DISSERTATIONES
INFORMATICAE
UNIVERSITATIS
TARTUENSIS
39

LIDIA FEKLISTOVA

Learners of an Introductory Programming MOOC: Background Variables, Engagement Patterns and Performance

DISSERTATIONES INFORMATICAE UNIVERSITATIS TARTUENSIS
39

DISSERTATIONES INFORMATICAE UNIVERSITATIS TARTUENSIS
39

LIDIA FEKLISTOVA

Learners of an Introductory Programming MOOC: Background Variables, Engagement Patterns and Performance

UNIVERSITY OF TARTU
Press
1632

Institute of Computer Science, Faculty of Science and Technology, University of Tartu, Estonia.

Dissertation has been accepted for the commencement of the degree of Doctor of Philosophy (PhD) in Computer Science on November 9, 2022 by the Council of the Institute of Computer Science, University of Tartu.

Supervisors

Assoc. Prof.	Marina Lepp University of Tartu Tartu, Estonia
Assoc. Prof.	Piret Luik University of Tartu Tartu, Estonia

Opponents

Prof.	Valentina Dagienė Vilnius University Vilnius, Lithuania
Prof.	Tobias Ley Tallinn University Tallinn, Estonia University for Continuing Education Krems Krems, Austria

The public defense will take place on December 13, 2022 at 12:15 in the University of Tartu Delta Centre, room 1037, Narva mnt 18, Tartu, Estonia.

The publication of this dissertation was financed by the Institute of Compute Science, University of Tartu.

ISSN 2613-5906 (print)
ISBN 978-9916-27-098-1 (print)

ISSN 2806-2345 (PDF)
ISBN 978-9916-27-099-8 (PDF)

University of Tartu Press
http://www.tyk.ee/

Viam supervadet vadens

Lucius Annaeus Seneca the Younger

ABSTRACT

In the modern world, people constantly improve their knowledge and skills to be successful. One opportunity to facilitate personal and professional development is to participate in massive open online courses (MOOCs). This course format has several advantages such as online participation or flexibility in time. However, given the huge number of participants and diversity of their backgrounds, it is a serious challenge for MOOCs instructors to engage them all in the learning process. The impact of background variables on engagement, which in turn may influence performance, remains understudied. Thus, research is required to understand the phenomenon of MOOCs in which participants, besides coming from different backgrounds, have greater autonomy compared to traditional physical classes, and to provide in future more cost-effective MOOCs and personalised learning experiences.

The doctoral thesis investigates MOOC participants' background variables and their impact on course enrolment and completion probability, and explores different engagement clusters among completers in terms of background variables and performance. Data from various sources were extracted to compose an overall picture.

The thesis is focused on a MOOC "About Programming" provided by the research group of informatics didactics (University of Tartu). The study begins with an investigation of the MOOC participants' and completers' background variables. Distinctive variables are identified in comparison to the participants and completers in other two programming MOOCs provided by the same research group. Next, the performance by the MOOC "About Programming" completers and non-completers is compared. The results indicate that completers do not always achieve better results in assessments than non-completers.

Thereafter, behavioural and cognitive engagement solely among the MOOC "About Programming" completers is studied. The results indicate that completers cannot be considered a homogeneous group. There are different engagement clusters that can be identified among them. In addition, members' background variables and performance vary between the identified clusters.

The results of the thesis can prove quite beneficial to the scientific literature to understand the phenomenon of MOOC. This comprehension in terms of a variety of background variables, engagement patterns and performance can be helpful for course instructors to develop cost-effective MOOCs and provide personalised learning where different course activities and help sources can be targeted at specific groups.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AP	MOOC “About Programming”
IT	Information technology
IP1	MOOC “Introduction to Programming I”
IP2	MOOC “Introduction to Programming II”
KC	Knowledge collector
MOOC	Massive open online course
STEM	Science, technology, engineering and mathematics

LIST OF ORIGINAL PUBLICATIONS

Publications included in the thesis

I. Luik, P., **Feklistova, L.**, Lepp, M., Tõnisson, E., Suviste, R., Gaiduk, M., Säde, M., & Palts, T. (2019). Participants and completers in programming MOOCs. *Education and Information Technologies*, *24*(6), 3689–3706. https://doi.org/10.1007/s10639-019-09954-8

II. **Feklistova, L.**, Lepp, M., & Luik, P. (2019). Completers' engagement clusters in programming MOOC: The case of Estonia. In L. Gómez Chova, A. López Martínez, I. Candel Torres (Eds.), *Proceedings of the 12th annual international conference of education, research and innovation (ICERI)* (pp. 1119–1126). IATED. http://doi.org/10.21125/iceri.2019.0341

III. **Feklistova, L.**, Luik, P., & Lepp, M. (2020). Clusters of programming exercises difficulties resolvers in a MOOC. In C. Busch, M. Steinicke, T. Wendler (Eds.), *European Conference on e-Learning* (pp. 563–569). Academic Conferences International Limited.

IV. **Feklistova, L.**, Lepp, M., & Luik, P. (2021). Learners' performance in a MOOC on programming. *Education Sciences*, *11*(9). https://doi.org/10.3390/educsci11090521

The author's contributions to the publications were as follows:

Publication I:	literature review, results interpretation and discussion, writing the paper in cooperation with other authors.
Publications II–III:	formulating research questions, developing a post-questionnaire, carrying out data collection and analysis, literature review, writing the paper as the main author.
Publication IV:	formulating the research questions, data analysis, literature review, writing the paper as the main author.

Publications not included in the thesis

Luik, P., Lepp, M., **Feklistova, L.**, Säde, M., Rõõm, M., Palts, T., Suviste, R., Tõnisson, E. (2020). Programming MOOCs – different learners and different motivation. *International Journal of Lifelong Education*, *39*(3), 305−318. http://doi.org/10.1080/02601370.2020.1780329

1. INTRODUCTION

This chapter provides an overview of the research problem that encouraged the author to investigate MOOC participants' background variables, engagement patterns and performance. At the end of this chapter four research questions that guided this study are posed.

1.1. Research problem

Acquisition of new knowledge and skills does not need to occur exclusively between the walls of a school, a university, or a similar institution. Personal or professional development can be supported by participating in training programs which can be provided in different formats. Massive open online course (hereinafter, MOOC) is one such format that encourages ongoing professional development and lifelong learning (Donitsa-Schmidt & Topaz, 2018).

Researchers (e.g., Dodson et al., 2015; Phan et al., 2016) have explained the abbreviation MOOC as follows. The *massiveness* of MOOC refers to the fact that, in general, the number of participants is not limited, and the generated amount of quantitative data about participants' activity and performance is huge. The *openness* refers to the availability of a course for everyone and everywhere since a course is usually free of charge and has no admission requirements. The *online* component refers to the fact that the learning process no longer takes place exclusively in a physical classroom. Due to using different gadgets, people from all parts of the world can participate in a course virtually, and engage with learning activities provided by course instructors. Given all these aspects, MOOCs expand educational opportunities to a large number of people. According to Shah's report (2021), by the end of 2021, about 220 million learners had acquired new knowledge and skills in MOOCs, nearly 1,000 universities offered about 70 MOOC-based degrees and more than 19,000 courses.

MOOCs are one of the ways to fill the gap in the training of new specialists. The Future of Jobs Report 2020 indicated the continued growing demand for experts in existing and newly emerging specialisations in the field of computer science (World Economic Forum, 2020). Different MOOC platforms offer a huge number of courses on computer programming which is the core of computer science. For example, a query made in February 2022, using the term "computer programming" and limiting results only to computer science-related fields, retrieved 1,762 courses on the Coursera platform. For the last few years, computer programming courses are among the top-ranking courses taught through MOOCs. On the list of "250 most popular online courses of all time", the top-10 includes four courses related to computer science and programming (Shah, 2022).

Given the high demand for specialists in computer science and for courses on programming, it is crucial to create MOOCs that facilitate application of the obtained knowledge and skills in the real world after completion. This requires course instructors not only to provide up-to-date learning materials but also consistently browse research results about MOOC participants, and the methodological and technical issues of such courses. Although ten years have passed since the New York Times heralded 2012 as "*The year of the MOOC*" (Pappano, 2012), researchers are still analysing participants' socio-demographic characteristics, engagement, performance, motivation, etc. and are suggesting different improvements in course design (e.g., Alonso-Mencía et al., 2021; Chen et al., 2020; de Jong et al., 2021; de Souza & Perry, 2021; Deng et al., 2019; Psathas et al., 2018; Zafras et al., 2020).

To help course instructors and researchers better understand findings, it is important to use terms on a similar basis. With regard to those who participate in MOOCs some researchers have used the general term "participant" (Kahan et al., 2017) or "learner" (Kang, 2020), while others have distinguished between *an enrolled student*, *a non-completing student* and *a completing student* (Cristea et al., 2018). In this doctoral thesis a MOOC population is classified as follows (Figure 1):

- non-starter – a person who registered on the course but has not made any submissions for assessment;
- non-completer – a person who failed to meet the course completion requirements;
- completer – a person who fulfilled all requirements for successful course completion;
- participant – the term encompasses all three abovementioned groups;
- learner – the term encompasses both non-completers and completers.

