausing online videos supports in the process of taking notes. Kirschner (2012) criticizes that instructor presence was only given through video lectures. Martin (2012) identified a learning experience similar to well-taught traditional courses due to short videos which explained ideas in a good way and because they were integrated in a meaningful way together with quizzes and further explanations. Also, feedback videos were described as favourable, were instructors would answer those questions rated highest by the active learners. Young (2013) identified the need for additional material besides video lectures due to the difficulty of finding back specific information within a video. Covach (2013) observed that most learners engaged with the videos and that video content had to be more accurate compared to classroom lessons.

Seaton et al. (2014) constitute that country of origin can account for video activity of certificate earners and describe how downloading lectures can account for findings related to video engagement in 11 MITx courses. Woodgate, Macleod, Scott and Haywood (2015) analysed 13 MOOCs from the University of Edinburgh and identified among all the tool used videos being those mostly used by learners. This applied for both, learners receiving a statement of accomplishment and for those not receiving one. However, the number for those receiving a statement of accomplishment was higher. A few studies inspected student behaviors and characteristics related to dropout in an effort either to predict dropout as in Halawa, Greene and Mitchell (2014) or to describe the types of students likely to dropout as in Kizilcec and Halawa (2015), but these studies did not examine course and lecture features.

Guo and Reinecke (2014) consider videos in a broader context of course materials and constitute that for four edX MOOCs learners which complete the course navigate in a non-linear way (not along the presented sequence) whereas those that do not complete are more likely to follow the linear navigation. The authors also compare Coursera, edX and Udacity in the hierarchical and sequential way those platforms present information. Perna, Ruby, Boruch and Wang (2014) in contrast identify mainly sequential learners in their characterisation of registrants and starters as well as their differentiation between sequential and user driven progress. More in detail, they draw a positive relation between accessing the a lecture in the first module and the last module for the 16 courses analysed. One of the few contributions shedding light into learner engagement in MOOCs from other areas than North America and Europe can be found with Zhong, Zhang, Li and Liu (2016).

In conclusion, research on student engagement with videos in MOOCs comes in diverse shapes and is grounded in various concepts and theories. For xMOOCs, the video interaction seems central and connected to learner success. Two ways of describing learners and their engagement with the course evolved: either clusters were build based on demographic data or on interaction patterns. Connection between those were made mostly post-hoc, so that for specific demographic clusters interaction patterns were identified. Some authors explored the overall pathways which learners take through a MOOC and how they interact with different components of the content.

Results are mixed, pointing either at the fact that learners follow the given linear path or skipping it and taking short cuts.

5. Theoretical model on learning and engagement in MOOCs through videos

With the rise of MOOCs as research objects for educational research, different perspectives on which learning theories suit best to account for the concept of learning and learner engagement emerged. Terras and Ramsay (2015) address the importance of individual learner attributes and the perception of learning online from a psychological stance. Siemens (2014) identifies connectivism as *the* learning theory for the digital age, which was already addressed by others (cf. Williams, Karousou, & Mackness, 2011; Downes, 2012). Ho et al. (2014) postulated that there will be no overarching theory of learning in MOOCs. Along the lines of Perry and Nichols (2014) as well as Glaser et al. (2001) different theoretical perspectives explaining the phenomenon of learning in (online) learning environments can be considered. Influenced by the complexity of the learning concept and the different actors involved in the learning process, those range from cognitive perspectives focusing on the individual learner over social perspectives highlighting learning networks. Tschofen and Mackness (2012) suggest the focus on the individual learning experience as fruitful for MOOC research.

The literature review has shown that being explicit in concepts being used is essential to frame research approach and data analysis. Engagement is related to concepts such as learner activity and learner interaction. Activity as in active learning from a pedagogical stance is not considered, rather the notion of a learner being cognitively active. It is a mental state, which can but does not necessarily have to be connected to interaction and engagement in the learning process. This being said, a learner can be cognitively inactive and interact or engage with content. Interaction is regarded as events in terms of Wagner (1994) whereas engagement refers to the resources (time and effort) learners invest with respect to their academic experience (Jennings & Angelo, 2006; Kuh, 2003). Anderson (2003) identifies interaction as significant for processing information, learner control and constructing knowledge. Engagement in interaction as well as the quality and quantity of interaction have been described as central elements of learning experience and learner satisfaction (Rhode, 2009; Khalil & Ebner, 2013). Rubens, Walz, and Koper (2014) describe three different level of interaction: learner-content, learner-instructor and learner-student. The first level happened to be described by Moore (1989) as characterizing feature of education in a way that it promotes internalization and processing of information.

The theoretical lense of this work focusses on the individual learner based on the concepts of motivation and engagement and their importance for learning, knowledge and knowledge transfer. Motivation is an important precondition for engagement. It is acknowledged that engagement with the learning content is an important part of learning in online environments. It is argued that a learning environment can trigger the same process of knowledge construction independently of the technology used to provide. Furthermore it is acknowledged, that the analysis of learner interaction with the content is one component of many and that the learning space is expanding beyond the online environment. As discussed in Perna et al. (2014) this thesis carefully selected resources for the theoretical model of engagement and motivation in MOOCs due to the difficulty of applying findings from distance-education and online-education. It is argued that traditional distance-education is not meant to attract a massive number of learners nor is it open in terms of entry requirements and (un-)enrollment. MOOCs are less homogenous, coordinated and monitored and learner motivation might be different.

Learner motivation and engagement as well as underlying factors are explained in line with Hew (2016/2015) who bases his model on the self-determination theory (SDT). Even though different frameworks for participation in online learning environments exist, the one developed by Hew appears particularly useful as it is based on research into MOOCs and the deviated results. Also, compared to other proposals, Hew does not cluster learners based on their interaction with the course but rather describes the complexity and interdependencies of this particular learning context (Hughes & Dobbins, 2015). Furthermore, this model acts independently of the learners success (for MOOCs often described as successful certification or passing) and emphasizes the motivation of the learner as underlying factor for success.

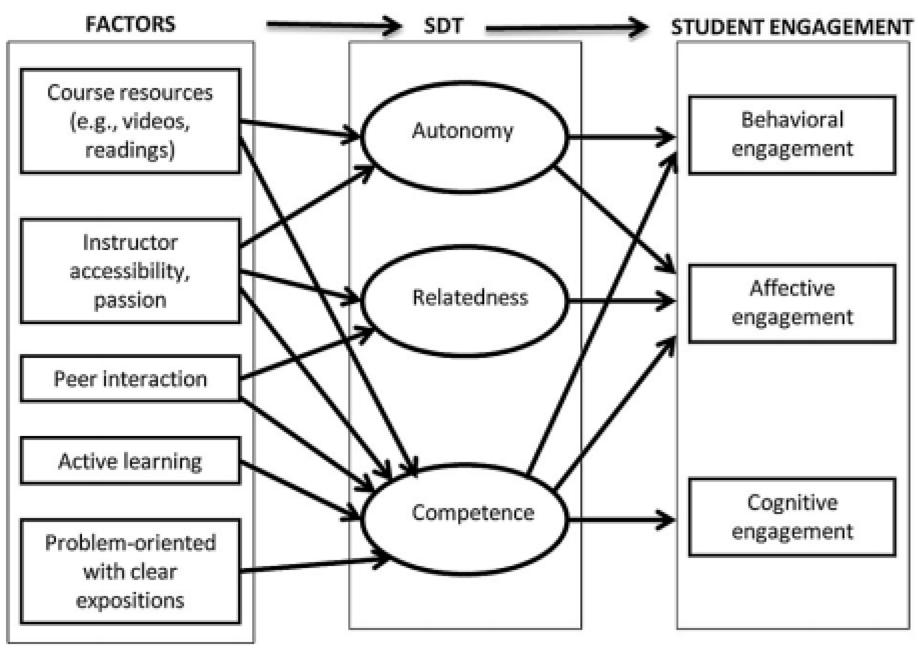

Figure 6: Model of student engagement in MOOCs (Hew, 2016, p. 337).

Figure 6 visualises the model. Learner engagement has been identified as important for retention, learning and achievement (Appleton, Christenson & Furlong, 2008; Fredricks, Blumenfeld & Paris, 2004). Based on a literature review conducted by Fredricks et al. (2004) the construct of learner engagement consists of three components, namely behavioral, affective and cognitive engagement. Behavioral engagement refers to what learners are doing and how they participate in activities. Affective (or emotional) engagement reads as how learners are feeling with respect to the course, the instructor, other peers or the learning process itself. Finally, cognitive engagement describes how and what learners are thinking task-specifically. Even though those different types of engagement are displayed as separated, they are indeed interrelated and connected with each other. Also, engagement is defined as the observable demonstration of motivation, whereas motivation is defined as the underlying reason for a response (Reeve, 2012; Darr, 2012).

In line with this, the different types of engagement can be explained by the self-determination theory. According to SDT, learner engagement depends on the three factors autonomy, relatedness and competence. These three fundamental psychological needs need to be fulfilled to generate a response by the individual. Autonomy in this context is the need to have freedom to act which when fulfilled translates into behavioral and affective engagement. The learner decides individually to take a MOOC and based on this, emotions are affected positively. The need of relatedness describes the feeling of belonging and connecting to a learning community. When fulfilled, it influences affective engagement with respect to a MOOC. The need of competence has effects on all three - behavioral, affective and cognitive engagement. The ability to master MOOC content motivates the learner to continue with the course, produces positive feelings and triggers the thought process of thinking about the performed activities.

With this baseline for learner motivation and engagement established, main factors of MOOCs as identified by Hew (2016) come into play. More specifically, those factors describe characteristics of a MOOC and refer to the course content, the instructor, peer interaction, active learning and problem-orientedness with clear expositions. Each of those characteristics influence learner motivation and engagement in a specific manner. For the purpose of this thesis, it is in particular the course resources that will be examined. More specifically, the main resources for the courses analysed, the video lectures, are examined. They are influencing autonomy and competence of the learner and hence are a significant underlying factor for behavioral and cognitive engagement.

Other models focus on different aspects in particular, such as the role of the instructor (Ma, Han, Yang, & Cheng, 2015) or on particular online learning environments different from MOOCs (Chen, Lambert, & Guidry, 2010; Henrie, Halverson, & Graham, 2015). Still, disagreement exists on how to measure academic progress and learner engagement through MOOCs and most probably further research is needed to address this issue (Raffaghelli, Cucchiara, & Persico, 2015). Engaging learning content especially in MOOCs and how learners approach this content can reveal meaningful insights about the learning process, and to some extent about the learning success. Whereas videos and the interaction with them is limited, they have repeatedly proven to be a component which is not only considered central by instructors in MOOCs but also mainly used in these courses by learners. Learning comes in different facettes, engagement with videos can represent a basic level of the learning process and knowledge transfer. They represent the baseline for deeper understanding, application and reflection in sequential learning environments. Deviations from the planned instruction might in addition indicate different learning preferences but effectiveness of supporting the learning process. Focussing on this video interaction, the active process of learning is highlighted in contrast to the pure focus of knowledge being a product (e.g. a quiz result or a passing grade). The model proposed by Hew (2016) represents a framework which incorporates all of these factors based on the self-determination theory.

6. Method

6.1. Literature Review

The aim of the literature review is to systematically gain an overview of existing academic literature associated with key aspects of this thesis. It answers the questions why the research question and hypotheses are still valid in the light of these key themes and which methodological approaches are useful to address them. The review does not aim at being exhaustive but at identifying substantial contributions, their results and implication for this research for filling existing gaps. Thus, it provides the argumentation for the relevance of this study with reference to scientific literature, analyses the results and applies the knowledge in a systematic and critical way. The goal is further to produce a well supported and significant contribution to the academic field. A broad perspective on the understanding of research field is taken, which results in contributions from different fields and thus recognizing the interdisciplinarity of educational technology and data analytics. Work of others is analysed in the light of the way of presentation, the content, the novelty of results and the methodology used. Throughout the literature search a tracking tool has been used to cluster and annotate the papers as well as make visible cross references. For the literature review, a mixed, three-fold approach has been chosen. This consisted of a lead-article as a starting point, key word search in relevant databases and the skimming of meta-articles covering research in MOOCs. Iterative reference checks have been conducted until no further relevant references came to light.

In the first place, resources used for the lead article (Guo et al., 2014) have been scanned systematically to identify potential resources in a snowball approach. Secondly, research databases have been fed with the keywords MOOC or Massive Open Online Courses respectively, and video or lecture and engagement and higher education. These keywords have been selected carefully after several iterations and test searches. The databases used were Web of Knowledge, ProQuest, Scopus, IEEE, ERIC, Education Research Complete, The Wiley Online Library and The Learning and Technology Library. Results were also crosschecked with publications from edX's own research consortium. Finally, three meta-articles generally analysing research in the field of MOOCs have been identified and their resources have been taken into account. Resources had to be peer-reviewed journal articles written in English and published until the end of March 2016. All together, the initial search identified over 800 articles, which have been scanned to identify those specifically focussing on learner engagement on videos and going beyond the analysis of course demographics.

A starting point was the cited related work, with resources being excluded which were not published in peer-reviewed journals. Keynote and proceedings from relevant conferences have been excluded from database research. The same applied for research not being conducted with data from MOOCs. Guo et al. (2014) point out that theses resources do not focus on videos in particular. This applies for Coetzee et al. (2014) focussing on forum interaction or Koutropoulos et al (2012) investigating emotive language in discussion forums. Another example is Mak, William, & Mackness (2010). Furthermore, articles aiming at anticipating or predicting student behavior have been excluded (cf. Yang, Sinha, Adamson, & Rose, 2013; Balakrishnan & Coetzee, 2013). Also, research focussing on learner contributions in social media and the connection to learner performance have been excluded, even if these contributions consisted of videos (e.g. Fournier, Kop, & Durand, 2014; Alario-Hoyos, Muñoz-Merino, Pérez-Sanagustín, Delgado Kloos, & Parada, 2016). The three meta-article skimmed for additional resources where Liyanagunawardena et al. (2013) and Gaševic, Kovanović, Joksimović, and Siemens (2014) as well as Veletsianos and Shepherdson (2016).

Kop (2011) found that engagement in synchronous visual session was lower than the engagement in discussion forums. Due to the differences in set-up and interactive components between xMOOCs and cMOOCs those findings have not been considered for the literature review (e.g. Fini, 2009; Kop & Carroll, 2011; Levi, 2011; Milligan et al., 2013; Rodriguez, 2012). As the learning environment of a MOOC is considered specific, articles focusing on closed online learning environments but on videos have not been considered either (e.g. Romanov & Nevgi, 2007; McNulty et al., 2009). In addition, due to limited results complying with the requirements of peer-reviewed journal articles, finally conference proceedings and other contributions were considered if they have been identified as often cited by other authors.

6.2. Research Approach

By combining the results from the literature review and Hew's (2016) theoretical model for learner motivation and engagement in MOOCs through videos, the research approach has been chosen. Based on Perry and Nichols' (2014) design continua for classifying research, this research approach can be classified as exploratory, quantitative and applied. The main reason for choosing an explorative data analysis is that it was the actual course data that should be analysed from different sources but beforehand it was nearly impossible to tell if the available data quantity and quality would be sufficient to answer generated research hypothesis.

As an example, in the early phase of the project, hypothesis focussed on interaction between learners and their impact on student engagement. However, in the iterative process of forming hypotheses and locating data to validate these hypotheses, it became obvious that the data generated based on learner interaction was not representative and satisfactory to respond to the research needs. Another line of development was the idea to analyse MOOCs to better understand how participants define course success and which aspects of the data correlate with course success. While the limitation by data availability was one point, another aspect was the limitation based on the edX platform and the data extraction. Some data points were not measured or it was not possible to extract them in a meaningful way. Also, some data points were pre-calculated by edX, which implies that not only structure and content is impacted by the platform but also data analytics and calculations. Results calculated by the platform tools had to double checked with the results calculated from raw data.

This explorative data analysis was beneficial from different points of view. First of all, it was the first time that the data for the three courses has been analysed in this way. Therefore, developed hypothesis and questions had to be tested against reality of the data quality and were adapted in a way that they could be answered by the existing data or by new data collection of the same data set. Those instances where data was not sufficient where documented to guide further research and hypothesis formulation. It was a development from generalistic statements and requirements towards specific hypothesis under the umbrella of student engagement in massive open online courses.

In a systematic approach, context, challenges and opportunities have been defined. The scope was to understand and analyse the data available. Involved in the project where researchers, teachers, course developers and professors. First and foremost, the challenge was to understand and cluster the data to generate an initial overview of potential sources and pitfalls in calculation and interpretation. Secondly, the understanding of analytical options was critical, as a wide range from conservative approaches to social learning analytics was available. Besides the availability own skills and experience as well as sources for support would determine these options. Another aspect was the challenge of over-interpretation. To avoid this, new ideas resulting from brainstorming would be tested directly and discussed to see if they are meaningful with respect to the overall picture. By doing so, the data was explored and expert discussions were used to validate the quality of the hypothesis and the potential if the data could support in investigating them. For the data access generally it was possible from the beginning to have indirect access to all sources via a data specialist and the project leader of the MOOC project. Questions emerged in which format the data would be available and which tools could be used for the analytic part.

In conclusion, the focus of the explorative data analysis moved from student performance to student engagement as the focal point. Due to close contact with the course responsibles it was also possible to get direct feedback on the analysis results. Generally speaking, from a data cycle perspective (cf. Siemens, 2013), the aspects of data analysis, representation and visualization as well as action and loop-back have been applied. Data collection and acquisition, storage and cleaning have been executed with the support of a data expert. In line with Veletsianos, Collier and Schneider (2015) the data log files of a MOOC platform are limited and there are no context variables available. From a data privacy perspective the consent from the project leader has been obtained and for all student data anonymisation has been guaranteed. For detailed discussions on ethics, privacy and data refer to Zimmer (2010) and Robbins (2013).

6.3. Data Structure

Data did result from three main sources: edX insight, edX studio and data from the edX data packages provided by a data specialist. Those sources were combined and then analysed in SPSS. In close collaboration with a data analyst and strictly following rules of anonymization, this thesis builds on edX data extracted from the edX data packages. In the course of the analysis, KI was collaborating with Chalmers to benefit from the established methods of data extraction and data processing. The script used for extracting the data from the edX analytics server was made publically available on github as "ChalmersX Analytic Toolkit". Additional information was manually extracted from edX insights (video name as on edX and position in the course) and edX studio (video uploads list for matching video length and video ID automatically). In a first step, this data collection, processing and the analysis of data with respect to the research question was executed for the two Chalmers courses ChM001x and ChM002x. The preliminary results were taken to the team at Karolinska Institutet to discuss them and identify potential issues and ensure feasibility of an aligned analysis of the for the KIx course. For this purpose, the data specialist supported with his knowledge and based on his scripts, consistency between all master data sets could be ensured and the process of analysis could be executed accordingly. A Master file was created for all three courses including the variables described in Table 2.

Name	**Description**	**Type of Value**
General Video information	Name, ID, Module and Section the video is part of	-
Video length (Time)	The length of the video in seconds, due to deviation in the database actually lengths can deviate by 1-3 seconds	Numeric, Scale
Completed Views		Numeric, Scale
Incompleted Views		Numeric, Scale

Completion Rate per Video		Numeric, Scale
Preceding Quiz		Numeric, Nominal (1/0)
Subsequent Quiz		Numeric, Nominal (1/0)
Feedback Video	Is the video a feedback video, meaning that it has been produced in while the course was running and responds to learners questions from the preceding module	Numeric, Nominal (1/0)
First in Module	Is the video first in the module as defined by the edX layout (left frame, main module)	Numeric, Nominal (1/0)
First in Section	Is the video first in the section as defined by the edX layout (left frame, module sections)	Numeric, Nominal (1/0)
Video Type	Based on Guo et al. (2014) all videos were classified as either classroom lecture, talking head, digital table drawing (Khan Academy Style), Power Point slide presentation, None of the above or Mixed styles. Comments were used to identify the exact style of the last two applied.	Numeric, Nominal (1-6)

Table 2. Overview of values analysed in the video analysis.

Below, the variables calculated by the edX platform are described more in detail. In addition, additional analytical work is indicated. All engagement numbers are based on edX calculations for the the time frame between end of April and beginning of May 2016. For two courses it was tested if the difference in actual end of the course and date of data collection resulted in significant changes in viewing numbers and completion rate. This was not the case, on average for ChM001x and ChM002x the difference in time inpoint resulted in three more completed views, four more incompleted views and a -0.35% change in completion rate. The engagement computations express the number of unique users which complete an activity during a week or during the time from course start to course end respectively.

Calculation of video length

For the calculation of the video length all video IDs have been identified and then paired with the video duration as available in the Video Uploads overview of edX studio. Compared to the video length shown to the learner via the player in the course this length can deviate from one two three seconds. In a test analysis, these differences for the correlation results of one course led to the difference of length of 01:53 minutes in total, which changed the correlation by maximum 0.001 points. Based on this test, it is assumed that time difference between time displayed by player to the user and actual time of the video as indicated in the Video Upload overviews does not have a significant effect on the results.

Completed Views

Completed views have been taken as calculated by edX insights Engagement for Videos (see edX, 2016). This variable refers to the number of unique viewers at the point the video is either 30 seconds apart from the end or at the 95% completion mark.

Incompleted Views

Incompleted views have been taken as calculated by edX insights Engagement for Videos (see edX, 2016). The incomplete views result from the number of unique viewers which started playing the video minus the number of those who were playing the video near its end.

Completion Rate per Video

Completion rate per video has been taken as calculated by edX insights Engagement for Videos (see edX, 2016).

Preceding / Subsequent Quiz

Indicates if a quiz is part of the same subsection as the video and does either precede the video directly or follows the video directly. If a video for example is followed by a quiz but this quiz is in the next subsection, this is not considered as a subsequent quiz.

Feedback Video

Video is produced during the running course and covers questions of the learners which emerged during the preceding week, The video explains details related to those questions. The content of the feedback video is not created in advance and it is released after the course start.

First in Module

Video is placed first in a module of a course. There can only be one video first in each module.

First in Section

Video is placed first in a section of a course. For a module there can be several videos first in section, however only one video can be first for each section.

Video type

All 169 videos have been watched and clustered into different video types based on Guo et al. (2014). The authors identified classroom lecture, talking head, digital tablet drawing and PowerPoint slide presentation as the dominant video styles. If video styles did not match those four groups, they were indicated as either none of the above or mixed styles with a detailed description of this mixed style. After all, the majority of the videos have been identified as talking head combine with PowerPoint slides or a whiteboard. It was consequently decided to exclude video type from the analysis due to the lack of sufficient data points.

In addition, for each course active users and the course components they engaged with have been extracted. An active user is each unique user who engages with one of the categories shown. In this context it does not matter, how often the user engages with the content. 20 completed problems result in the increase by 1. For the attribute “active”, any page in the course counts as an activity (e.g. a forum contribution, an answer submission, the view of a video). Enrolling and unenrolling are excluded. For videos specifically, this means that a learner has clicked play for at least one course video on the edX platform. For attempted problems the same applies: each unique learner who gave an answer to at least one problem is counted.

Finally, a decision had to made with respect to the appropriate correlation factor based on the different hypothesis and variables. For video length correlated to completion rate and completed views, Pearson correlation is used due to the fact that this data is continuous data. Further, for completion rate and completed views in correlation to feedback video, first in module, first in section, preceding quiz and subsequent quiz Kendall’s Tau correlation coefficient is used.

6.4. Data Basis

6.4.1. Summary of MOOCs analysed[4]

This thesis analyses data from three MOOCs by two Swedish Universities which ran on edX between March and October 2015. Two course were released by Chalmers University of Technology in Gothenburg (Chalmers) and are called ChM001x: Introduction to Graphene Science and Technology as well as ChM002x: Sustainability in Everyday Life. Karolinska Institutet University (KI) - a medical and health science university based in Stockholm - run KIx: KIUrologyx Introduction to Urology. All three MOOCs are hosted on the edX platform. All data are extracted based on edX log files with the support of a data expert to ensure anonymity and consistency. This introduction illustrates an overview of these courses, whereas more detailed course descriptions and demographic information can be found in the respective subsection of this chapter. Both, Chalmers and KI presented their general motivation for their MOOCs initiatives. Whereas the former aimed at responding to three educational trends globalisation, individualism and digitalisation; the latter grounded the initial motivation for a healthcare course in the need for an introduction for non-EU/EES citizens to the Swedish healthcare system and to health and caring careers within this environment respectively (UKÄ, 2016b). In their survey summary which was carried out together with all Swedish HEI, UKÄ summarised the main motivations for MOOCs initiatives in Sweden. Besides marketing and branding for the university, development of technology, pedagogy and teaching materials, broader student

[4] Parts of this chapter have been taken from two submitted manuscripts (Stöhr, Stathakarou, Müller, Nifakos, & Mcgrath, 2016; Müller & Stöhr, 2016).

recruiting and collaboration with other universities are named among those universities already giving or planning to give MOOCs (UKÄ, 2016a, p. 20). These motivations complete the perspective given during the report presentation. From a broader perspective these motivations do not deviate from those of other European HEI: In their survey conducted for the European University Association Gaebel et al. (2014) found international visibility followed by student recruitment to be the main motivators, whereas financial aspects where less more common.

Table 3 summarize the main important aspects of the three course at a glance. Two courses lasted seven weeks, one nine weeks. The highest number of enrolled students hovered around 5.000 and 9.000, the completion rate for the KIx: KIUrology was more than twice as high as for the other two courses. The number of analysed videos for the courses ranged between 70 for ChM001x: Introduction to Graphene and Technology, 52 for ChM002x: Sustainability in Everyday Life and 47 for KIx: KIUrology. Total video time was highest for the ChM001x course with nine hours and 20 minutes, lowest for the KIx with three hours and 51 minutes. The average length per video was lowest for the KIx course, with the highest average video completion rate. The two other courses had higher average video lengths and lower video completion rates.

Course Code: Course Name (Term)	Duration in weeks	Highest number enrolled students	Course completion rate	Number of videos analysed	Total video time	Ø video length	Ø video completio n rate
ChM001x: Introduction to Graphene Science and Technology (1T2015)	9	9.733	4%	70	09:20:15	08:00	73.53%
ChM002x: Sustainability in Everyday Life (2T2015)	7	8.377	2%	52	06:32:51	07:33	73.26%
KIx: KIUrology (3T2016)	7	4.916	10%	47	03:51:44	04:56	88.60%

Table 3. Overview of key figures for the analysed MOOCs.

The course descriptions are presented in an adapted version of the MOOC canvas developed by Alario-Hoyos et al. (2014) in Table 4. A general course description is given based on the description publicly available on edX. The focus lies on age distribution, gender, educational background and country of origin as well as socio-economic status, ICT profile and motivation and intention of the students (Kalz et al., 2015). By doing so, a common framework enables easier comparison of student characteristics and a critical debate if comparison of the courses it reasonable. Additionally, enrolled learner data is contrasted with the same demographics for the learners finishing the course where applicable. The guiding question here was if the population starting the course differs from the population receiving a certificate.