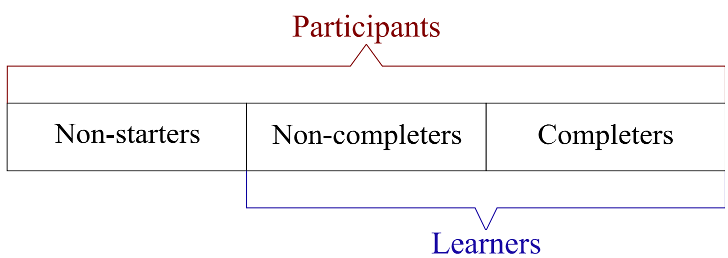

Figure 1. MOOC population classification used in the current thesis

Given the *massiveness* and *openness* of MOOCs, it is natural that participants have different background variables. There is still a lack of research about MOOC participants' background variables and how background variables, such as gender, education, etc., influence the possibility to participate and complete a MOOC (Zafras et al., 2020). In addition, it is claimed that background may change over time and, therefore, new studies are needed (Duran et al., 2020). Thus, participants' and completers' background variables in the context of a MOOC on programming is an intriguing area to explore.

Learners' course completion status may be determined from their performance which is usually calculated based on assessment results. The types of assessment in MOOCs can include creation of digital stories, writing an essay, working on a mini project, submitting a quiz, etc. (de Barba et al., 2016; Phan et al., 2016; Psathas et al., 2018). However, some researchers observe performance through the lens of engagement and desire to continue the course (Cobos & Garcia, 2020), or grades and desire to continue the course (de Souza & Perry, 2021). In this doctoral thesis learners' performance in the MOOC is interpreted as the number of attempts per programming task and per quiz, and the received quiz scores.

De Barba et al. (2016) call to investigate the factors that contribute to learners' performance in MOOCs. One of such factors may be engagement since learners who were more engaged were more likely to succeed in the course (Goldberg et al., 2015). Engagement encompasses mental energy and effort that MOOC learners invest into the learning process to achieve planned results (Jung & Lee, 2018), and refers to learners' readiness and degree of activity in interaction with MOOC content activities (Walji et al., 2016). Learners demonstrate different engagement patterns (Deng et al., 2020b; Khalil & Ebner, 2017). The patterns can be defined as different approaches adopted by learners in learning environments (Herrington et al., 2003) or ways in which learners' activity can be manifested (Lawson & Lawson, 2013). Researchers have described learners' engagement either as a unidimensional construct that is seen mostly as behavioural engagement (Kuo et al., 2021; Sun et al., 2019) or as a multidimensional construct that encompasses dimensions like behavioural, emotional, cognitive and social engagement (Deng et al., 2019; Lan & Hew, 2020). Researchers have also pointed out the importance of learners' background variables in MOOC engagement (Deng et al., 2020b; Kuo et al., 2021; Li & Baker, 2018). Despite the large amount of studies about engagement it is still unclear how MOOC learners engage with a course (Deng et al., 2019; Kuo et al., 2021; Lan & Hew, 2020). This understanding can help gain insights why there are completers and non-completers in a MOOC (Pursel et al., 2016; Sun et al., 2019), and what course activities and support mechanisms are most in demand. Thus, the current doctoral thesis contributes to the comprehension of engagement with MOOCs through addressing engagement from a behavioural and cognitive point of view.

Knowledge of behavioural engagement is important since it is associated with better performance (Deng et al., 2019). Researchers have employed different data collection methods to examine learners' behavioural engagement with multiple activities such as reading materials, watching videos, taking quizzes, submitting assessments and participating in forum discussions (e.g., Barthakur et al., 2021; Kahan et al., 2017; Kang, 2020; Rizvi et al., 2020) or with one particular activity such as watching a lecture video (Li & Baker, 2018). With regard to programming courses, researchers have studied the amount of compilations, executions and code generated (Gallego-Romero et al., 2020). In this doctoral thesis the behavioural engagement is considered from a MOOC completer's perspective and is interpreted as the amount of activities a completer engaged with throughout the course.

Based on behavioural engagement, researchers have identified various groups of learners in MOOCs (e.g., Arora et al., 2017; Khalil & Ebner, 2017; Kizilcec et al., 2013). Besides other groups, so-called group of *perfect* learners (Khalil & Ebner, 2017) or *achievers* (Arora et al., 2017) has been identified, however the completion ratio of this group was not 100%. The author of the thesis has not found in scientific databases any studies related to behavioural engagement of solely MOOC completers.

Although previous research (e.g., Kahan et al., 2017; Li & Baker, 2018; Gallego-Romero et al., 2020) has undoubtedly enriched our knowledge of behavioural engagement in MOOCs, there are still calls by researchers to examine the engagement with different activities that could influence learners' performance (Mubarak et al., 2021) and to explore behaviour of learners with different background variables in the same MOOC (Deng et al., 2019; Phan et al., 2016). Inspired by these calls and keeping in mind the diversity of MOOC participants, the author of the doctoral thesis examines the possibility of identifying different groups of completers based on behavioural engagement in programming MOOCs, and the variability of completers' background variables and performance across the identified groups.

Cognitive engagement is also vital in online learning (Kuo et al., 2021) since it is critical for knowledge acquisition (Li & Baker, 2018) and achievement of desired outcomes (Li & Baker, 2016). While behavioural engagement can be observed, cognitive engagement involves invisible processes (Li & Baker, 2018). Researchers are still looking for appropriate ways to measure cognitive engagement. Some researchers have utilised scales which indicated MOOC learners' reported level of agreement with statements such as "*If I do not know about a concept when I am learning in the online class, I do something to figure it out*" (Kuo et al., 2021) or "*If I watched a video lecture that I did not understand at first, I would watch it again to make sure I understood the content*" (Deng et al., 2020a). Others have examined quiz reattempts (Do et al., 2013, as cited in Li & Baker, 2018) or content generated by learners (Galikyan et al., 2021). In this thesis cognitive engagement is considered from a MOOC completer's perspective and is interpreted as the frequency with which a completer engages with help sources to resolve a difficulty while solving a programming task.

Despite the need to be aware of different learner groups in order to optimise the learning environment (Hennis et al., 2016), there is a limited number of studies (Li & Baker, 2016) about identifying different learner groups in MOOCs solely by cognitive engagement. Also, the author of the thesis has not found any scientific research carried out with MOOC completers investigating their cognitive engagement.

Although previous research (e.g., Galikyan et al., 2021; Kuo et al., 2021; Li & Baker, 2016) has shed some light on cognitive engagement in MOOCs, this dimension of engagement remains rarely researched (Deng et al., 2019; Kuo et al., 2021). Cognitive engagement requires research (Atapattu et al., 2019; Deng et al., 2019), including examining the relationship between cognitive engagement and desired learning outcomes (Deng et al., 2019). This thesis contributes to the topic by investigating the different groups of completers that can be identified on the basis of cognitive engagement in programming MOOC, and the variability of completers' background variables and performance across the identified groups.

The doctoral thesis examines MOOC participants' background variables, performance, and completers' behavioural and cognitive engagement. The novelty of the study is the identification of behavioural and cognitive engagement clusters solely among MOOC completers, analysing data that include the engagement not only with activities required for course completion, but also with voluntary ones. In addition, the originality of this thesis lies in examining these clusters in terms of background variables and performance using the data from the same MOOC. This can be helpful for understanding how completers have actually learnt in MOOCs and can provide a deep insight of how heterogeneous the group of completers is, how different completers' sub-groups engage with a course and to what extent engagement patterns can cause differences in completers' performance. Understanding of these sub-groups and their engagement patterns can be useful for planning future MOOCs in a more cost-effective way, and improving the quality of acquired knowledge through a more personalised learning experience.

1.2. Focus of the research

The doctoral thesis aimed to explore participants' background variables and their impact on course enrolment and completion probability, identify completers' behavioural and cognitive engagement clusters and investigate them in terms of background variables and performance. The main focus was on an introductory programming MOOC named "About Programming". "Introduction to Programming I" and "Introduction to Programming II" were the other two programming MOOCs that were studied as well. The aim was approached by answering the following research questions:

1. What background variables of participants in introductory programming MOOC can be identified compared to other programming MOOCs?
   The research question 1 is addressed in Publication I.

2. How did completers perform compared to the non-completers of introductory programming MOOC?
   The research question 2 is addressed in Publication IV.

3. What behavioural engagement clusters of completers can be identified in introductory programming MOOC, and how do completers' background variables and performance vary between clusters?
   The research question 3 is addressed in Publications II and IV.

4. What cognitive engagement clusters of completers can be identified in introductory programming MOOC, and how do completers' background variables and performance vary between clusters?
   The research question 4 is addressed in Publications III and IV.

2. THEORETICAL BACKGROUND

The chapter provides an overview of the theoretical background of the doctoral thesis and is organised as follows: in sub-chapter 2.1. an overview about MOOC participants' and completers' background variables is given. Sub-chapter 2.2. considers behavioural and cognitive engagement. Sub-chapter 2.3. highlights the relationships between two engagement dimensions and learners' achievements.