Overall course description	
General description	*Name* *Duration* *Field/Area*
Target learners	*General course requirements*

Pedagogical approaches	*xMOOC / cMOOC*
Objectives and Competences	*Course objectives and outcomes*
Learning Contents	*Contents of modules / Short description of modules*
Assessment Activities	*Description of assessment activities*
Complementary Technologies	*Additional components apart from the characteristic MOOC components*
Demographics	
	Age, gender, educational background, country of origin (comparison of enrolled and graduating students)
	Motivation and intention (where data available)
	Socio-economic background (where data available)
	ICT profile (where data available)

Table 4. Framework for MOOC description and demographic comparison.

The demographic data is based on a commonly structured data file for demographic data of enrolled students. The size of the total sample is indicated, for further analysis the sample size is reduced by blank and “null” fields and if applicable fields are grouped to give a better overview.

6.4.2. ChM001x: Introduction to Graphene Science and Technology

ChM001x: Introduction to Graphene Science and Technology was offered by Chalmers University of Technology and was at the same time the first MOOC offered by this university (Stöhr, 2016). Even though the course title indicates otherwise, the course covered advanced content with no complementary course resources available (as for example a course book). Overall, ChM001x aimed at forming an elemental understanding of graphene characteristics, synthesis process and future areas of application, with a focus on electronics. As can be seen from Table 5, the course ran from March 23rd until June 1st 2015 and consisted of the common course components offered by the edX platform: video lectures, self-assessment quizzes, weekly assignments and automatically graded multiple-choice questions as well as question including exercises for calculation. In the final exam, an article about graphene had to be read and interpreted. There were weekly update videos based on learner feedback and questions during the week, which were made accessible at the beginning of each module. For an honor code certificate, at least 60% of the graded assignments had to be completed. The overall workload of the course was indicated with six hours per week, nine hours and 20 minutes in total were comprised by videos. The sample size for ChM001x based on the data file extracted was 10572 unique students. In the end, 360 learners would finish the course with a certificate, the average passing grade being 80%. When specific attributes have not been available, the sample size (N) has been adapted as indicated respectively. Besides the data file, some demographics have been extracted from the post-survey of the course. In addition to the basic information on the course, this chapter covers a comparison of the distribution of demographics in the enrolled student population and the cohort the finally finished the course.

Overall course description	
General description	*Name:* ChM001x: Introduction to Graphene Science and Technology *Duration:* 9 weeks, 23rd March 2015 - June 01, 2015 *Field/Area:* Physics
Target learners	- Appropriate knowledge of english - University knowledge of mathematics - Advanced knowledge of general physics and chemistry - Workload: 6 hours per week, 54 hours in total, ~9 hours videos
Pedagogical approach	xMOOC
Objectives and Competences	- Basic understanding of the fundamental characteristics of graphene, the synthesis process and the future applications of graphene, particularly within electronics - Explain how to perform graphene material synthesis experiments - Assess if and how graphene can be used in different electronic products to improve, enhance or add new properties - Comprehend scientific articles on graphene research - Identify existing and new areas of graphene application and evaluate its potential - Evaluate whether or not graphene can contribute to your career advancement – be it in academia or industry - Learn about peer assessment
Learning Contents	Module 1: Introduction to graphene Module 2: Graphene synthesis 1 Module 3: Graphene synthesis 2 (chemical vapor deposition) Module 4: Graphene transparent electrodes in optoelectronics 1 Module 5: Graphene transparent electrodes in optoelectronics 2 Module 6: Electronic structure of graphene Module 7: Experiments: Graphene synthesis by CVD
Assessment Activities	● Interspersed exercises or problems ○ Quizzes first ○ Peer assessment ● Homework assignment for the first six weeks ● Final assignment ● Certificate passing grade: 60%
Complementary Technologies	Weekly update videos based on student feedback during the week
Overall structure of the course	● Lecture videos paired with quizzes (most of the lecture videos were a combination of talking head and presentation of powerpoint slides) ● Feedback videos referring to learner questions during the week were released at the beginning of each module starting from module 2
Demographics (N= 10572)	
Gender (N=9061)	83% male, 17% female (enrolled) 87% male, 13% female (certificates generated)
Age (N=8901)	Average age: 30, Under 25: 44%, 26-40: 41%, Above 41: 15% (enrolled)

	Similar distribution for certificates generated
Educational Background (N=9029)	Bachelor: 33%, Master: 27%, High school: 25% (enrolled) Bachelor: 32%, Master: 32%, High school: 18%, PHD: 12% (certificated generated)
Country of Origin (N=8966)	India: 8%, USA: 8%, Spain: 3%, Sweden: 3% (enrolled) Spain: 26%, Sweden: 15%, USA: 8%, India: 7% (certificates generated)
Motivation and Intention (N=1300)	*Content Familiarity* Mathematical concepts, Basic chemical concepts, Basic electronical elements, Modern physics concepts: more or less equal distribution between novice and expert (rating from 1-5) Nanofabrication and characterization methods, basic concepts in material science, basic solid state physics concepts: novice category around 40% for each item (rating from 1-5) *Motivation* Over 60% of the learners wanted to extend their current knowledge of the topic and 25% thought the course would be fun and enjoyable. *Intention* I plan to watch all lectures, complete all activities, engage actively in the forum and earn a state of accomplishment: 31% I plan to watch all lectures, complete all activities, and earn a statement of accomplishment: 44%
Socio-economic background	I am geographically isolated from educational institutions Not at all true: Around 70% I cannot afford to pursue a formal education Not at all true: Around 60%
ICT profile	*Familiarity with MOOCs* 75% of the survey participants were not familiar with MOOCs or had heard of them but paid little attention. Whereas almost 75% had to some extent engaged with MOOCs, more than 50% never completed a MOOC.

Table 5. Description of ChM001x.

6.4.3. ChM002x: Sustainability in Everyday Life

ChM002x: Sustainability in Everyday Life was also offered by Chalmers University of Technology. It ran over seven weeks from the 8th of June to the 27th of July and covers topics broadly related to the field of environmental science. More detailed information can be extracted from Table 6. The motivation to developed this course from scratch as well, was the university's emphasize of the importance of sustainable development which is the driver and influencer of many other areas. At introductory level, this course aimed at conveying the complexity of sustainability in everyday life as well as enabling learners to make profound and well-informed decision in relation to this topic. The setup of the course was very similar to the one of ChM002x. Also, there was no supplementary course material in form of e.g. a book available. Offered via the edX platform, it contained video lectures, self-assessment quizzes, weekly assignments and automatically graded multiple-choice questions. Each module was build around a so called hotspot videos and different mini lectures with graded multiple-choice assignments and the final exam in form of a peer review. As in ChM001x, there were weekly update videos based on learner feedback and questions during the week, which were made accessible at the beginning of each module. For an honor code certificate, at least 60% of the graded assignments had to be completed. The overall workload of the course was indicated with six hours per week, six hours and 30 minutes in total were comprised by videos. The sample size for ChM002x based on the data file extracted was 8153 unique students. When specific attributes have not been available, the sample size (N) has been adapted as indicated respectively. Besides the data file, some demographics have been extracted from the post-survey of the course. In addition to the basic information on the course, this chapter covers a comparison of the distribution of demographics in the enrolled student population and the cohort the finally finished the course.

Overall course description	
General description	*Name:* ChM002x: Sustainability in Everyday Life *Duration:* 7 weeks, 8th June 2015 - 27th July 2015 *Field/Area:* Environmental Science
Target learners	- Learners generally interested in the topic of sustainability and how to relate it to everyday decisions - Passed compulsory school of at least 9 years - Be comfortable using electronic devices - Workload of 6 hours per week, 42 hours in total, 6,5 hours videos
Pedagogical approach	xMOOC
Objectives and Competences	- Introduce sustainable development to support everyday life sustainability decisions - Appreciate the complexity of sustainable development and understand how it relates to everyday life - Critically evaluate and reflect on the information flow from the public media - Develop cognitive and decision making skills that can be applied to issues and problems in everyday life - How to discuss these topics appropriately, and encourage others to make informed decisions regarding sustainable living
Learning Contents	Module 1: Course Welcome and Energy Module 2: Globalisation Module 3: Climate Module 4: Chemicals Module 5: Food Module 6: Final Exam
Assessment Activities	Quiz and exercises, peer-assessment, certificate passing grade: 60%, the quizzes and exercises related to each introductory hotspot lecture contribute 4% each week making up 20% of the total score, the assignments related to the mini-lectures make up 40% (8% each week), and the final exam also makes up 40% of the total score.
Complementary Technologies	Weekly update videos based on student feedback during the week
Demographics (N=9180)	
Gender (N=8153)	46% male, 54% female (enrolled) 54% female, 46% male (certificates generated)
Age (N=4747)	Average age: 29, Under 25: 33%, 26-40: 47%, Above 41: 20% (enrolled) Similar distribution for certificates generated
Educational	Bachelor: 33%, Master: 27%, High school: 25% (enrolled)

Background (N=9029)	Bachelor: 31%, Master: 37%, High school: 23% (certificates generated)
Country of Origin (N=5853)	USA: 16%, India: 10%, Sweden: 9% (enrolled learners) Sweden: 8%, USA: 7%, India: 5% (certificates generated)
Motivation and Intention (N=)	*Content Familiarity* Not assessed. *Motivation* Over 92% of the learners wanted to extend their current knowledge of the topic and 74% thought the course would be fun and enjoyable. *Intention* I plan to watch all lectures, complete all activities, engage actively in the forum and earn a state of accomplishment: 31% I plan to watch all lectures, complete all activities, and earn a statement of accomplishment: 48%
Socio-economic background	I am geographically isolated from educational institutions Not at all true: Around 68% I cannot afford to pursue a formal education Not at all true: Around 52%
ICT profile	*Familiarity with MOOCs* 78% of the survey participants were not familiar with MOOCs or had heard of them but paid little attention. Whereas almost 75% had to some extent engaged with MOOCs, more than 44% never completed a MOOC.

Table 6. Description of ChM002x..

6.4.4. KIx: KIUrologyx Introduction to Urology

The last course which is part of the MOOC analysis is the course KIUrologyx Introduction to Urology. It is the fifth MOOC prepared, offered and conducted by Karolinska Institute. The course lasted seven weeks between 15th of September and 27th of October 2015. Learners should establish a general basic understanding of urology through the course which was targeted at anyone interested in clinical medicine. The five course sections where comprised of videos followed by quizzes for self-assessment and a discussion board. The final exam decided on the passing grade and those learners scoring 60% or more would receive a certificate. Complementary course components were 3D models, two virtual patient cases and a live webinar. Out of 4916 learners, 514 finally completed the MOOC and received a certificate. Table 7 describes the learner population more detailed.

Overall course description	
General description	*Name:* KIUrologyx *Duration:* 7 weeks, 15th September 2015 - 27th October 2015 *Field/Area:* Clinical Medicine
Target learners	- Learners generally interested in clinical medicine - Workload of 6 hours per week, 42 hours in total, ~4 hours videos
Pedagogical approach	xMOOC

Objectives and Competences	- Develop a basic understanding of urology
Learning Contents	Course Welcome Section 1 - part A : Introduction to Lower Urinary Tract Symptoms (LUTS) Section 1 - part B : Treatment of Lower Urinary Tract Symptoms (LUTS) Section 2: Hematuria - Blood In The Urine Section 3: Pain and/or Lumps in the Scrotum Section 4: Erectile Dysfunction Final Exam
Assessment Activities	Certificate passing grade: 60%
Complementary Technologies	3D models, two virtual patient cases and a live webinar
Demographics	
Gender (N=4747)	38% female, 61% male (enrolled) 28% female, 72% male (certificates generated)
Age (N=4627)	Average age: 34, Under 25: 32%, 26-40: 47%, Above 41: 21% (enrolled) Similar distribution for certificates generated
Educational Background (N=4784)	Bachelor: 33%, Master: 25%, High school 23% (enrolled) Bachelor: 27%, Master: 32%, High school 18% (certificates generated)
Country of Origin	USA: 19%, India: 7%, Sweden: 4% No sufficient data available for certificates generated
Motivation and Intention	No sufficient data available
Socio-economic background	No sufficient data available
ICT profile	No sufficient data available

Table 7. Description of KIx:KIUrologyx.

The main reason for the differences in the demographic data was the fact, that the universities used different survey tools which did not cover the same questions. However, in the exit survey by KI students rated the video lectures as the most helpful component when it comes to reaching the learning outcome.

7. Results

7.1. ChM001x: Introduction to Graphene Science and Technology

Overall, 83% male learners and 17% female learners enrolled in the course. With an average age of 30 years, 44% of the student population were 25 and younger and 41% between 26 and 40.

Also, around 90% came with either a high school diploma, a bachelor's degree or a master's degree. For the learners earning a certificate, the educational background was even higher. The main countries of origin for enrolled learners were the United States and India followed by Spain and Sweden. In contrast, for the learners with a certificate, it was Spain and Sweden contributing to around 40% of the certificates generated. Survey participants were asked to rate their familiarity with basic concepts in physics, electronics, chemistry, material science, mathematics, and nanofabrication and characterisation methods. On a scale from 1 (Novice) to 5 (Expert), for all topics most survey participants would rate themselves between 2 and 3,5. However, with respect to nanofabrication and characterization methods, basic concepts in material science and basic solid state physics concepts the novice group was the strongest. With respect to the motivation, 60% of the survey participants wanted to extend their current knowledge, 25% thought the course would be fun and enjoyable. 31% decided to indicate the highest intention level in the survey, whereas 44% planned to watch all lectures, complete all activities, and earn a statement of accomplishment. The majority of survey participants did not indicate isolation from educational institutions nor that they could not afford formal education. Most of the survey participants were familiar with MOOCs.