2.1. Diversity of MOOCs participants and completers

Given the huge number of MOOCs participants, wide range of course topics, different levels of difficulty, variety in courses duration and completion requirements, etc. it is natural that backgrounds of those who participate and complete MOOCs vary. Usually, background variables encompass gender, age, education, previous experience with a course subject, etc. Researchers receive such data from learning platforms like Coursera (Pursel et al., 2016) or voluntary questionnaires (Gil-Jaurena et al., 2017; Kang, 2020), although high-achievers are more likely to respond to a questionnaire (Kovanović et al., 2019; Pursel et al., 2016).

Participants' difference in one of the main background variables, gender, may relate to the MOOC field. While females are major participants in social science and humanities fields (Gil-Jaurena et al., 2017), males are the majority in science, technology, engineering, and mathematics (hereinafter, STEM) courses (de Souza & Perry, 2021; Hennis et al., 2016; Jiang et al., 2018) and computer science (e.g., Alonso-Mencía et al., 2021; Duran et al., 2020; Kizilcec et al., 2013; Psathas et al., 2018). But there are some areas in technological MOOCs, such as sentiment analysis or usability evaluation, where the number of enrolments by females is similar to males (de Souza & Perry, 2021). Over recent years there has been a slight increase in the proportion of females participating in introductory programming MOOCs which are the core of computer science (Duran et al., 2020).

Findings about completion by females and males have differed across studies. While some studies reported the dominance of females among MOOC completers (e.g., Goldberg et al., 2015; Pradubwate et al., 2020), others, based on systematic literature review, concluded that males tended to complete MOOCs far more frequently than females (Zafras et al., 2020). It has been also pointed out that males tend to more frequently complete higher-level MOOCs (Kizilcec et al., 2013). However, in several studies no significant difference has been found, neither in completion by gender (Hone & Said, 2016) nor in gender impact on course completion probability (Bonafini, 2017; Pursel et al., 2016). To some extent, these contradictions may be related to both the course topic of a MOOC (or course content as assumed by Morris et al. (2015)) and the unequal proportions of enrolled females and males. However, considering STEM and computer science courses, where females are less likely to enrol, both males and females are equally likely to complete such courses (de Souza & Perry, 2021; Jiang et al., 2018). It is interesting to note that, while females had a higher likelihood of dropout at the beginning and approximately until the midpoint of the programming course, the gender difference is not noticeable towards the end of the course (Duran et al., 2020; Chen et al., 2020).

Participants at various ages take part in MOOCs. It is quite problematic to compare participants' age since some researchers operate with the term "average age" (Pursel et al., 2016), while others present results in terms of "median age" (Morris et al., 2015). In addition, since not all participants respond to questionnaires, the age of those who actually participate in the course and those who fill out questionnaires may be slightly different (Gil-Jaurena et al., 2017; Pursel et al., 2016). The diversity in the ages of the participants may be related to the course topic (Morris et al., 2015) or the group that a MOOC is primarily targeting. While the same MOOCs are offered for the university students and for the general public, the university students are, naturally, younger than general participants (Duran, et al., 2020; Khalil & Ebner, 2017; Watted & Barak, 2018). Defining a dominant age group is also complicated since studies have defined different boundaries for age groups (e.g., de Jong et al., 2021; Greene et al., 2015; Gil-Jaurena et al., 2017) or might even exclude certain age groups for some reasons (Hennis et al., 2016). Despite these issues, some researchers (Cabedo et al., 2018; Gil-Jaurena et al., 2017) have reported the dominance of participants in their 30s and 40s.

Regarding completion, no significant effect of age has been detected (Hone & Said, 2016). Meanwhile, older MOOC participants are less likely to drop out (Greene et al., 2015; Morris et al., 2015; Semenova & Rudakova, 2016). It might be due to a "sense of responsibility" – older participants may be more careful when selecting a course and work harder since the new skills can be useful in work duties (Semenova & Rudakova, 2016). However, Guajardo Leal et al. (2019) noticed that for each year of increase in the learner's age, the chance of successful course completion is decreasing.

The majority of MOOC participants are employed people (Cabedo et al., 2018; Greene et al., 2015; Kahan et al., 2017; Morris et al., 2015). This is consistent with the abovementioned findings about dominance of 30-40-year-old participants since people from this age group are a major labour force. The proportions of groups of unemployed people and students vary (Cabedo et al., 2018; Morris et al., 2015). It should be noted that comparison of groups with different employment status can be can be challenging due to the fact that a course may be intended for the general public or for one specific group (e.g., university students, people from a specific field of activity, etc.) and different groups can have different goals for participating in MOOCs. For example, employed learners treat the course as a training resource (Cabedo et al., 2018; Gomez-Zermeno & Garza, 2016), an opportunity for knowledge refreshment, professional advancement or improving future employability in the case of changing career fields (Watted & Barak, 2018). But university students, apart from personal benefits such as satisfaction of curiosity, deepening interest in a subject, personal enrichment and challenge (Watted & Barak, 2018), seek for academic benefit since the participation in MOOC expand their regular curriculum (Barak et al., 2016; de Jong et al., 2021).

Findings about completion by groups of different employment status have varied across studies. In the study by Morris et al. (2015) unemployed people were more likely to complete more of their course since, as assumed by Morris and colleagues, they probably had more free time. While in a study by Pradubwate et al. (2020) the majority of completers were university students, Gomez-Zermeno and Garza (2016) found employed participants more likely to succeed in a MOOC.

Given the absence of requirements, in general, for previous education certificates and the massive nature of MOOCs, the diversity of participants' education level is undoubted. However, many studies (e.g., Cabedo et al., 2018; Hennis et al., 2016; Morris et al., 2015; Psathas et al., 2018) have indicated that the majority of MOOC participants have already obtained at least a bachelor's degree.

With regard to relation between education level and MOOC completion, findings are contradictory. Previous studies have not found any significant difference in MOOC completion rates between those who had a university degree and those without it (Goldberg et al., 2015; Hone & Said, 2016). Also, education level is not a significant predictor of course completion (Bonafini, 2017). However, several studies have shown that participants with higher prior education level were more likely to complete a course (Morris et al., 2015; Pursel et al., 2016; Semenova, 2020).

Previous experience with the course subject is also one of background variables used to characterise MOOC participants. The number of participants who are somehow familiar with the topic vary (Hennis et al., 2016; Psathas et al., 2018). With regard to gender, males have more previous experience in the subject (Duran et al., 2020; Hennis et al., 2016).

While some researchers (Bonafini, 2017) have found previous experience being not related to course completion, others (Kennedy et al., 2015) found it to be a key indicator of learners' success in a course, increasing their chances of completing a MOOC (Semenova, 2020). This latter may be one of explanations why in the abovementioned study by Duran et al. (2020) females who had about two times less previous experience in programming than males were more likely to drop out at the early stage.

Although MOOCs have been actively used over the past years, learners' previous experience with participating in MOOCs vary (Gil-Jaurena et al., 2017; Greene et al., 2015; Hennis et al., 2016; Kang, 2020). Participants with previous experience with online learning and with MOOCs in particular are less likely to drop out of a course (Gomez-Zermeno & Garza, 2016; Greene et al., 2015; Morris et al., 2015). It is quite logical, since such people may have already obtained successful strategies that lead them to course completion (Kang, 2020). However, some studies showed that previous experience with online learning had no impact on completion probability (Pursel et al., 2016; Semenova, 2020).

Participation in MOOCs usually involves completing assessments. The number of exercise resubmissions and, therefore, exercise points have been used as a basis to distinguish groups of learners in university courses (Karavirta et al., 2006). With regard to completion, the more attempts MOOC learners make on a problem set, the lower is the likelihood of dropout (Chen et al., 2020). In addition, completers submit more assessments than non-completers but these differences vary across online course subjects (Soffer & Cohen, 2019).

In sum, the abovementioned studies and those provided in Publication I indicate that MOOCs with different topics, duration, requirements, etc. attract participants with different background variables. Contradictory results about the impact of particular background variables on the probability of MOOC completion have also been observed. While some researchers noted no relationship between particular background variable and completion, others demonstrated an increasing likelihood of successful completion. The contradictory results in the abovementioned studies may be related to diversity of course designs, learning approaches or organisational issues. Nevertheless, this once again highlights the *massiveness* and *openness* of MOOCs and the importance of exploring diverse backgrounds of MOOC participants and completers.

2.2. Engagement in MOOCs

Engagement is one of the vital components of learning, including in MOOCs. It has been mostly studied in the context of traditional classroom courses (Fredricks et al., 2004) and online courses (Cole et al., 2021; Pursel et al., 2016). Engagement by MOOC participants differs from engagement in traditional classes (Kennedy et al., 2015; Ramesh et al., 2013) and online learning engagement (Kuo et al., 2021; Pursel et al., 2016; Sun et al., 2019) because participants' engagement in MOOCs is considered to be less complex but limited by time, space and educational technologies (Deng et al., 2020a). In addition, given the *massiveness* of MOOCs and diversity of participants' backgrounds, making sure that they are all engaged in learning is a challenge (Hew, 2014; Walji et al., 2016). Nevertheless, in the context of MOOCs, engagement has been positively associated with better performance (Deng et al., 2019) and course completion likelihood (Goldberg et al., 2015). It should be noted that further in this thesis engagement is considered in the context of MOOCs.

Engagement is a complex and abstract construct that has many different interpretations. In the context of MOOCs, it has been studied, for example, through the lens of mental energy and effort (Jung & Lee, 2018) or willingness (Walji et al., 2016) that MOOCs learners demonstrated in activities to achieve learning outcomes. Differences in approaches may be related to the multidimensionality of engagement. While many studies have distinguished behavioural, emotional and cognitive dimensions of engagement (e.g., Jung & Lee, 2018; Lan & Hew, 2020), others have also identified a fourth dimension – social engagement (Daniels et al., 2016; Deng et al., 2020a). Despite this multidimensionality, studies have mostly considered only the behavioural dimension (Kuo et al., 2021; Sun et al., 2019), but focusing solely on it can lead to incomplete and incorrect understanding of how exactly MOOC participants engage with a course (Li & Baker, 2016). Thus, investigation of MOOCs engagement from different perspectives allows capturing the complex construct of engagement (Deng et al., 2019) and gaining a deeper understanding of the actual involvement of participants in a course. The current thesis contributes to this comprehension through investigating engagement from two perspectives – behavioural and cognitive. The thesis is based on the engagement theory of learning, according to which learners are "*meaningfully engaged in learning activities through interaction with others and worthwhile tasks*" (Kearsley & Shneiderman, 1998).

2.2.1. Behavioural engagement in MOOCs

Behavioural engagement in MOOCs is one of the most frequently covered dimensions of engagement (Deng et al., 2019; Jung & Lee, 2018; Kuo et al., 2021) and has been considered essential for achieving positive outcomes (Li & Baker, 2018). MOOC researchers (e.g., Jung & Lee, 2018; Poellhuber et al., 2019) use the concept, proposed by Fredricks et al. (2004) for traditional classrooms, that behavioural engagement is related to learners' participation in a course and involvement in academic activities.

To understand behavioural engagement in MOOCs, researchers have used questionnaires (Phan et al., 2016), data logs and learning analytics (Barthakur et al., 2021; Khalil & Ebner, 2017; Li & Baker, 2018; Tseng et al., 2016), combined logs with questionnaires (Kang, 2020; Samoilova et al., 2018), etc. This variety of employed methods can be an indication of the complexity of measuring and understanding learners' engagement with a MOOC. Researchers continue to look for ways of measuring behavioural engagement since there is no ideal method and each of them should be considered with a touch of criticism. Questionnaires can be used to collect learners' opinions and perception regarding a particular point but they are usually optional. Data collected in such a way can be biased due to the high number of non-respondents (Alonso-Mencía et al., 2021; Samoilova et al., 2018) since learners with higher levels of participation and performance (de Barba et al., 2016) as well as more active and engaged MOOC learners are more likely to complete the questionnaires (Kovanović et al., 2019; Samoilova et al., 2018; Torres & Beier, 2018). In addition, questionnaires may be conducted at the end of the course, their wording can be too difficult to understand (Henrie et al., 2015) and learners may over- or underestimate their level of behavioural engagement. Although learning analytics can be helpful in resolving some of the aforementioned problems (Greene et al., 2015) to enhance engagement and provide recommendations (Khalil et al., 2016), it still has some issues. Activities that are performed outside the MOOC platforms, such as watching downloaded videos (Kahan et al., 2017), reading printed-out materials, opening materials through bookmarks, etc., are not tracked. Problems can also arise in relation to security, privacy, transparency, mistaken interpretation (Khalil et al., 2016) and bias produced by unusual behaviour of learners who abuse the *openness* of MOOCs (Alexandron et al., 2019). Although data logs enable quick fixing and measuring of learning activities of a huge number of MOOC learners (Li & Baker, 2018), providing deep insights into learners' interaction with various activities (Rizvi et al., 2020) and as a result, permitting dynamic intervention to improve learners' success (Wen & Rosé, 2014a), this method has some problems as well. Keeping in mind the huge volume of generated data, it is quite difficult to understand what parts of these data are meaningful and valuable for understanding learning processes (Rizvi et al., 2020). Also, it remains uncertain whether an activity was taken as planned or whether a learner interacted (i.e., clicked, opened) with it accidentally. A learner may use some activities (e.g., textual materials) more than once only because the content can be hard or too long to comprehend in one go. Nevertheless, utilising different approaches, researchers have measured behavioural engagement in MOOCs from the point of view of quantified engagement with course pages, videos, forum discussions, quizzes, assessments, etc. (e.g., Barthakur et al., 2021; Goldberg et al., 2015; Li & Baker, 2018).

The diversity of MOOC learners who interact with various course activities differently leads to the emergence of groups with different behaviour engagement (Table 1). For a more thorough overview about the MOOC activities that were used to identify groups, see Publication II. To identify different groups based on learners' behavioural engagement with activities, most studies have employed cluster analysis (e.g., Kizilcec et al., 2013; Khalil & Ebner, 2017; Poellhuber et al., 2019). However, some researchers have utilised latent class analysis (Barthakur et al., 2021; Kang, 2020) as well.

Although each identified group demonstrates a unique mode of behavioural engagement, in all of the studies provided in Table 1 there exists quite a big group of MOOC participants who can be called *disengagers*. These are participants who merely enrolled on a course but have never essentially engaged with activities. The same generalised findings were made by Barthakur et al. (2021) and Li and Baker (2018) who have overviewed numerous studies, including some presented in Publication II and Table 1. Given that enrolment on MOOCs is usually free of charge and entails no specific commitments, the existence of such a group is consistent.

The second group, which is usually quite a small one, consists of learners who can be described as *ideal* learners. They engage with all course components – watch videos, submit assessments, participate in forum discussions, use additional materials provided by instructors, etc. Their active engagement confirms their intention to get the most out of the course (Pursel et al., 2016). As a result, most of them fulfil course requirements and pass the MOOC. This group has also been highlighted by Barthakur et al. (2021) and Li and Baker (2018).

Table 1. MOOC participants' behavioural engagement typologies

Authors	Data analysis	Studied activities	Identified groups
Kizilcec et al., 2013	cluster analysis (*k*-means)	Video lectures, assessments	Auditing (6%–9%)*, completing (5%–27%), disengaging (6%–28%), sampling (39%–80%)
Li and Baker, 2016	cluster analysis (*k*-means)	Video lectures	Disengagers (84%), quick decreasers (8%), slow decreasers (7%), strong enders (1%)
Arora et al., 2017	cluster analysis (*k*-means)	Video lectures, course material, discussion forums, assessments	Uninterested (72%), casuals (16%), explorers (5%), performers (4%), achievers (3%)
Kahan et al., 2017	hierarchical cluster analysis	Video lectures, discussion forums, assessments	Tasters (65%), disengagers (12%), downloaders (9%), online engagers (7%), moderately social engagers (4%), offline engagers (4%)
Khalil and Ebner, 2017	cluster analysis (*k*-means)	Video lectures, discussion forums, assessments	Dropout (87%), gaming the system (11%) and perfect learners (2%)
Samoilova et al., 2018	cluster analysis (partitioning around medoids)	Video lectures, quizzes	Non-active (52%), browsers (31%), auditors (9%) and diligent (8%) learners
Poellhuber et al., 2019	two-step cluster analysis	Video lectures, required readings, PDF lecture slides, quizzes, discussion forums	Ghost (30%)**, active-independent (21%), browser (16%), serious reader (16%), self-assessor (11%), active-social learners (6%)
Kang, 2020	latent class analysis	Video lectures, quizzes	Enrolling (63%), sampling (21%), completing (7%), auditing (5%), disengaging learners (5%)

Note. * Here and hereinafter, percentage of course participants

** Manually removed from the sample, not associated with the cluster analysis

Further generalisation of MOOC learners (based on Table 1) is not so straightforward since researchers can focus on different aspects of engagement, resulting in a different number of identified groups (e.g., 3 groups in the study by Khalil and Ebner (2017) vs 6 groups in the study by Poellhuber et al. (2019)). Nevertheless, benefitting from previous works, Barthakur et al. (2021) have aggregated the rest of learners under the third group and called them *selective* learners. These learners engaged with specific learning activities, depending on their goals and motivations. Li and Baker (2018), exploring previous works, have attempted a further distinction of groups based on priority in activities – some learners (*auditors*) mainly engaged with lecture videos and submitted a small number of assessments, others (*quiz-takers*), oppositely, submitted assessments and watched a small number of videos. It should be mentioned that despite different behavioural engagement patterns, MOOC completers can be found in every group. However, to the best of the thesis author's knowledge, there has been so far no scientific research aimed to investigate solely MOOC completers to understand whether behavioural engagement within this group varies or not. Therefore, unlike previous studies that mostly considered MOOC learners, the current doctoral thesis focuses solely on completers and attempts to distinguish completers' different groups by their behavioural engagement through exploring the amount of activities completers engaged with in the MOOC.

Despite researchers having emphasised the importance of learners' background variables in MOOC engagement (Deng et al., 2020b; Kuo et al., 2021; Li & Baker, 2018), there are not many studies in scientific databases that explore this relationship in the context of behavioural engagement. With regard to gender, some researchers have noted that it is not associated with engagement patterns (Kizilcec et al., 2013) or level of engagement (Whitmer et al., 2015). Meanwhile, taking into account course fields, it has been found that males engage with more activities than females in MOOCs on humanities and liberal arts, but in STEM MOOCs the difference by gender is smaller (Williams et al., 2018).