Figure 7 presents the overall active students for ChM001x and in what way they engaged with the content. Similar to trends identified in MOOC research, activity declined during the course with videos being watched more than problems attempted. The rapid decline in the beginning (until week five) is slowing down a bit, and also the delta between attempted problems and played videos is decreasing.

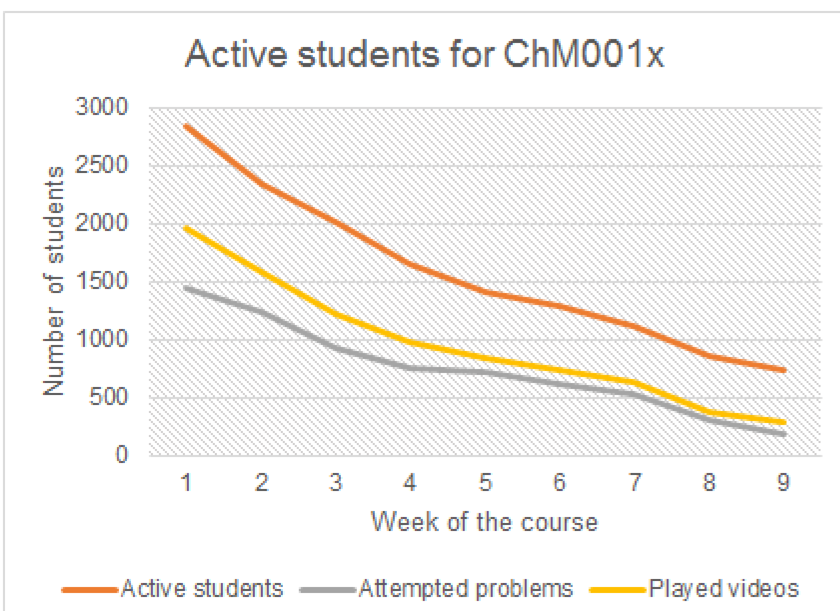

Figure 7: Active students for ChM001x.

For 70 videos in the ChM001x course the following results in Table 8 and Table 9 were calculated. The correlation between Video Length in Seconds and Completion rate was significant at the 0.01 level. As indicated already in the overview of all videos and completion rates, this calculation strengthens the hypothesis that there is a strong negative correlation between how long a video is and the proportion of learner completing the video. The longer the video, the lower the completion rate for the video; a moderate downhill linear relationship exists.

Table 8: Completion rate and correlations to different video characteristic for ChM001x

	Video Length (Sec)	**Feedback Video**	**First in Module**	**First in Section**	**Preceding Quiz**	**Subsequent Quiz**
Pearson Correlation	-.643**					
Sig. (2-tailed)	.000					
Kendall's tau_b		-.116	-.096	-.059	-.044	-.209*
Sig. (2-tailed)		.244	.334	.552	.659	.035
N	70	70	70	70	70	70

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

The hypothesis for the positive effect of completion rate in correlation to feedback videos, early position in modules and sections as well as connection to preceding quizzes has not been strengthened. In other words, if a video is a feedback video, if it is placed early in the syllabus and if it comes after a quiz is not significantly related to how many learners watch this video close to its end. However, for subsequent quizzes a significant correlation has been identified with a weak downhill linear relationship. Generally speaking, a video followed by a quiz has lower completion rates.

Table 9: Completed views and correlations to different video characteristic for ChM001x.

	Video Length (Sec)	**Feedback Video**	**First in Module**	**First in Section**	**Preceding Quiz**	**Subsequent Quiz**
Pearson Correlation	-.388**					
Sig. (2-tailed)	.001					
Kendall's tau_b		-.236*	-.095	.041	.112	-.115
Sig. (2-tailed)		.017	.339	.680	.259	.244
N	70	70	70	70	70	70

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

As completion rate is a calculation based on completed and incompleted views, the correlation in Table 9 follows the trends indicated in Table 8 earlier. Whereas generally, completed and incompleted views do decline in the process of the course, there is no significant correlation when it comes to first in module and first in section position of the video. Both hypotheses related to absolute completed views have not been supported on a significant level, but results indicate a trend. One significant correlation at 0.05 level which has not been hypothesized for this course is the weak negative relationship between completed views and feedback video. This would mean that if a video is a feedback video it has a negative effect on completed views.

7.2. ChM002x: Sustainability in Everyday Life

Overall, 47% male learners and 53% female learners enrolled in the course. The gender distribution for the certificate owners was slightly more even. With an average age of 29 years, 33% of the student population were 25 and younger and 47% between 26 and 40. The majority of learners was between 20 and 30 years old. Also, around 85% came with either a high school diploma, a bachelor's degree or a master's degree. The educational background was even higher for those learners earning a certificate. The main countries of origin for enrolled learners were the United States and India followed by Sweden. In contrast, for the learners with a certificate, most learners came from Sweden, followed by the US and India.

With respect to the motivation, 92% of the survey participants wanted to extend their current knowledge, 74% thought the course would be fun and enjoyable. 31% decided to indicate the highest intention level in the survey, whereas 48% planned to watch all lectures, complete all activities, and earn a statement of accomplishment. The majority of survey participants did not indicate isolation from educational institutions nor that they could not afford formal education. Most of the survey participants were familiar with MOOCs.

Figure 8 presents the overall active students for ChM002x and in what way they engaged with the content. Again, activity declined during the course with videos being watched more than problems attempted. The rapid decline in the beginning (until week five) is slowing down slightly, and also the delta between attempted problems and played videos is decreasing. Compared to ChM001x the activity decline is not as steep in ChM002x.

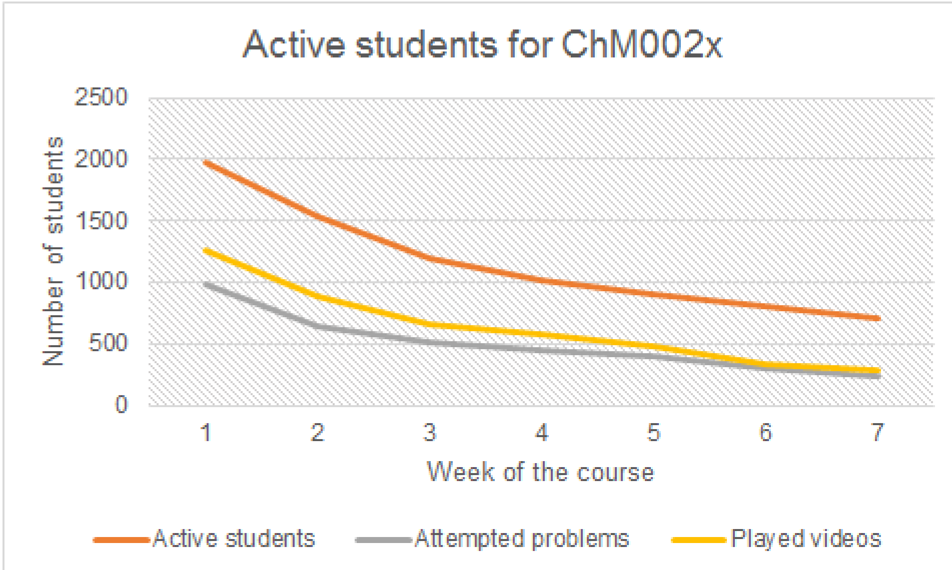

Figure 8: Active students for ChM002x.

For 52 videos in the ChM002x course the following results in Table 10 and Table 11 were calculated. As already for the videos of ChM001x, the correlation between Video Length in Seconds and Completion rate was significant at the 0.01 level for the Pearson correlation. Again, this strengthens the hypothesis that there is a moderate negative relationship between how long a video is and the proportion of learner completing the video. In this case however, the correlation is slightly lower. The longer the video, the smaller the completion rate for the video.

Table 10: Completion rate and correlations to different video characteristic for ChM002x.

	Video Length (Sec)	**Feedback Video**	**First in Module**	**First in Section**	**Preceding Quiz**	**Subsequent Quiz**
Pearson Correlation	-.509**					
Sig. (2-tailed)	.000					
Kendall's tau_b		-.028	-.216	-.177		-.323**
Sig. (2-tailed)		.812	.066	.133		.006
N	52	52	52	52	52	52

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
b. Cannot be computed because at least one of the variables is constant.

A positive correlation between completion rate and feedback video has not been identified. The characteristics First in Module and First in Section do not seem to have an effect neither. There is no result for preceding quizzes and their relation to completion rates, because such videos did not exist in the course. However, there is a weak negative relationship between completion rate and subsequent quiz, which is significant at the 0.01 level. Video completion rates correlate negatively with the characteristic of being followed by a quiz.

Table 11: Completed views and correlations to different video characteristic for ChM002x.

	Video Length (Sec)	**Feedback Video**	**First in Module**	**First in Section**	**Preceding Quiz**	**Subsequent Quiz**
Pearson Correlation	-.132					
Sig. (2-tailed)	.350					
Kendall's tau_b		-.028	-.216	-.177		-.323**
Sig. (2-tailed)		.812	.066	.133		.006
N	52	52	52	52	52	52

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

With respect to completed views and different characteristics, there are no significant correlations but the weak negative relationship between completed views and subsequent quiz. Completion rate and to some extend completed views for this course seem to correlate negatively with videos having a subsequent quiz. In other words, videos followed by a quiz have lower completion rates and completed views. The subsequent quiz correlation for completed views is significant at the 0.01 level, indicating that there is a negative effect between completed views and the characteristic of a subsequent quiz in connection to a video.

7.3. KIx: KIUrologyx Introduction to Urology

Overall, 62% male learners and 38% female learners enrolled in the course. The gender distribution for the certificate owners was even stronger with 72% male and 28% female learners. The learner population had an average age of 34, with 32% under 25 and 47% between 26 and 40. Also, around 80% came with either a high school diploma, a bachelor's degree or a master’s degree. The majority of enrolled learners came from the United States, India and Sweden.

Figure 9 presents the overall active students for KIUrologyx and in what way they engaged with the content. Activity declined during the course with videos being watched more than problems attempted. Compared to the two Chalmers courses however, the activity decline was not as steady and steep and almost reached two plateaus between week two and three as well as between week four and five.

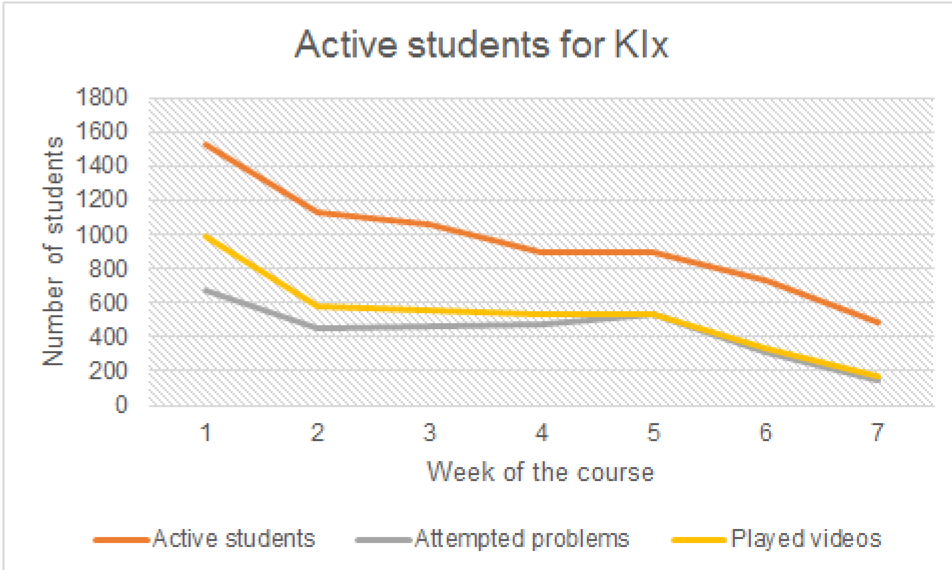

Figure 9: Active students for KIx.

For 47 videos in the KIx course the following results in Table 12 and Table 13 were calculated. Compared to ChM001x and ChM002x the correlation between completion rate and video length is not significant. All in all, this weakens the hypothesis of a negative correlation between those two variables. Also for the other characteristics - feedback video, fist in module, first in section, preceding and subsequent quiz as well as video type, no significant relationship has been detected.

Table 12: Completion rate and correlations to different video characteristic for KIx.

	Video Length (Sec)	**Feedback Video**	**First in Module**	**First in Section**	**Preceding Quiz**	**Subsequent Quiz**
Pearson Correlation	-.262					
Sig. (2-tailed)	.075					
Kendall's tau_b			.158	.104	.038	-.165
Sig. (2-tailed)			.197	.396	.753	.176
N	47	47	47	47	47	47

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
b. Cannot be computed because at least one of the variables is constant.

A week uphill relationship significant at the 0.01 level between completed views and preceding quiz exists, indicating that completed views for videos after a quiz might be higher. In line with the results for completion rate, there is no significant relationship between completed views and one of the other video characteristics.

Table 13: Completed views and correlations to different video characteristic for KIx.

	Video Length (Sec)	Feedback Video	First in Module	First in Section	Preceding Quiz	Subsequent Quiz
Pearson Correlation	-.275					
Sig. (2-tailed)	.061					
Kendall's tau_b			.143	-.136	.266*	-.007
Sig. (2-tailed)			.238	.264	.029	.957
N	47	47	47	47	47	47

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
b. Cannot be computed because at least one of the variables is constant.

In contrast to the other courses, KIx included an AB test with the aim to split a longer video into two separate ones to present to random groups of learners with the same video content in two different ways of sectioning. This AB-test was placed within Section (Module) 1- part B, in the beginning of both, a module and a section. In agreement with the overall hypothesis, it was expected that the longer video would have lower completion rates and possibly also a lower number of incompleted views compared to the two shorter videos.

Table 14: Results from AB-testing for one video for KIx.