While Whitmer et al. (2015) have claimed that learners' age had no relationship to the level of engagement, Williams et al. (2018) have reported a positive influence of age on course engagement. Older learners take part in more MOOCs activities (Torres & Beier, 2018) and tend to engage with more activities (Kahan et al., 2017).

Whereas professional status has not been found to be related to the level of engagement (Whitmer et al., 2015), learners with different employment status can vary in their engagement with activities (Kahan et al., 2017). The percentage of working learners is highest among those who engage with fewer activities and the percentage of non-working learners is highest among those who engage with more course activities and particularly with the discussion forum (Kahan et al., 2017).

Although a correlation between education and the level of engagement has not been detected (Whitmer et al., 2015), some differences have been found in the context of course fields. In STEM MOOCs, more educated learners engage with more activities but in MOOCs on humanities and liberal arts no relationship between a higher level of attained education and engagement is found (although PhD learners are less likely to stay engaged with the course) (Williams et al., 2018).

In sum, the variety of data collection and analysis methods employed highlights the complexity of comprehending MOOC learners' behavioural engagement. Researchers have identified different learners' groups by behavioural engagement and have attempted to find out how learners' background variables relate to behavioural engagement. Since the results are still contradictory, further investigation is needed.

2.2.2. Cognitive engagement in MOOCs

Cognitive engagement in MOOCs is a rarely studied dimension of engagement (Deng et al., 2019; Kuo et al., 2021) but it plays an important role in online learning (Kuo et al., 2021), is critical for knowledge acquisition (Li & Baker, 2018) and achieving desired outcomes (Li & Baker, 2016). Although there is yet no general agreement on what cognitive engagement means in the context of MOOCs, researchers (e.g., Jung & Lee, 2018; Li & Baker, 2016; Liu et al., 2018) have often operated with the definition proposed by Fredricks et al. (2004) for traditional classrooms: "*thoughtfulness and willingness to exert the effort necessary to comprehend complex ideas and master difficult skills*". Deng et al. (2020b) attempted to specify cognitive engagement in the context of MOOCs and extended the abovementioned definition to "*learners' mental investment in the study of a MOOC to comprehend complex ideas, master difficult skills and strengthen learning and performance*". Other MOOC researchers have noted that cognitive engagement can be expressed, for example, in searching for supplemental information on the material (Daniels et al., 2016), reworking any remaining unclear concepts (Deng et al., 2020a) or undertaking an action to cope with a difficulty (Kuo et al., 2021). Some MOOC researchers have also linked cognitive engagement to learners' motivational goals and self-regulated learning skills, although their ideas came from traditional education (Joksimović et al., 2018).

Data collection about cognitive engagement is difficult due to the fact that this dimension of engagement involves invisible processes (Li & Baker, 2018). Keeping in mind a suggestion given in the context of traditional classrooms, that data about cognitive engagement can be collected with self-reporting (Fredricks et al., 2004; Henrie et al., 2015), MOOC researchers have used questionnaires (e.g., Deng et al., 2020b; Jung & Lee, 2018). In addition, there have been attempts to infer cognitive engagement with MOOCs from learners' behaviour (Galikyan et al., 2021; Li & Baker, 2018; Liu et al., 2018). These behavioural indicators of cognitive engagement have been collected through a data log, which recorded the number of video interaction events like pause or seeking (Li & Baker, 2018; Liu et al., 2018), or content generated by learners during learning (Galikyan et al., 2021). The latter implies that researchers have investigated the level of language abstraction (Wen et al., 2014b) in learners' forum posts or learners' posts content contributions (Galikyan et al., 2021). Quiz reattempts can also be indicative of learners' cognitive engagement with the course (Do et al., 2013, as cited in Li & Baker, 2018).

Although researchers are still looking for indicators of cognitive engagement and ways to collect data about it, the current thesis proposes that utilisation of help sources can be considered as a proxy indicator of cognitive engagement in MOOCs. Activities like seeking for help (Nelimarkka & Hellas, 2018) or searching for information (Daniels et al., 2016) may indicate that a MOOC learner struggles with a difficulty in the learning process but, nevertheless, applies effort to overcome it. This suggests that data about cognitive engagement could be collected in terms of engaging with help sources. Although course instructors offer different technical solutions to get help (e.g., Publication III; Duran et al., 2020; Lepp et al., 2018; Nelimarkka & Hellas, 2018), and learners seek help from YouTube videos (Lausa et al., 2021) and friends (Nelimarkka & Hellas, 2018), there are only a few studies in scientific databases about how frequently or actively learners engage with help sources. For example, Carter et al. (2015) observed the amount of help learners received from an instructor (although they did not specify the form of the course), Lepp et al. (2018) looked at the number of MOOC learners who get an answer from a troubleshooter per task. The current thesis attempts to investigate the frequency of engaging with help sources by completers as an indicator of cognitive engagement in MOOCs.

Despite there being a variety of approaches to measuring cognitive engagement, each of them should be treated with caution. Issues with using questionnaires overlap with those discussed in sub-chapter 2.2.1. With regard to data logs, for example, video interaction events in MOOCs, Li and Baker (2018) noted that data about pausing or seeking in videos did not allow them to measure cognitive engagement as it is – these data helped researchers infer probable cognitive engagement. They also added that videos can be paused due to reasons unrelated to learning, such as getting a rest or doing something else. Some researchers have concluded that, at least at the moment, data logs cannot be applied to measure learners' mental effort (Li & Baker, 2018) or learning analytics is not capable of capturing non-behavioural states (Samoilova et al., 2018), even though data logs and learning analytics have positive aspects already discussed in sub-chapter 2.2.1. Nevertheless, researchers make a compromise in their approaches and attempt to investigate learners' cognitive engagement with learning.

While knowledge of different learners' groups helps optimise the learning environment (Hennis et al., 2016), scientific databases only include a limited number of studies in which learners' groups by cognitive engagement have been considered. A few researchers (Deng et al., 2020b; Liu et al., 2018) have identified MOOC learners' groups but they have simultaneously employed different engagement dimensions, including cognitive. Unfortunately, due to this complex approach, it is difficult to comprehend how groups of learners differ precisely in terms of cognitive engagement, without taking into account the influence of other dimensions. Although Kew and Tasir (2021) identified different learners' groups solely by cognitive engagement, they did it in the context of e-learning. Meanwhile, the study by Li and Baker (2016) was focused precisely on identifying MOOC learners' groups based solely on cognitive engagement.

Li and Baker (2016) measured cognitive engagement by the number of pauses made by MOOC learners while watching video lectures. By applying *k*-mean cluster analysis they distinguished four groups. Learners who almost never paused the videos or did not watch any videos over several weeks formed the largest group (89% of MOOC learners). The smallest group (less than 1%) consisted of learners who paused each of the videos more than 10 times during most of the course. The other two groups differed in terms of the distribution of pausing over the course period – some continued to pause the videos at an average rate during most of the course (2%) while for others this number decreased quickly as the course progressed (9%). Kew and Tasir (2021) have focused on cognitive engagement in the context of e-learning and investigated learners' effort and attention in analysing and synthesising forum posts. After applying a content analysis method and measuring the percentage of cognitive contributions (such as asking a question that required an explanation, answering a question without any explanation or giving information with elaboration), the authors identified three groups. While about half of learners (52%) demonstrated low cognitive engagement level in discussion forums (i.e., while writing forum posts they used less cognitive contribution, mental effort and did not provide any explanations in their posts), about one third (39%) demonstrated a high level. The third group was placed between the two. However, neither of the abovementioned studies reported how learners from different groups performed in a course. In addition, the author of the doctoral thesis has not found any scientific research aimed to investigate solely completers' cognitive engagement with a MOOC. The current thesis addresses this research gap and attempts to distinguish completers' cognitive engagement based on the frequency of engagement with different help sources while encountering a difficulty.

Whereas learners' background variables are important in engagement in MOOCs (Deng et al., 2020b; Kuo et al., 2021; Li & Baker, 2018), there were no studies found in scientific databases attempting to investigate a relationship between MOOC learners' background variables and solely cognitive engagement. However, several studies have focused on this topic in the context of traditional classrooms or online learning, but with inconsistent results. While some researchers (Kew & Tasir, 2021; Pietarinen et al., 2014) claimed that females demonstrate higher cognitive engagement than males, others (Teng & Wang, 2021; Watt et al., 2017), on the contrary, have observed cognitive engagement of males being higher. Meanwhile, no difference by gender in cognitive engagement (Wang et al., 2011) and no impact on cognitive engagement by gender (Manwaring et al., 2017) was found. In terms of age, younger school students have been found to have higher cognitive engagement (Watt et al., 2017), whereas no effect of age on cognitive engagement has been found among online learners (Wysocki, 2007). Studying online learners, Wysocki (2007) observed that previous online experience and employment status had no effect on cognitive engagement.

In, cognitive engagement is a complicated construct with ongoing disagreement about its definition. MOOC researchers attempt to comprehend the multi-faceted nature of cognitive engagement by employing different data collection and analysis methods. Despite this, little is still known about the groups that can be identified among MOOC learners based on differences in cognitive engagement, and even more so when it comes to completers. Relationships between learners' background variables and cognitive engagement have been barely examined as of yet.

2.3. Relationships between engagement and learners' achievements

Previous studies have indicated contradictory results about the relationship between behavioural engagement in the context of MOOCs and learners' achievements (in terms of performance or course completion). Li and Baker (2018), based on a relevant literature review, concluded that there is a positive relationship between behavioural engagement in MOOCs and performance in a quiz. Results by Lei et al. (2018), based on a meta-analysis of 69 studies published between 2003 and 2015, also reveal a positive relationship between behavioural engagement and performance. However, it should be noted that types of courses (i.e., a course in a traditional classroom, e-learning course, MOOC, etc.) considered in analysed studies were not specified. Meanwhile, some researchers have found no relation between behavioural engagement and test scores (Wang & Sui, 2020) or have concluded that it is not a significant predictor of MOOC completion (Lan & Hew, 2020).

Regarding cognitive engagement in the context of MOOCs, results have also varied across studies. While some researchers (Wang & Sui, 2020) have indicated a positive relationship between cognitive engagement and test scores, Li and Baker (2018) reviewed relevant literature and concluded that cognitive engagement in MOOCs can have both positive and negative relationship with performance in quiz. Meanwhile, Lei et al. (2018) in their meta-analysis found the relationship between cognitive engagement and performance being straightforwardly positive (although types of courses were not specified). In addition, cognitive engagement has been observed to be a strong significant positive predictor of MOOC completion (Lan & Hew, 2020).

A closer look at engagement with particular course components showed intriguing results. In the case of videos, behavioural engagement – measured by the number of whole videos watched – had a positive relationship with MOOC learners' quiz scores and was a positive strong significant predictor for quiz scores (Liu et al., 2018) and course completion (Pursel et al., 2016). However, in a study by Bonafini (2017) the number of videos watched was a non-significant predictor for course completion. In the case of measuring cognitive engagement with videos by the number of pauses (Li & Baker, 2016; Liu et al., 2018) or seeking in videos (Liu et al., 2018) a positive relationship with MOOCs learners' performance was revealed. These indicators of cognitive engagement were also significant predictors for quiz scores (Liu et al., 2018). Interestingly, without naming any specific engagement dimension, MOOC researchers have claimed that the number videos watched per week is positively associated with course completion rate (Pursel et al., 2016), while the number of backward and forward video seeks has weak negative correlation with grade (Dissanayake et al., 2018).

Behavioural engagement with forums, measured by the total number of posts, is a positive significant predictor for course completion (Bonafini, 2017; Pursel et al., 2016). Analysing the content of MOOC forums, Galikyan et al. (2021) distinguished several levels of cognitive engagement in forums with different relationships to course grade.

Researchers have attempted to find a correlation between engagement in assessments and learners' performance (de Barba et al., 2016; Kennedy et al., 2015). However, it should be noted that the authors did not specify what dimension of engagement they considered. While learners who are more inclined to resubmit their assessments are more likely to perform well, assessments submissions have a weak positive correlation with final grades (Kennedy et al., 2015). Interestingly, while assessments submissions do not predict final grades (Kennedy et al., 2015), the number of quiz attempts is the strongest predictor of learners' final grades (de Barba et al., 2016). This inconsistency may be related to the type of assessments since Kennedy et al. (2015) considered programming assessments on problem solving, while de Barba et al. (2016) considered quizzes with non-compulsory multiple-choice questions.

In sum, relationships between behavioural and cognitive engagement in the context of MOOCs and learners' achievements are contradictory. A closer look at relevant literature about engagement with a specific course component and its impact on learners' achievements revealed that results were not straightforward. Given the variety of learners' groups with different behavioural (e.g., Kahan et al., 2017; Kizilcec et al., 2013; Poellhuber et al., 2019) and cognitive (Li & Baker, 2016) engagement, and available results about their achievements, it would be interesting to gain insight into how different groups of MOOCs completers differ in performance.

3. RESEARCH METHODOLOGY

This chapter provides an overview of the methodology used in the doctoral thesis and is organised as follows: sub-chapter 3.1. presents generalised information about research design and proposed research model. Sub-chapter 3.2. provides a detailed overview of the context of the study. Sub-chapter 3.3. presents the study sample. Sub-chapters 3.4. and 3.5. describe the employed data collection and data analysis methods.

3.1. Research design

The study was guided by a research model (Figure 2), the development of which was based on a review of literature. The model was designed to understand how background variables, engagement and performance are related to each other because literature has hitherto revealed contradictory results about the relationship between these three parts.

In the model used in the current study, background variables include socio-demographics characteristics (i.e., age, gender, education, employment status) and previous experience (i.e., experience with programming and web-based education). These variables were selected because, despite being one of the most frequently studied ones, there is still ongoing disagreement among researchers about their influence on participation and course completion as well as engagement. These background variables were used to characterise course participants and completers' engagement clusters.

Engagement is considered from completers' point of view and encompasses their engagement with course activities (to facilitate subsequent identification of behavioural engagement clusters) and help sources (to facilitate subsequent identification of cognitive engagement clusters). In previous research, a limited number of activities, such as video lectures, discussion forums or assessments (Table 1), has been used to investigate behavioural engagement. In this thesis, however, the list of studied course activities is longer and, besides mandatory activities like reading learning materials on programming, includes activities that are not required to complete the course (e.g., demos, weblinks, encouraging emails, belletristic stories, etc.). Whereas cognitive engagement is rarely studied and the few studies that have been done on this topic have considered it through video interactions or forum posts (sub-chapter 2.2.2.), in the current thesis cognitive engagement is studied through completers' engagement with different help sources (such as helpdesk, troubleshooters, friends, etc.).

Performance is assessed through programming tasks and quizzes because they were used to set completion status. Besides comparing completers' and non-completers' performance, the study also focuses on investigating the impact of completers' behavioural and cognitive engagement on their performance.

Thus, the research model represents the idea that background variables can be used to characterise MOOC participants and may influence course completion and engagement, which in turn may influence performance.

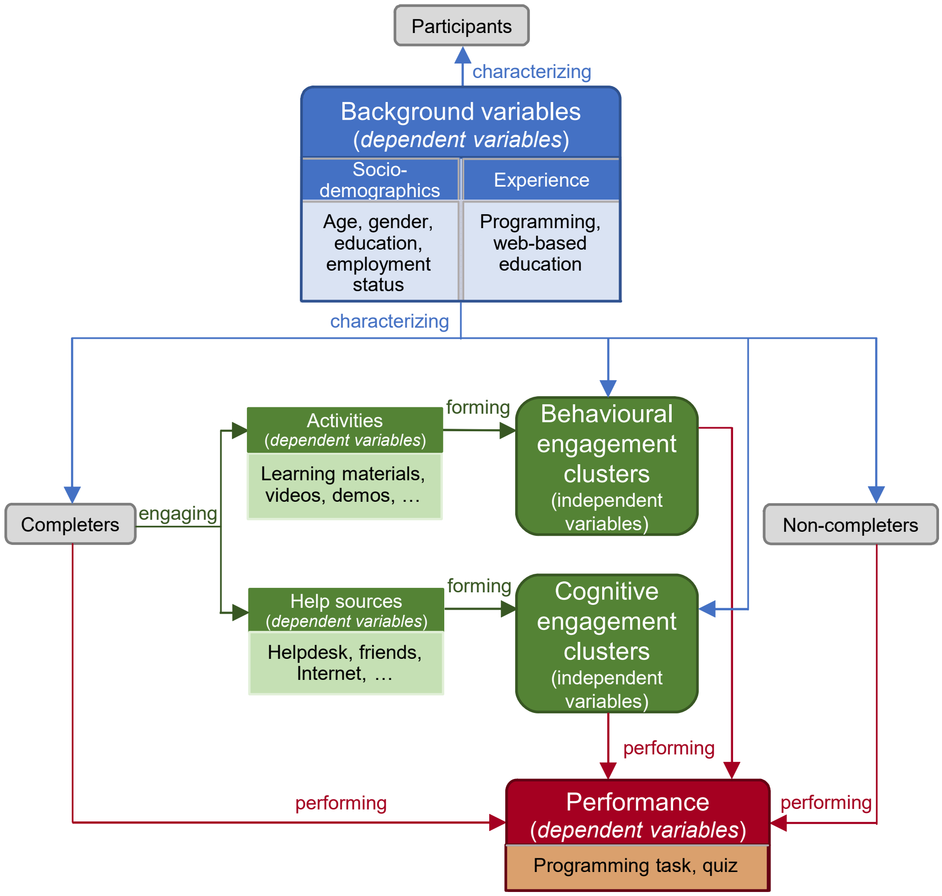

Figure 2. Research model for the doctoral study

In the current study, four research questions were posed. Various data collection and analysis methods were employed to answer them (Table 2).