	Video Length	Completed Views	Incompleted Views	Completion Rate
Long	17:34	304	63	0.738
Short 1	11:54	170	43	0.828
Short 2	05:44	265	94	0.798
Average Short 1+2	08:49	237	53	0.813

Overall, the longer version had a completion rate with 14 percent points below the course average (87%). The completion rates for the two shorter versions were higher than the longer version, the first part having a close to 10 percent point higher completion rate, and the second part still six percent points higher than the longer version. Completed views for the longer video were 67 views higher compared to the average of the two shorter ones. Even though the second shorter video was shorter in minutes compared to the first one, it did not have a higher completion rate, but higher absolute completed and incompleted views.

8. Discussion and Limitations

All in all, the results have both, supported and weakened the hypothesis made. Some general comments with respect to the overall validity of the results shall be made before the different hypothesis are discussed in more detail. In line with the argumentation of other authors, it is disputable if the edX learner population analysed can account for a general online learner population using videos for learning purposes. It might be that the characteristics described and the results presented characterise a specific subpopulation of (online) learners which for example is curious about new learning formats and might benefit from the setup of a MOOC.

Furthermore, the limitation of the data available also limits the frame of the data analysis. In line with Perna et al. (2014) data plethora can not per se be regarded as a guarantee for qualified insights, especially not if complex relationships and correlations are to be tested. Big data is not necessarily the same as better data (Ho et al., 2014). Limitation of data thus refers to two attributes - the limitation based on the data quality and the limitation of the data interpretation. Connected to these aspects are factors of privacy and ethics. Were further hypothesis might have been meaningful to develop, they could not be tested due to the anonymisation of data. One example is the considered interviewing of learners, which was not possible due to data privacy and a non-existing data privacy agreement.

Another aspect is the interplay of learner intention as planned behaviour and actual behaviour, the pedagogical approaches of the instructor as well as to which extend those aspects can be reflected by the data. Learning is not watching videos, learning is a little bit of everything. The task was to dig deeper into the data provided and establish stable (sub)constructs of learning in the form of engagement with video lectures which could give reason for further analysis. In comparison to other work described in the literature review the methods for data analysis used were manual and basic. The intention of this work was to establish a sound ground for future research and indicate which direction would make sense for this future work. Another justified objection is the question if this work has measured learning or video quality. It is countered, that by strictly separating both - content quality and learning - the construct of learning can not be mapped in online learning environments. With respect to the demographic data and especially to the survey data, is needs to be mentioned that parts might not be representative for the overall learner population in the courses. Additionally, for some parts the comparability between the courses is limited. Even though the same data extraction method was used, survey and additional data sources were not aligned so that for some parts data is missing or to a different extend available for the KIUrologyx compared to the both Chalmers courses. This represents the overall problematic with the comparison of different MOOCs, where only a slightly different set-up results in the question if and to which extent comparison is meaningful.

The descriptive statistics of all three courses reinforce several existing research results and critics in connection to MOOCs but also shed light into differences which might be connected to the fact that these have been offered by a European/Swedish institution. As Grainger (2014) points out, learner profiles are tightly connected to the MOOC subject but general conclusions might still hold for some of them. In the 2014 report, the 2013 prognosis have been confirmed that learners in MOOCs are well educated, often working individuals looking to complement or expand already existing knowledge. He additionally mentions that MOOC offering change over time, as they adapt to insights from earlier course developments. In the course of this progress, it's the early MOOC learners which are likely to be described by more education, better socioeconomic background and higher financial supplies.

All three courses had relatively high dropout rates with completion rates hovering between 2% and 10%. these completion rates are in line with the range of 5% to 12% identified by Perna et al. (2014). With respect to gender distribution, there is now overall conclusion possible, rather the observation, that the Graphene course and the Urology course had a higher part male students, whereas the Sustainability course was more even close to a 50/50 split. For all three courses, the educational background was high, with around 90% of the survey respondents having a high school, bachelor or master degree. When it comes to country of origin for the enrolled learners, the high numbers based in the US and India are in line with research results, too. What does deviate however, is, that learners which received a certificate, are mainly based in Europe, more specifically in Spain. But also learners from Sweden engaged in the courses and represented a substantial part of the learner population receiving a certificate. An educated guess would be that there is a connection between origin of the MOOC offering institution and importance of certification for those learners geographically close to the institution. Most of the socio-economic backgrounds for participants did not represent underprivileged learners. Whereas ICT profiles seem to indicate that MOOCs were known, only few learners finished a MOOC. In addition, motivation related to enjoying the course topic, not so much to the ambition to earn a certification. Even though most participants indicated a high engagement with the course material in the pre-survey, this engagement was not visible from the course data. Grades for the courses were spread out with most learners reaching medium-high grades.

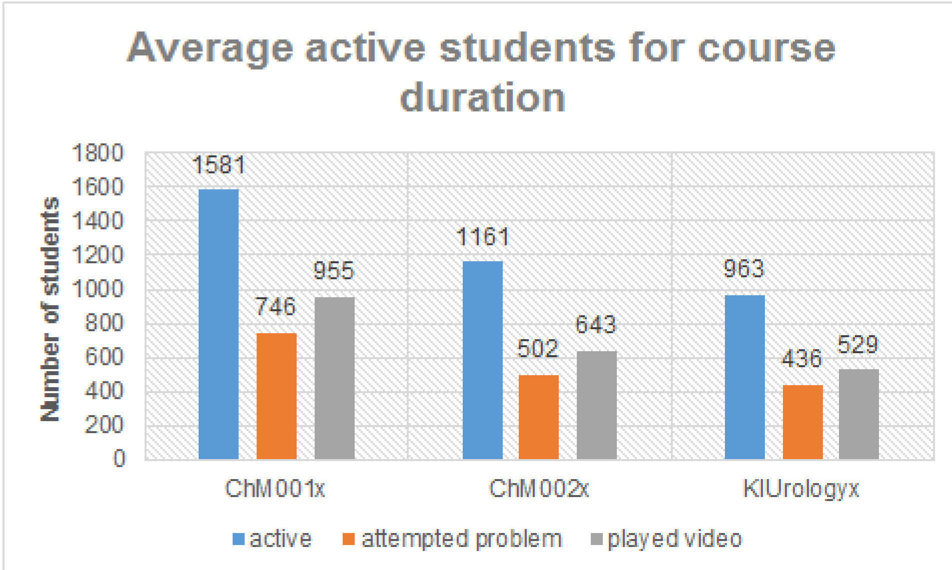

Figure 8: Average active students for all three courses and different course components.

Figure 8 visualises a comparison of average active students for all three courses. Thus, for ChM001x, the most unique active learners were identified. More unique learners played a video than attempted a problem. This distribution can be found back in ChM002x and KIUrologyx, although the differences are a bit smaller. An average of 1581 active students represents only a small fraction of all enrolled students. When breaking activity down into different course components this effect becomes even stronger. Especially for videos and the completed views per video, it has to be mentioned that the significance of the results diminish during the process of the course, as active students and completed views dropped sharply. Nevertheless, results hold for indicating a trend for all three courses. The set-up of the course in a xMOOC format seems to reinforce video watching as a popular learner activity for all courses.

Figure 9 displays a scatterplot framing all video completion rates and video length for the three courses. All in all, the relationship between video length and learner engagement was not as strong as expected. The identified moderate downhill relationship from the previous chapters becomes obvious. Generally speaking, the KIUrologyx course had higher completion rates for the videos compared to the Chalmers course which were nearly similar. However, no reasonable explanation for this could be found based on the analysed characteristics. The scatter plot also indicates potential thresholds with respect to video lengths which could potentially promote video completion. Whereas for ChM001x a drop can be observed at around 12 minutes, for ChM002x around 9 minutes video length. For the Urology course, this threshold hovers around 7 minutes. Indeed it can be discussed, if the pure video length is a meaningful indicator for completion rate and student engagement. It shall be argued that nevertheless, video length is a dominant characteristic of a video and one of the first aspects learner in online courses might look at. The indicated time might be a decision factor for immediate playing of the video, skipping or watching it later. In addition, even though only a small population was involved, the AB testing within the Urology course strengthens the assumption, that video length independently of other characteristics might be an influencing factor for engagement with video lectures.

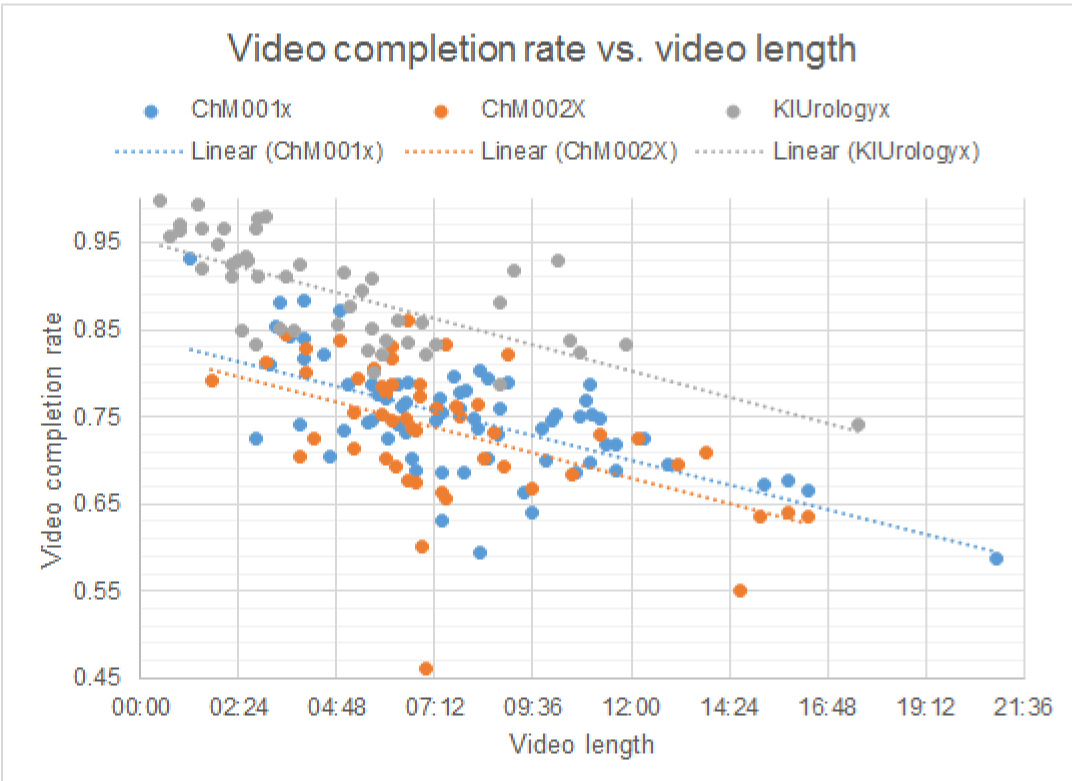

Figure 9: Overview of video completion rate in correlation to video length for all courses and videos.

Another finding for discussion are the feedback videos and their correlation to video completion rate. It was expected that due to their strong alignment and link to learners input and question as well es due to their "newness" learners would interact more with these videos than with normal video lectures. This was not the case and it could not be supported by the empirical data. At least not for both courses containing feedback videos (ChM001x and ChM002x). The weak negative relationship in ChM001x would indicate the opposite of the hypothesis, more specifically that there is a negative correlation between the characteristic of a feedback video and completion rates for the video. Other characteristics such as position in the MOOC or direct relation to a quiz do not seem to have significant impact on video completion rate and completed views. At least for feedback videos, one possible explanation could be, that the feedback videos are not relevant for the entire learner cohort. Those learners starting the video but already knowing the answers or not being interested in the feedback drop out. Due to the limitation in number of questions which could be asked and answered, it is likely that those were only relevant to a minority of the learners.

With respect to video styles it can be noted that there was not significant correlation based on the video styles chosen. The cluster identified by Guo et al. (2014) seem to narrow and not applicable for all video lectures which shall be analysed. Most of the videos were designed in a combination of talking head and the presentation of PowerPoint Slides. Lecturers would vary, so that they changed frequently, discussed together or invited others for interviews.

In conclusion, Hew (2016) proposed a theoretical model where course content is one important factor for learner motivation and engagement in MOOCs. With videos being one course element, those can account for learner autonomy and competence. The results of this thesis indicate, that some characteristics of videos lead to different tendencies in learner interaction with them. Even though videos are only one part of learning in MOOCs and the major part might happen through videos, interaction data indicates how well video lectures are accepted by the learners and also how useful they are. In addition, in MOOCs where learner interaction is limited, videos might also act as a katalysator for instructor accessibility. Also, depending on the quality of the video and the impression of the instructor, videos can transmit passion. Both of these factors are referred to in Hew's model as underlying factors for relatedness. When videos can account for all three motivational needs, they can also trigger the three types of learner engagement. Thus, they create the baseline for the learning process.