Table 2. Research questions with data sources and analysis methods

Research questions	Data sources	Data analysis methods
RQ1: What background variables of participants in introductory programming MOOC can be identified compared to other programming MOOCs?	Study Information System Pre-questionnaire	*Descriptive statistics* of study participants *Comparison tests* to compare learners' background variables
RQ2: How did completers perform compared to the non-completers of introductory programming MOOC?	Moodle learning management system	*Descriptive statistics* of study participants and performance *Comparison test* to compare completers' and non-completers' performance
RQ3: What behavioural engagement clusters of completers can be identified in introductory programming MOOC, and how do completers' background variables and performance vary between clusters?	Study Information System Pre-questionnaire Post-questionnaire Moodle learning management system	*Descriptive statistics* of study participants, identified clusters and performance *k-means cluster analysis* to group completers based on behavioural engagement *Comparison tests* to compare the groups' background variables, behavioural engagement, and performance
RQ4: What cognitive engagement clusters of completers can be identified in introductory programming MOOC, and how do completers' background variables and performance vary between clusters?		*Descriptive statistics* of study participants, identified clusters and performance *k-means cluster analysis* to group completers based on cognitive engagement *Comparison tests* to compare the groups' background variables, cognitive engagement, performance

It should be noted that participants' background variables will be compared between three programming courses (described in sub-chapter 3.2.), while behavioural and cognitive engagement clusters of completers will be analysed only in the context of the MOOC "About Programming".

3.2. Study context

This sub-chapter provides an overview of the three programming MOOCs considered in this thesis. All three MOOCs were developed by the research group of informatics didactics from the Institute of Computer Science, University of Tartu. The group focused on the xMOOC model which suggests information distribution to learners from a centralised source (Overstreet, 2017), usage of computer-marked assessments and provision of supporting materials (Bates, 2019). The course developers relied on cognitive constructivist theory (Drew, 2019; Graduate Student Instructor Teaching & Resource Center, n.d.) according to which learners will be responsible for their own learning process and take an active role during the construction of knowledge. Following this theory, course instructors acted as facilitators who provided resources, guided learners throughout the course and offered opportunities to check learners' own understanding of the topics (e.g., by providing self-assessment questions or practical programming tasks). The author of this thesis joined the research group in September 2018. By this time, the MOOCs design and content had already been developed and courses had been run several times (Lepp et al., 2017a).

All three courses were taught in Estonian and focused on learning programming in Python. The first Estonian-language introductory programming MOOC was launched in December 2014 and was named "About Programming" (hereinafter, AP). Students and school teachers provided some assistance to maintain it (Lepp et al., 2017a). A year later, in 2015, the second MOOC named "Introduction to Programming I" (hereinafter, IP1) was introduced. In April 2017 the research group presented the third MOOC named "Introduction to Programming II" (hereinafter, IP2). The courses had different target groups and vary in terms of available credit points, level of difficulty, requirements, etc. (Table 3).

The MOOCs were held in the Moodle learning management system. Moodle provided a rich learning environment and possibilities to inform participants about organisational issues, track their progress, upload programming tasks solutions, submit quizzes, offer different support mechanisms like forums, etc. The course page also included links to learning materials. The course instructors made them open through a public platform developed and managed by the host institute. Since the studied MOOCs had common features in their design and organisation, the thesis includes a detailed overview of only AP. For other programming MOOCs only the main differences are highlighted.

Table 3. The summarised overview of programming MOOCs

	About Programming (AP)	Introduction to Programming I (IP1)	Introduction to Programming II (IP2)
Launch year	2014	2015	2017
Main target group	Learners with no previous experience in programming	Learners with no or little previous experience in programming	Learners with previous experience in programming at least to the level of IP1 (or equivalent)
Credit points	1 ECTS	3 ECTS	3 ECTS
Duration	4 weeks	8 weeks	8 weeks
Number of mandatory programming tasks per week	1-2	2-4	2-4
Number of mandatory quizzes per week	1	1	1
Total number of participants (completion rate)*	12,018 (62.2%)	7,182 (55.5%)	2,737 (32.6%)
Specific features	• Helpdesk availability	• Helpdesk availability	• Each learner must present the final project • No helpdesk available

Note. ECTS stands for "European Credit Transfer and Accumulation System" * As April 2022, https://didaktika.cs.ut.ee/moocid/

3.2.1. MOOC "About Programming"

This sub-chapter provides a detailed overview of the MOOC "About Programming" that is the main focus of the doctoral thesis. This information is useful to get an insight into the context in which the studied data were collected – what course was offered, what activities and help sources learners were provided to engage with and what assessments were used to evaluate learners' performance.

3.2.1.1. Course design

The MOOC "About Programming" lasted for 4 weeks with a possibility to earn 1 ECTS. The amount of work expected from a learner was 26 hours. The registration and participation in the MOOC were free of charge. This course mainly targeted adults but school kids could also participate. A learner did not have to have previous experience with programming. The course introduced the basic concepts of programming in Python and learner's obtained skills and knowledge were assessed with different programming tasks and quizzes (Table 4).

Table 4. Topics and number of mandatory assessments in the MOOC "About Programming"

Week	Topics	Number of programming tasks	Number of quizzes
1	Algorithms. Program. Variables. Data types	1	1
2	Conditionals. Strings. Turtle graphics	2	1
3	Loops. Regular expressions	2	1
4	Functions. Data exchange and files	1	1

The course consisted of eight blocks – two blocks per week. Each block started with an overview of activities that were recommended to learners to undertake during the particular week (Figure 3).

Figure 3. One of the blocks with guidelines to learners

To receive a certificate of completion, a learner had to successfully complete all six mandatory programming tasks and all four quizzes. Completion status was set automatically using the appropriate formula in the Moodle grade book. If a learner passed all mandatory assessments, his/her completion status was recorded as "passed". If a learner passed part of mandatory assessments, his/her completion status was recorded as "failed". But if a learner merely enrolled on the MOOC but never provided any solution to a programming task or submitted a quiz, the respective completion status was recorded as "not started".

3.2.1.2. Provided course activities

In the MOOC, different activities were provided. The data on the amount of each activity a completer engaged with were used to explore behavioural engagement. The course activities included learning materials that were composed for each block.

Recommended textual learning materials included explanations of new topics and overviews of IT-related topics like self-driving cars, programmable domestic appliances, labyrinths, etc. Learning materials included video clips with audio-visual explanations, links to additional resources and demo codes that demonstrated the usage of particular programming topics learnt.

In the course, there were optional self-assessment questions as well. The course instructors used different types of questions and did not store the results. Self-assessments helped learners to self-evaluate their understanding of programming aspects covered in particular block. All answers, including the right ones, returned feedback that explained why they were correct or not.

The course materials included supplementary resources in the form of belletristic stories about programming. Learners could also engage with activities such as troubleshooters, forums, encouraging videos and emails that are described in sub-chapter 3.2.1.4.

3.2.1.3. Mandatory assessments

During the course a learner demonstrated his/her understanding about the programming by passing mandatory assessments in the form of programming tasks and quizzes. The number of attempts per assessment and received scores per quiz were used to explore learners' performance. There were six mandatory programming tasks in total (Table 4). A learner had to solve one or two programming tasks each week to demonstrate the practical skills obtained. Learners from the same course run were given the same tasks. Each task had a problem description with input (and constraints, if necessary), and a picture of expected program output. Some tasks also had general clues to potentially help learners better understand the context of the task (Figure 4). A learner was expected to create a program and to upload a solution file to the Virtual Programming Lab. This is a Moodle plugin that course instructors used to set up solution checking tests (Rodríguez-del-Pino, 2012). Thereby, the solution provided by a learner was checked by an automated assessment system. In order to receive a "passed" grade, the solution output had to be consistent with the task requirements. Depending on a task, the system could qualify a wrong variable name or a missing function as an error. If the system detected an error, the solution was automatically graded as "failed". In the case of a failed checking test the appropriate feedback from the system was displayed. A learner was expected to read the feedback comments, improve the code and upload the file with the solution once again. The number of attempts to submit a solution was unlimited. Those who wanted to practice some more could solve optional programming tasks.