9. Future Research and Implications for Design

Recommendations for future research can go in two distinct directions with respect to videos promoting learner engagement in MOOCs. The micro or the macro level. Micro level in this case refers to a more granular analysis of videos, for example by analysing viewing patterns by seconds. Why is it exactly that learners stop the video, and when. Why in this case could related to reinforcing factors for pausing patterns. A superficial first skimming through these characteristics for the analysed videos indicated for example, that some videos showed a lot of stopping and playing for those videos preceeding a multiple-choice quiz with the quiz exactly relating to the content covered in the video. Viewing patterns in this case would simply indicate a strong connection between video content and assessment of this content in the quiz. The question here would be, if the connection of quiz and video only reinforces reviewing patterns, does video watching really promote learning? Or does video watching in this constellation refers to how learners use course resources to pass assessments with the least effort possible. Another interesting future direction could be, if the simple displaying of a video time influences the learner and the decision to watch a video or not. This could be set up for example by hiding the time for one group and displaying it for the other. When exactly is the threshold for video length which influences the decision to start the video. And what kind of content should be covered in the crucial first seconds of the video to sustain learners attention. Is it possible to adapt video styles according to individual viewing patterns? Could the same content be covered in different styles and would this increase video completion rates? Another micro level perspective could be the in-depth analysis of video styles and how they related to engagement figures. This work has shown, that a four-cluster framework for video lectures is too narrow and does not consider mixed styles, different lecturers and more interactive formats such as simulation. With respect to the further development of video quality and content, it seems likely that adapted frameworks can shed light into the question which production styles exactly influence viewing habits and improve video viewer retention. As for the current research status, video styles seem insufficient for using them as an analysis framework. Which video style cluster are suited best for learning environments where new mixed styles are produced and tested frequently? Which style mix promotes learner engagement best? Which characteristic of the style mix could be conceptualised and transferred to other style clusters?

Multi-dimensional models could be of interest as well. Tackling video completion rate it would be interesting which additional variable would result in significant correlations and what this would say about video lenght in connection to other characteristics. The same applies for specific target groups and how they engage with the content. Is there are particular student cluster that engages strongly with videos? What does this say about the importance of videos? Which other influencing factors could be connected to this question?

A debated topic which might deliver valuable qualitative insights into these question is a possible triangulation based on tracking individual student engagement patterns for videos and then interviewing them with specific questions to their engagement patterns. How do they react to video length? What makes them pause a video? Why are they pausing videos? Which role do videos play in the learning process in a xMOOC?

Some design implications identified during the course of the thesis can be supported. It seems like short videos are the key for high completion rates and thus high learner engagement. How short videos specifically have to be, might depend on the content and the context of the course. The results of this thesis indicate, that completion rates drop significantly at video length between 7 and 12 minutes. It appears that shortness is the key for retaining video watchers. This of course affects the content presented in those videos. With time limitations it is not possible to present the same content as one could present in longer videos. Thus, these shorter “learning bites” could only touch the surface of detailed and complex topics and would act as an introduction. Other course components instead would take over the role of a content deliverer, such as texts, discussions, chats or other suitable artifacts. This would also have implications for the overall importance of videos in xMOOCs. Will they still be the center of attention in future? Will MOOCs generally develop to a more video-decentralised set-up?

Finally, the production of feedback videos as analysed in this work does not seem to be worth the effort when success is defined by completion rate compared to normal videos. It became clear, even though those videos might be highly relevant for active learners, they are not affecting the viewing patterns of the learners. Do those feeback videos really enhance learning for the average xMOOC population? What characteristics, if not feedback videos and positions in the MOOC foster learner engagement and the learning process?

10. Summary

In conclusion, this work has shown that Swedish MOOCs can contribute to the research in the field with valuable insights on learner demographics, motivations and intentions as well as with insights on how learners engage with different course components in MOOCs. It has been argued, that course components and their quality are an important factor in the learning process and thus worth being analysed and explored with respect to how learners engage with them. Although this work does only represent an initial investigation of the topic, future research can benefit from approaches firmly grounded in the learning science and with the learner as the focus of all attention and research effort. Also, this research should be closely related to the design of online learning environment to changes focus and procedures on three levels (Reich, 2015). Firstly, research needs to be not only about engagement but about learning. This thesis has shown how complex the construct of learning is and how many diverse factors come into play when operationalizing and analysing it. It has further indicated a way, how to argument for and establish a research approach which attempts considering and uniting this complexity. Secondly, research has to move from investigating individual courses to comparisons across contexts. In the course of this work, it has been highlighted, how challenging a comparison across courses is, even though those courses have been produced and conducted geographically close to each other. Research into MOOCs does not only call for sound research approaches but also frameworks which make it possible to compare diverse MOOCs across institutions, countries, regions and platform providers. Data quality and consistency is a focal point of future research, where in future also agility and adaptability of data sets comes into play. The future of MOOC research depends on experience and knowledge of researchers within the field of learning sciences who embrace data analysis and make sense of the huge amount of data available. Thirdly, the focus on post-hoc analysis will shift towards multidisciplinary, experimental design. The A/B testing in this thesis was only an appetizer of how future research on MOOCs, learner engagement and learning could look like. The aim was to exemplify, how such research designs could look like and to establish a baseline for transferring those designs from a post-hoc to a prediction level. This thesis has also highlighted the importance of interdisciplinary and inter-institutional collaboration. Without the expertise of the many researchers and specialists involved and their genuine interest in learning within MOOC environments, this thesis would not have been possible. All in all it has been demonstrated that massive open online course are and will be an interesting research object for questions related to learning, learner engagement and course component quality.

11. Reference List

Accenture. (2015). Accenture Technology Vision 2015. Accenture.

Adams, C., Yin, Y., Madriz, L. F. V., & Mullen, C. S. (2014). A phenomenology of learning large: the tutorial sphere of xMOOC video lectures. *Distance Education*, 35(2), 202–216.

Alario-Hoyos, C., Muñoz-Merino, P. J., Pérez-Sanagustín, M., Delgado Kloos, C., & Parada G, H. A. (2016). Who are the top contributors in a MOOC? relating participants' performance and contributions. *Journal of Computer Assisted Learning*, 32(3), 232-243.

Alario-Hoyos, C., Pérez-Sanagustín, M., Cormier, D., & Kloos, C. D. (2014). Proposal for a Conceptual Framework for Educators to Describe and Design MOOCs. *Journal of Universal Computer Science*, *20*(1), 6–23.

Allen, I.E. & Seaman. J. (2014). Grade Change: Tracking Online Education in the United States. Babson Survey Research Group and The Sloan Consortium. Retrieved from http://www.onlinelearningsurvey.com/reports/gradechange.pdf.

Allon, G. (2012). 'Operations Management', Udemy. The *Chronicle of Higher Education*, 59(6), B10–11.

Alstete, J. W. (2014). Revenue generation strategies: Leveraging higher education resources for increased income. *ASHE Higher Education Report*, 41(1), 1-138.

Anders, A. (2015). Theories and applications of massive online open courses (MOOCs): The case for hybrid design. *The International Review of Research in Open and Distributed Learning*, 16(6) doi:10.19173/irrodl.v16i6.2185.

Anderson, T. (2003). Getting the Mix Right Again: An Updated and Theoretical Rationale for Interaction. *The International Review of Research in Open and Distance Learning*, Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/149/230.

Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *The International Review of Research in Open and Distributed Learning*, *12*(3), 80–97.

Appleton, J. J., Christenson, S. L. & Furlong, M. J. (2008). Student engagement with school: critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45, 369–386.

Baker, R., & Siemens, G. (2014). Educational data mining and learning analytics. Cambridge Handbook of the Learning Sciences.

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*.

Bali, M. (2014). MOOC pedagogy: Gleaning good practice from existing MOOCs. *Journal of Online Learning and Teaching*, 10(1), 44.

Beaven, T., Hauck, M., Comas-Quinn, A., Lewis, T., & de los Arcos, B. (2014). MOOCs: Striking the right balance between facilitation and self-determination. *Journal of Online Learning and Teaching / MERLOT*, 10(1), 31.

Belanger, Y., & Thornton, J. (2013). Bioelectricity: A quantitative approach Duke University's first MOOC. Retrieved from http://dukespace.lib.duke.edu/dspace/handle/10161/6216.

Bokor, J. (2012). University of the future: A thousand year old industry on the cusp of profound change. Ernst and Young Report. Australia: Ernst and Young.

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom research into edX's first MOOC*. Research & Practice in Assessment*, 8(1), 13.

Brinton, C. G., Rill, R., Ha, S., Chiang, M., Smith, R., & Ju, W. (2015). Individualization for Education at Scale: MIIC Design and Preliminary Evaluation. *IEEE Transactions on Learning Technologies*, *8*(1), 136–148.

Burch, S. L., & Harris, S. E. (2014). A Massive Open Online Course on climate change: the social construction of a global problem using new tools for connectedness. *Wiley Interdisciplinary Reviews: Climate Change*, *5*(5), 577–585.

Burd, E. L., Smith, S. P., & Reisman, S. (2015). Exploring business models for MOOCs in higher education. *Innovative Higher Education*, 40(1), 37-49.

Butin, D. W. (2012). What MIT Should Have Done, *2012*(6). http://doi.org/10.1145/2241156.2263018.

Campagnaro, S. & Müller, F. (2016). Kiron Open Higher Education: Refugee Education Using MOOCs. *Proceedings of the 2th MOOCs in Scandinavia Conference*.

Carlile, O., & Jordan, A. (2005). It works in practice but will it work in theory? The theoretical underpinnings of pedagogy. *Emerging Issues in the Practice of University Learning and Teaching*, *1*, 11–26.

Carr, D. (2013, March 14). edX MOOC Software Goes Open Source. Retrieved from http://www.informationweek.com/software/edx-mooc-software-goes-open-source/d/d-id/1109084?.

Chen, P. -S.D., Lambert, A.D., & Guidry, K.R. (2010). Engaging online learners: the impact of web-based learning technology on college student engagement. *Computer & Education*, 54, 1222–1232.

Clark, D. (2013, April 16). MOOCs: taxonomy of 8 types of MOOC. Retrieved from http://donaldclarkplanb.blogspot.de/2013/04/moocs-taxonomy-of-8-types-of-mooc.html.

Clark, R. C., &. Mayer, R. E. (2011). E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning. San Francisco, CA: Pfeiffer.

Coetzee, D., Fox, A., Hearst, M. A., & Hartmann, B. (2014). Should Your MOOC Forum Use a Reputation System? In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 1176–1187). New York, NY: ACM.

Collins, D., Weber, J., & Zambrano, R. (2014). Teaching business ethics online: Perspectives on course design, delivery, student engagement, and assessment. *Journal of Business Ethics*, 125(3), 513-529. doi:10.1007/s10551-013-1932-7.

Conole, G. G. (2013). MOOCs as disruptive technologies: strategies for enhancing the learner experience and quality of MOOCs. *Revista de Educación a Distancia*, (39). Retrieved from http://revistas.um.es/red/article/view/234221/0.

Cormier, D., & Siemens, G. (2010). Through the open door: Open courses as research, learning, and engagement. *Educause*, 45 (4),(pp. 30-39). Retrieved from http://er.educause.edu/articles/2010/8/through-the-open-door-open-courses-as-research-learning-and-engagement.

Covach, J. (2013). To MOOC or Not To MOOC? *Music Theory Online*, *19*(3). Retrieved from http://www.mtosmt.org/issues/mto.13.19.3/mto.13.19.3.covach.pdf

Daniel, J. (2012). Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education*, 3(0). Retrieved from http://jime.open.ac.uk/articles/10.5334/2012-18/.

Darr, C. W. (2012). Measuring student engagement: the development of a scale for formative use. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds), Handbook of research on student engagement (pp. 149–172). New York: Springer.

De Freitas, S. I., Morgan, J., & Gibson, D. (2015). Will MOOCs transform learning and teaching in higher education? Engagement and course retention in online learning provision. *British Journal of Educational Technology: Journal of the Council for Educational Technology*, *46*(3), 455–471.

Dillahunt, T. R., Wang, B. Z., & Teasley, S. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *International Review of Research in Open and Distance Learning*, 15(5), 15-5.

Dodd, T. (2014, October 6). Top MOOC provider edX no longer free for all. Retrieved from http://www.afr.com/news/policy/education/top-mooc-provider-edx-no-longer-free-for-all-20141005-jlu72.

Doolittle, P. E., Bryant, L. H., & Chittum, J. R. (2015). Effects of degree of segmentation and learner disposition on multimedia learning: Segment length and learner disposition. *British Journal of Educational Technology*, 46(6), 1333-1343. doi:10.1111/bjet.12203.

Downes, S. (2012) Connectivism and Connective Knowledge. Retrieved from http://www.downes.ca/files/books/Connective_Knowledge-19May2012.pdf.

Drachsler, H., & Kalz, M. (2016). The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges. *Journal of Computer Assisted Learning*. http://doi.org/10.1111/jcal.12135.

edX (2016). Computation References. Retrieved from http://edx.readthedocs.io/projects/edx-insights/en/latest/Reference.html#all-computations.

Eisenberg, M., & Fischer, G. (2014). MOOCs: A perspective from the learning sciences. *Proceedings of ICLS 2014*.

European Commission. (2015). Education and Training Monitor 2015: Sweden. Luxembourg: Publications Office of the European Union.

European Commission. (2016). Open Education Scoreboard. Retrieved from http://openeducationeuropa.eu/en/open_education_scoreboard.

Evans, D. (2012). 'Introduction to computer science', Udacity. *The Chronicle of Higher Education*, 59(6), B11.

Evans, B. J., Baker, R. B., & Dee, T. S. (2016). Persistence patterns in massive open online courses. *Journal of Higher Education*, 87(2), 206.

Fischer, G. (2014). Beyond hype and underestimation: identifying research challenges for the future of MOOCs. *Distance Education*, *35*(2), 149–158.

Fini, A. (2009). The Technological Dimension of a Massive Open Online Course: The Case of the CCK08 Course Tools. *The International Review of Research in Open and Distributed Learning*, *10*(5). Retrieved from http://www.irrodl.org/index.php/irrodl/article/viewArticle/643.

Fournier, H., Kop, R., & Durand, G. (2014). Challenges to research in MOOCs. *Journal of Online Learning and Teaching / MERLOT*, *10*(1), 1.

Frank, S. (2012). Review: MITx's online circuit and analysis course. *IEEE Spectrum*.

Fredricks, J. A., Blumenfeld, P. C. & Paris, A. (2004). School engagement: potential of the concept: state of the evidence. *Review of Educational Research*, 74, 59–119.

Gaebel, M. (2014). MOOCs Massive Open Online Courses - An update of EUA's first paper (January 2013). Brussels, Belgium: European University Association.

Gaebel, M., Kupriyanova, V., Morais, R. & Colucci, E. (2014). E-learning in European Higher Education. Institutions: Results of a mapping survey conducted in October-December 2013. Retrieved from http://www.eua.be/Libraries/Publication/e-learning_survey.sflb.ashx.

Gašević, D., Kovanovic, V., Joksimovic, S., & Siemens, G. (2014). Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative. *The International Review of Research in Open and Distributed Learning*, *15*(5). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1954.

Gašević, D., Zouaq, A., & Janzen, R. (2013). "Choose your classmates, your GPA is at stake!": The association of cross-class social ties and academic performance. *American Behavioral Scientist*, 57(10), 1460-1479.

Glance, D. G., Forsey, M., & Riley, M. (2013). The pedagogical foundations of massive open online courses. *First Monday*, 18(5). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/viewArticle/4350/3673.

Glaser, R., Chudowsky, N., Pellegrino, J. W., & Others. (2001). Knowing what students know: The science and design of educational assessment. National Academies Press.

Grainger, B. (2014). Massive open online course (MOOC) report. University of London International Programmes. Retrieved from http://www.londoninternational.ac.uk.

Greeno, J. G., Collins, A. M. & Resnick, L. B. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.), Handbook of Educational Psychology (pp. 15–46). New York: Macmillan.

Grossman, S. (2013). American MOOC providers face international competition. *The Chronicle of Higher Education*, *5*(2013), 8–07.

Grünewald, F., Meinel, C., Totschnig, M., & Willems, C. (2013). Designing MOOCs for the Support of Multiple Learning Styles. In *Scaling up Learning for Sustained Impact* (pp. 371–382). Springer Berlin Heidelberg.

Grunspan, D. Z., Wiggins, B. L., & Goodreau, S. M. (2014). Understanding classrooms through social network analysis: A primer for social network analysis in education research. *CBE Life Sciences Education*, 13(2), 167-178. doi:10.1187/cbe.13-08-0162.

Guo, P. J., Kim, J., & Rubin, R. (2014). How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. In *Proceedings of the First ACM Conference on Learning @ Scale Conference* (pp. 41–50). New York, NY, USA: ACM.

Guo, P. J., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs. In *Proceedings of the First ACM Conference on Learning @ Scale Conference* (pp. 21–30). ACM.

Haggard, S. (2013). The maturing of the MOOC. *BIS Research Paper Number*, *130*.

Halawa, S., Greene, D., & Mitchell, J. (2014). Dropout prediction in MOOCs using learner activity features. *Experiences and Best Practices in and around MOOCs*, *7*. Retrieved from http://www.moocsandco.com/sites/default/files/elearning%2037.pdf#page=7.

Halevy, A., Norvig, P., & Pereira, F. (2009). The Unreasonable Effectiveness of Data. *IEEE Intelligent Systems*, *24*(2), 8–12.

Hansen, J. D., & Reich, J. (2015). Socioeconomic Status and MOOC Enrollment: Enriching Demographic Information with External Datasets. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 59–63). New York, NY: ACM.

Harris, P. L. (2013, April 10). SJSU/EdX Adds More Campuses, Courses. Retrieved from http://blogs.sjsu.edu/newsroom/2013/sjsuedx-expansion/.

Hashmi, A. H., & Shih, C. W. (2013, March 3). EdX To Seek New Revenue Models. Retrieved from http://www.thecrimson.com/article/2013/3/5/edx-funding-virtual-independent/.

Henrie, C.R., Halverson, L.R., Graham C.R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90(1), 36-53.

Hew, K. F. (2015). Towards a model of engaging online students: lessons from MOOCs and four policy documents. *International Journal of Information and Education Technology,* 5(6), 425.

Hew, K. F. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology*, 47(2), 320-341. doi:10.1111/bjet.12235.

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*, 12, 45–58.

Hey, A. J. G., Tansley, S., Tolle, K. M., & Others. (2009). The fourth paradigm: data-intensive scientific discovery (Vol. 1). Microsoft research Redmond, WA.

Hill, P. (2013, March 6). Emerging student patterns in MOOCs: A graphical view. E-Literate. Retrieved from http://mfeldstein.com/emerging_student_patterns_in_moocs_graphical_view.

Ho, A. D., Reich, J., Nesterko, S. O., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). HarvardX and MITx: The First Year of Open Online Courses, Fall 2012-Summer 2013. http://doi.org/10.2139/ssrn.2381263.

Hoadley, C., & Van Haneghan, J. (2011). The Learning Sciences: Where they came from and what it means for instructional designers. *Trends and Issues in Instructional Design and Technology (3rd Ed. , Pp. 53-63). New York: Pearson*. Retrieved from https://www.steinhardt.nyu.edu/scmsAdmin/uploads/006/742/Hoadley-VanHaneghan-draft.pdf.

Hollands, F. M., & Tirthali, D. (2014a). Resource requirements and costs of developing and delivering MOOCs. *International Review of Research in Open and Distance Learning*, 15(5). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1901.

Hollands, F. M., & Tirthali, D. (2014b). MOOCs: expectations and reality. Full report. Center for BenefitCost Studies of Education, Teachers College, Columbia University, NY. Retrieved from: http://cbcse.org/wordpress/wpcontent/uploads/2014/05/MOOCs_Expectations_and_Reality.pdf.

Hollands, F. M., & Tirthali, D. (2015). MOOCs in Higher Education: Institutional Goals and Paths Forward. Palgrave Macmillan.

Holmberg, B. (2001). *Distance education in essence: An overview of theory and practice in the early twenty-first century*. Bibliotheks-und Informationssystem der Carl von Ossietzky Universität Oldenburg (BIS).

Hughes, G., & Dobbins, C. (2015). The utilization of data analysis techniques in predicting student performance in massive open online courses (MOOCs). *Research and Practice in Technology Enhanced Learning*, 10(1), 1-18. doi:10.1186/s41039-015-0007-z.

Hyman, P. (2012). In the Year of Disruptive Education. *Communications of the ACM*, *55*(12), 20–22.

Impey, C. D., Wenger, M. C., & Austin, C. L. (2015). Astronomy for astronomical numbers: A worldwide massive open online class. *International Review of Research in Open and Distance Learning, 16*(1).

Jansen, D., & Schuwer, R. (2015). Institutional MOOC strategies in Europe. EADTU.

Jennings, J.M., & Angelo, T. (Eds.). (2006). Student engagement: Measuring and enhancing engagement with learning [Proceedings of a symposium]. Wellington: New Zealand Universities Academic Audit Unit.

Johnstone, S. M. (2005). Open educational resources serve the world. *Educause Quarterly*, *28*(3), 15.

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, *15*(1). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1651

Jordan, K. (2015, June 12). MOOC Completion Rates: The Data. Retrieved from http://www.katyjordan.com/MOOCproject.html.

Kalantzis, M., & Cope, B. (2012). *New learning: Elements of a science of education*. Cambridge University Press.

Kalz, M., & Specht, M. (2013). If MOOCS are the answer, did we ask the right questions? Implications for the design of large-scale online-courses. Retrieved from http://dspace.ou.nl/handle/1820/5183.

Kalz, M., & Specht, M. (2014). Assessing the crossdisciplinarity of technology-enhanced learning with science overlay maps and diversity measures. *British Journal of Educational Technology: Journal of the Council for Educational Technology*, *45*(3), 415–427.

Kalz, M., Kreijns, K., Walhout, J., Castaño-Munoz, J., Espasa, A., & Tovar, E. (2015). Setting-up a European Cross-Provider Data Collection on Open Online Courses. *The International Review of Research in Open and Distributed Learning*, 16(6). http://doi.org/10.19173/irrodl.v16i6.2150

Karsenti, T. (2013) The MOOC. What the research says, *International Journal of Technologies in Higher Education*, 10, 23–37.

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

Kelly, A. P. (2014). Disruptor, distracter, or what?: a policymaker's guide to massive open online courses (MOOCs).

Khalil, H., & Ebner, M. (2013). Interaction Possibilities in MOOCs – How Do They Actually Happen? International Conference on Higher Education Development. Mansoura: Mansoura University. 1-24.

Kim, P. (2014). *Massive Open Online Courses: The MOOC Revolution*. Routledge.

Kirschner, A. (2012). A pioneer in online education tries a MOOC. *The Chronicle of Higher Education*, 59(6), B21–22.

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1).

Kizilcec, R. F., Bailenson, J. N., & Gomez, C. J. (2015). The instructor's face in video instruction: Evidence from two large-scale field studies. *Journal of Educational Psychology*, 107(3), 724. doi:10.1037/edu0000013.

Kizilcec, R. F., & Halawa, S. (2015). Attrition and Achievement Gaps in Online Learning. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale* (pp. 57–66). New York, NY: ACM.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 170–179). New York, NY: ACM.

Knight, S., Buckingham Shum, S. & Littleton, K. (2014). Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. *Journal of Learning Analytics*, 1(2), 23-47.

Koedinger, K. R., D'Mello, S., McLaughlin, E. A., Pardos, Z. A., & Rosé, C. P. (2015). Data mining and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(4), 333-353. doi:10.1002/wcs.1350.

Kolowich, S. (2013a). How EdX plans to earn, and share, revenue from its free online courses. *The Chronicle of Higher Education*, *21*. Retrieved from http://www.immagic.com/eLibrary/ARCHIVES/GENERAL/CHRON_HE/C130221K.pdf

Kolowich, S. (2013b). Why professors at San José State won't use a Harvard professor's MOOC. *The Chronicle of Higher Education*.

Kop, R. (2011). The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *International Review of Research in Open and Distance Learning*, 12(3).

Kop, R., & Carroll, F. (2011). Cloud computing and creativity: Learning on a massive open online course. *European Journal of Open, Distance and E-Learning*, *14*(2). Retrieved from http://www.eurodl.org/index.php?p=special&sp=articles&article=458&article=457

Korn, M. (2014, October 1). Corporate Training Gets an Online Refresh. Retrieved from http://www.wsj.com/articles/corporate-training-gets-an-online-refresh-1412194344.

Kortemeyer, G. (2013). Ten years later: why open educational resources have not noticeably affected higher education, and why we should care. *Educause Review*, *48*(2).

Koschmann, T. D. (1996). *CSCL, theory and practice of an emerging paradigm*. Routledge.

Koutropoulos, A., Gallagher, M. S., Abajian, S. C., de Waard, I., Hogue, R. J., Keskin, N. O., & Rodgriguez, C. O. (2012). Emotive Vocabulary in MOOCs: Context & Participant Retention. *European Journal of Open, Distance and E-Learning*, 1.

Kovanović, V., Joksimović, S., Gašević, D., Siemens, G., & Hatala, M. (2015). What public media reveals about MOOCs: A systematic analysis of news reports. *British Journal of Educational Technology: Journal of the Council for Educational Technology*, *46*(3), 510–527.

Kuh, G. D. (2003). What we're learning about student engagement from NSSE. *Change, 35*(2), 24.

Kulkarni, C., Wei, K. P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., Klemmer, S. R. (2013). Peer and self assessment in massive online classes. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(6), 1-31. doi:10.1145/2505057.

Lane, A. (2013). The potential of MOOCs to widen access to, and success in, higher education study. In: *The Open and Flexible Higher Education Conference 2013*, 23-25.

Laney, D. (2001). 3D Data management: Controlling data volume, velocity and variety. Meta Group.

Levy, D. (2011). Lessons learned from participating in a connectivist massive online open course (MOOC). In *Proceedings of the Chais conference on instructional technologies research 2011: Learning in the technological era,* 31–36.

Lin, L. (2012, October 26). EdX platform integrates into classes. Retrieved from http://tech.mit.edu/V132/N48/801edX.html.

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distributed Learning*, *14*(3), 202–227.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral*. Retrieved from http://abs.sagepub.com/content/early/2013/03/11/0002764213479367.abstract.

Lombardi, M. M. (2013). The inside story: Campus decision making in the wake of the latest MOOC tsunami. *Journal of Online Learning and Teaching*, 9(2), 239.

Ma, J., Han, X., Yang, J., & Cheng, J. (2015). Examining the necessary condition for engagement in an online learning environment based on learning analytics approach: The role of the instructor. *The Internet and Higher Education*, 24, 26–34.

Mak, S., Williams, R., & Mackness, J. (2010). Blogs and forums as communication and learning tools in a MOOC. In *Networked Learning Conference.University of Lancaster*, 275-285.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. Retrieved from http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation.

Martin, F. G. (2012). Will massive open online courses change how we teach? *Communications of the ACM*, *55*(8), 26–28.

Mauro, A. D., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review. Great Britain. Forestry Commission*, *65*(3), 122–135.

McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). The MOOC model for digital practice. Retrieved from http://www.davecormier.com/edblog/wp-content/uploads/MOOC_Final.pdf.

McGuire, R. (2014, June 19). The Best MOOC Provider: A Review of Coursera, Udacity and Edx. Retrieved from http://www.skilledup.com/articles/the-best-mooc-provider-a-review-of-coursera-udacity-and-edx.

Macleod, H., Haywood, J., Woodgate, A., & Alkhatnai, M. (2014). Emerging patterns in MOOCs: Learners, course designs and directions. *TechTrends*, *59*(1), 56–63.

McNulty, J. A., Hoyt, A., Gruener, G., Chandrasekhar, A., Espiritu, B., Price, J., Ron, & Naheedy, R. (2009). An analysis of lecture video utilization in undergraduate medical education: Associations with performance in the courses. BMC Medical Education, 9(1), 6-6. doi:10.1186/1472-6920-9-6.

Mehaffy, G.L. (2012). Challenge and change. EDUCAUSE Review, September/October 2012, 25-41. Retrieved from http://net.educause.edu/ir/library/pdf/ERM1252.pdf.

Mendoza-Gonzalez, R. (2016). *User-Centered Design Strategies for Massive Open Online Courses (MOOCs)*. IGI Global.

Meyer, R. (2013, October 21). Harvard and MIT's Online Education Startup Has a New Way to Make Money. Retrieved from http://www.theatlantic.com/technology/archive/2013/10/harvard-and-mits-online-education-startup-has-a-new-way-to-make-money/280700/.

Milligan, C., Margaryan, A., & Littlejohn, A. (2013). Patterns of engagement in massive open online courses. *Journal of Online Learning with Technology*.

Moore, M. G. (1989). Editorial: Three types of interaction. *American Journal of Distance Education*, 3(2), 1–7. doi:10.1080/08923648909526659.

Morse, G. (2015, February 23). The Problem With Big Data. Retrieved April 3, 2016, from https://channels.theinnovationenterprise.com/articles/the-problems-with-big-data

Müller, F. & Stöhr, C. (2016). The role of video as a learning component in two MOOCs in Swedish higher education. *Proceedings of the 2th MOOCs in Scandinavia Conference*.

Ng'ambi, D., & Bozalek, V. (2015). Editorial: Massive open online courses (MOOCs): Disrupting teaching and learning practices in higher education. *British Journal of Educational Technology: Journal of the Council for Educational Technology*, *46*(3), 451–454.

NRC (1999). High Stakes: Testing for Tracking, Promotion, and Graduation. Heubert, Jay P., and Robert M. Hauser, eds. Washington, D.C.: National Academy Press, 1999. Retrieved from http://www.nap.edu/catalog/6336.html.

Norberg, A., Händel, Å., Ödling, P. (2015). Using MOOCs at learning centers in northern sweden. *International Review of Research in Open and Distance Learning*, 16(6), 137-151. doi:10.19173/irrodl.v16i6.2035.

O'Sullivan, D. (2016). MOOCs and the middle ages: England in the time of king richard III. *Historian,* (129), 12.

Pappano, L. (2012, November 2). The year of the MOOC. The New York Times. Retrieved from http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=all&_r=0.

Perna, L. W., Ruby, A., Boruch, R. F., & Wang, N. (2014). Moving through MOOCs understanding the progression of users in Massive Open Online Courses. *Educational Researcher* , 0013189X14562423.

Perry, F. L., Jr., & Nichols, J. D. (2014). Understanding Research in Education: Becoming a Discerning Consumer. Routledge.

Peters, O. (2010). Distance education in transition: Developments and issues. BIS-Verlag der Carl-von-Ossietzky-Univ.

Pethuraja, S. (2015). Massive Open Online Courses (MOOCs) For Everyone.

Porter, S. (2015). To MOOC or Not to MOOC: How Can Online Learning Help to Build the Future of Higher Education? *Elsevier Science*.

Pring, R. (2000). The “false dualism”of educational research. *Journal of Philosophy of Education*, *34*(2), 247–260.

Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2016). Understanding MOOC students: Motivations and behaviours indicative of MOOC completion: MOOC student motivations and behaviors. *Journal of Computer Assisted Learning*, 32(3), 202-217. doi:10.1111/jcal.12131.

Raffaghelli, J. E., Cucchiara, S., & Persico, D. (2015). Methodological approaches in MOOC research: Retracing the myth of proteus: MOOC research: Retracting the myth of proteus. *British Journal of Educational Technolog*y, 46(3), 488-509. doi:10.1111/bjet.12279.

Rai, L., & Chunrao, D. (2016). Influencing factors of success and failure in MOOC and general analysis of learner behavior. *International Journal of Information and Education Technology*, 6(4), 262-268.

Rees, J. (2013, July). The MOOC racket. Retrieved April 18, 2016, from http://www.slate.com/articles/technology/future_tense/2013/07/moocs_could_be_disastrous_for_students_and_professors.html.

Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds), Handbook of research on student engagement (pp. 149–172). New York: Springer.

Reich, J. (2015). Rebooting MOOC Research. *Science*, *347*(6217), 34–35.

Rehfeldt, R. A., Jung, H. L., Aguirre, A., Nichols, J. L., & Root, W. B. (2016). Beginning the dialogue on the e-transformation: Behavior analysis’ first massive open online course (MOOC). *Behavior Analysis in Practice,* 9(1), 3-13. doi:10.1007/s40617-015-0102-z.

Rhode, J. F. (2009). Interaction Equivalency in Self-Paced Online Learning Environments : An Exploration of Learner Preferences. *The International Review of Research in Open and Distance Learning*, 10(1), 1–23. Retrieved from http://www.irrodl.org/index.php/irrodl/article/viewArticle/603.

Rice, J. (2013). What I learned in MOOC. NATL COUNCIL TEACHERS ENGLISH 1111 KENYON RD, URBANA, IL 61801 USA.

Robbins, J. (March 25, 2013). The ethics of MOOCs. Retrieved from http://www.insidehighered.com/blogs/sounding-board/ethics-moocs.

Roberts, G., Waite, M., Lovegrove, E. J., & Mackness, J. (2013). xvc: Hybridity in through and about MOOCs. *Creating a Virtuous Circle: Proceedings of OER13*.

Rodriguez, C. O. (2012). MOOCs and the AI-Stanford Like Courses: Two Successful and Distinct Course Formats for Massive Open Online Courses. *European Journal of Open, Distance and E-Learning*. Retrieved from http://eric.ed.gov/?id=EJ982976.

Romanov, K., & Nevgi, A. (2007). Do medical students watch video clips in eLearning and do these facilitate learning?, *Medical Teacher*, 29(5), 490-494. doi:10.1080/01421590701542119.

Romero, C. & Ventura, S. (2010), "Educational Data Mining: A Review of the State of the Art," in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601-618, doi: 10.1109/TSMCC.2010.2053532.

Roos, D. (2014, October 17). How edX Works. Retrieved from http://people.howstuffworks.com/edx1.htm.

Rosewell, J., & Jansen, D. (2014). The OpenupEd quality label: benchmarks for MOOCs. *INNOQUAL: The International Journal for Innovation and Quality in Learning*, *2*(3), 88–100.

Round, C. (2013). The best mooc provider: A review of coursera, udacity and edx. Retrieved from http://www.skilledup.com/articles/the-best-mooc-provider-a-review-of-coursera-udacity-and-edx.

Rubens, W., Kalz, M., & Koper, R. (2014). Improving The Learning Design of Massive Open Online Courses. *The Turkish Online Journal of Educational Technology (TOJET)*, 13(4), 71–80.

Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D. E. (2014). Who does what in a massive open online course? *Communications of the ACM*, *57*(4), 58–65.

Seaton, D. T., Coleman, C., Daries, J., & Chuang, I. (2015). Enrollment in MITx MOOCs: Are We Educating Educators. *Educause Review (February 2015)*.

Selingo, J. (2014, Oct 29). Demystifying the MOOC. New York Times. Retrieved from http://www.nytimes.com/2014/11/02/education/edlife/demystifying-the-mooc.html.

Selwyn, N., Bulfin, S., & Pangrazio, L. (2015). Massive open online change? Exploring the discursive construction of the "MOOC" in newspapers. *Higher Education Quarterly*, 69(2), 175–192.

Shah, D. (2013, December 23). MOOCs in 2013: Breaking down the numbers. Retrieved from https://www.edsurge.com/news/2013-12-22-moocs-in-2013-breaking-down-the-numbers.

Shah, D. (2014, December 26). MOOCs in 2014: Breaking down the numbers. Retrieved from https://www.edsurge.com/news/2014-12-26-moocs-in-2014-breaking-down-the-numbers.

Shum, S. B., & Ferguson, R. (2012). Social learning analytics. *Journal of Educational* Technology & Society,15(3), 3.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380-1400.

Siemens, G. (2014). Connectivism: A learning theory for the digital age. Retrieved from http://er.dut.ac.za/handle/123456789/69.

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, 46(5), 30.

Singer, S.R. & Bonvillian, W.B. (2013). Two Revolutions in Learning. *Science 22*, 339(6126), 1359.

Snijders, C., Matzat, U., & Reips, U.-D. (2012). Big data: Big gaps of knowledge in the field of internet science. *International Journal of Internet Science*, *7*(1), 1–5.

Steffens, K. (2015). Competences, learning theories and MOOCs: Recent developments in lifelong learning. *European Journal of Education*, 50(1), 41-59. doi:10.1111/ejed.12102.

Stephens, M., & Kyle M L Jones. (2014). MOOCs as LIS professional development platforms: Evaluating and refining SJSU's first not-for-credit MOOC. *Journal of Education for Library and Information Science*, 55(4), 345.

Stöhr, C. (2016). STEM MOOCs in practice – experiences from ChalmersX first MOOC on graphene science and technology. In. Khalil, Ebner, Kopp, Kalz (eds.). *Proceedings of the European Stakeholder Summit on experiences and best practices in and around MOOCs (EMOOCS 2016)*. 527-531.

Stöhr, C., Stathakarou, N., Müller, F., Nifakos, S., & Mcgrath, C. (2016). Videos as learning objects in MOOCs - A study of specialist and non-specialist users' engagement in MOOCs". Manuscript submitted for publication.

Sursock, A. (2015). Trends 2015: Learning and teaching in European universities. European University Association.

Suthers, D., & Rosen, D. (2011). A Unified Framework for Multi-level Analysis of Distributed Learning. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 64–74). New York, NY: ACM.

Suthers, D., & Verbert, K. (2013). Learning analytics as a middle space. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 1–4). ACM.

Terras, M. M., & Ramsay, J. (2015). Massive open online courses (MOOCs): Insights and challenges from a psychological perspective. *British Journal of Educational Technology: Journal of the Council for Educational Technology*, *46*(3), 472–487.

The Four V's of Big Data. (2013). The Four V's of Big Data. Retrieved from http://www.ibmbigdatahub.com/infographic/four-vs-big-data.

Tschofen, C., & Mackness, J. (2012). Connectivism and dimensions of individual experience. *The International Review of Research in Open and Distributed Learning*, *13*(1), 124–143.

UKÄ. (2015). Higher education in Sweden 2015 Status Report. Stockholm, Sweden: Swedish Higher Education Authority.

UKÄ (2016a). Öppna nätbaserade kurser (MOOCs) i svensk högskola. Stockholm, Sweden: Swedish Higher Education Authority.

UKÄ (2016b). MOOCs i svensk högskola - möjligheter och utmaningar. Seminarium Berns 27 januari 2016. [Video File]. Retrieved from http://presenter.qbrick.com/?presentationGuid=117d8ebb-49ec-44cc-b8cc-702b641993b2

Universities UK. (2013). *Massive open online courses: Higher education's digital moment?*

Vardi, M. Y. (2012). Will MOOCs destroy academia? *Communications of the ACM*, *55*(11), 5.

Veletsianos, G., Collier, A., & Schneider, E. (2015). Digging deeper into learners' experiences in MOOCs: Participation in social networks outside of MOOCs, notetaking and contexts surrounding content consumption. *British Journal of Educational Technology*, 46(3), 570-587, doi:10.1111/bjet.12297.

Veletsianos, G., & Shepherdson, P. (2016). A Systematic Analysis and Synthesis of the Empirical MOOC Literature Published in 2013–2015. *The International Review of Research in Open and Distributed Learning*, 17(2). http://doi.org/10.19173/irrodl.v17i2.2448.

Voss, B. D. (2013). Massive open online courses (MOOCs): A primer for university and college board members. AGB Association of Governing Boards of Universities and Colleges. Retrieved from http://www.valpo.edu/vital/assets/docs/AGB-report_2013_MOOCs.pdf.

Wagner, E. D. (1994). In support of a functional definition of interaction. *The American Journal of Distance Education*, *8*(2), 6–29. doi:10.1080/08923649409526852.

Wiebe, E., Thompson, I., & Behrend, T. (2015). MOOCs From the Viewpoint of the Learner: A Response to Perna et al. (2014). *Educational Researcher*, *44*(4), 252–254.

Williams, R., Karousou, R., & Mackness, J. (2011). Emergent learning and learning ecologies in web 2.0. *International Review of Research in Open and Distance Learning*, 12(3).

Woodgate, A., Macleod, H., Scott, A., & Haywood, J. (2015). Differences in online study behaviour between sub-populations of MOOC learners. *Educación XX1*, 18(2), 147-163.

Woolf, M. (2014, July 10). Who Performs the Best in Online Classes?. Retrieved from http://minimaxir.com/2014/07/online-class-charts/.

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 Data-driven education workshop* (Vol. 11, p. 14).

Young, J. R. (2013). What professors can learn from "hard core" MOOC students. *The Chronicle of Higher Education*, *59*(37), A4.

Yuan, L., & Powell, S. (2013). MOOCs and disruptive innovation: Implications for higher education. *Special Edition*, 60.

Yuan, L., Powell, S. & Olivier, B. (2014). Beyond MOOCs: Sustainable Online Learning in Institutions. CETIS, Bolton. Retrieved from http://publications.cetis.ac.uk/2014/898.

Zemsky, R. (2014). With a MOOC MOOC Here and a MOOC MOOC There, Here a MOOC, There a MOOC, Everywhere a MOOC MOOC. *The Journal of General Education*, *63*(4), 237–243.

Zimmer, M. (2010). "But the data is already public": on the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313–325.

Zhong, S.-H., Zhang, Q.-B., Li, Z.-P., & Liu, Y. (2016). Motivations and Challenges in MOOCs with Eastern Insights. *International Journal of Information and Education Technology*, 6(12), 954.