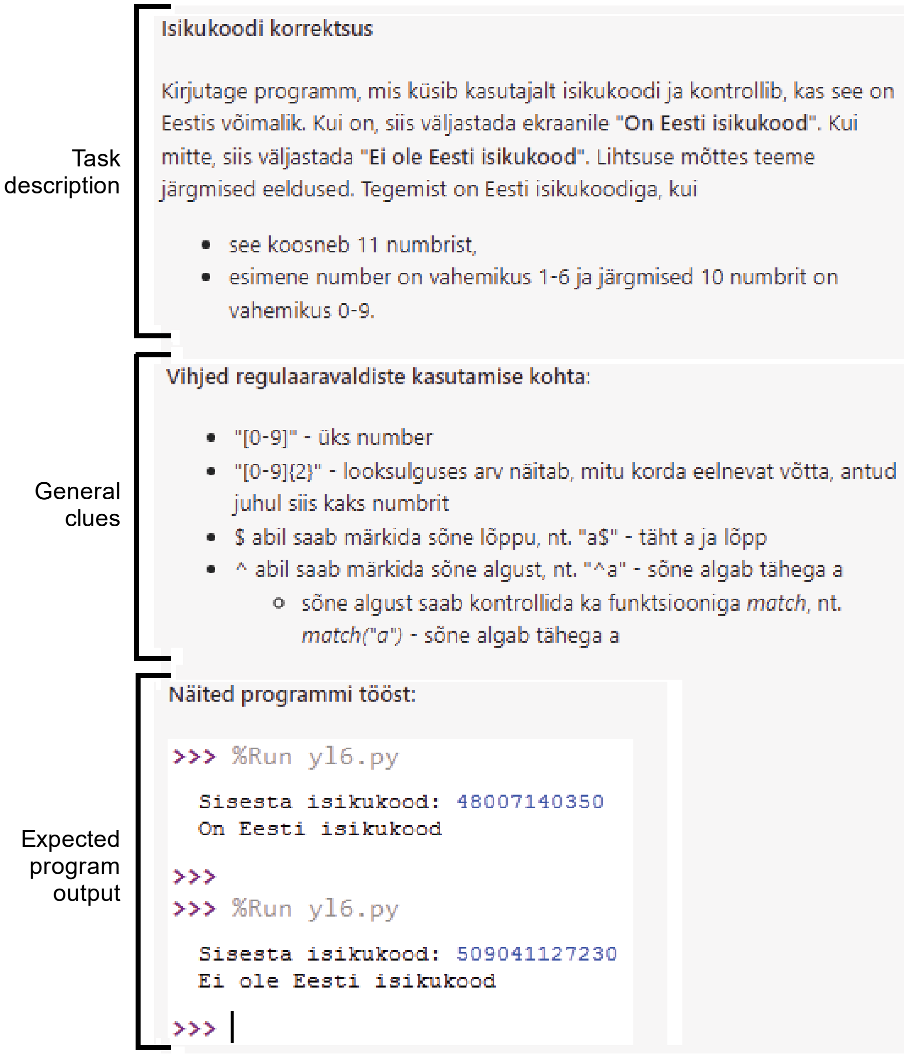

Figure 4. A structure of a mandatory programming task

There were four mandatory quizzes in total (one per week) with ten questions in each. The quizzes included theoretical questions, questions related to IT subjects; there were also questions which required from a learner a given code analysis and prediction of possible output (Lepp et al., 2017b). Learners from the same course run got the same questions and in the same order. A required threshold value for each quiz was 9 points out of 10. Multiple-choice, short answer, and matching question types were used in the automatically checked quizzes (Figure 5). After finishing a quiz, a learner had an opportunity to see where he/she went wrong and what the correct answers should be. Each answer to a multiple-choice question was provided with feedback, explaining why the answer was wrong or right. There was no time limit assigned per quiz attempt and a learner could solve a quiz as many times as he/she saw fit.

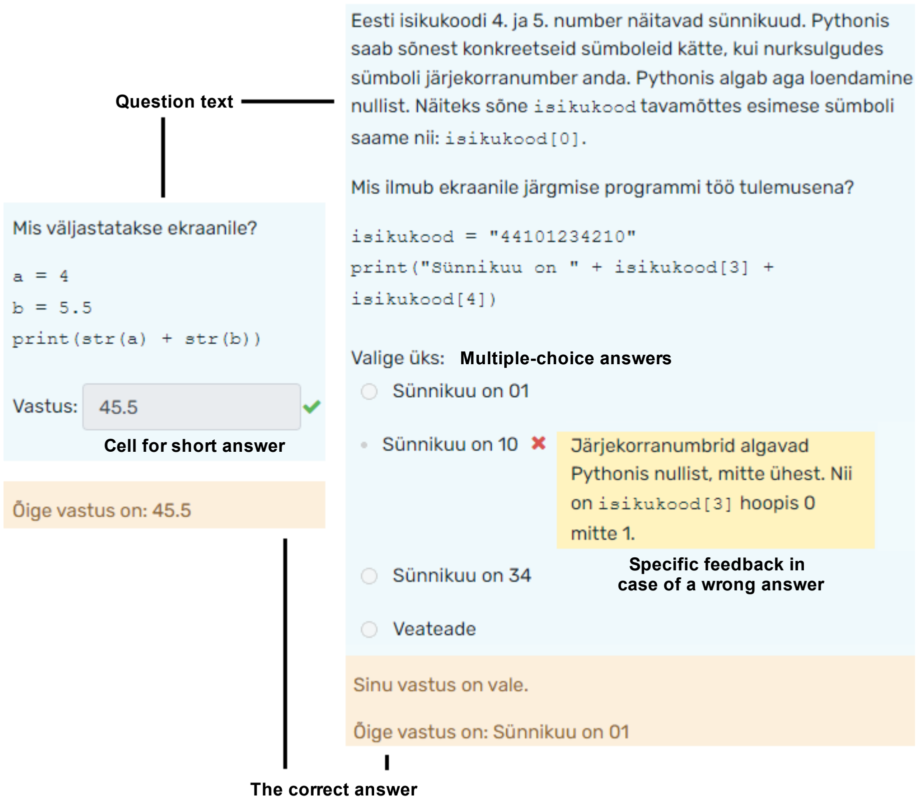

Figure 5. Example of short answer (left) and multiple-choice (right) question types and of feedback returned in case of a wrong answer (right)

3.2.1.4. Support mechanisms

Various support mechanisms were provided in the MOOC. The data about the frequency with which a completer engaged with each of them in case of a difficulty were used to explore cognitive engagement. A learner could seek help from forums and if he/she wanted to discuss some aspects of programming, had a question or a difficulty he/she could make a post in the forum. Each week had its own separate forum to avoid a mess in topic discussions. It was forbidden to share solutions and if somebody still posted them the course instructors deleted those posts. However, clues in the form of pointing to the appropriate place in the learning materials, asking guiding questions, providing links to additional materials were welcomed.

For each programming task, course instructors provided a troubleshooter that was implemented in the form of a tree (Lepp et al., 2018). It did not show the whole solution but contained clues on certain aspects for every programming task. At every step, learners were asked about possible problems they might have overlooked at any given moment (e.g., "Have you used the function *upper*?" or "Have you imported the module *re*?"). Those questions could help a learner figure out the source of the problem and, depending on the response, a learner could be forwarded to a page with explanation on how to solve the problem or to a page with the next possible reason of the problem (Figure 6).

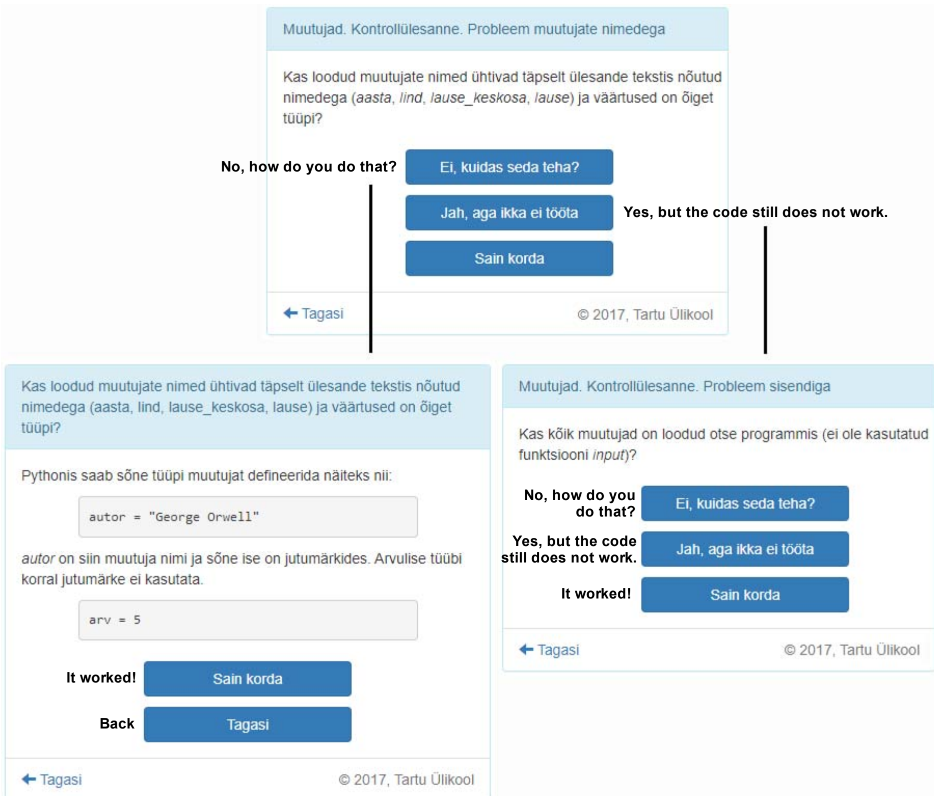

Figure 6. Troubleshooter pages with a possible problem statement and solution suggestion

In the course, there was also a list of frequently asked questions provided with a typical set of questions. The presence of a helpdesk with a quick response time of 8 hours (even at weekends) from the course teaching staff distinguished this course from traditional MOOCs. Similar to a forum, a learner was expected to describe the problem and the teaching staff gave clues but not the whole solution.

Weekly encouraging videos and emails were sent to learners to motivate them, draw their attention to the most difficult moments and provide an overview of what was happening in the course.

3.2.2. MOOCs "Introduction to Programming I" and "Introduction to Programming II"

Both courses were longer in time compared to AP and lasted for 8 weeks. The amount of work expected from a learner was 78 hours (worth 3 ECTS). Like AP, these two MOOCs also provided learners with weekly guidelines.

IP1 was targeting learners who had no or little previous experience with programming. Completion of AP was welcomed but not absolutely necessary. Although some topics were repeated, concepts and structures of programming were explored more deeply. The course introduced the basics of algorithmic thinking, arrays and graphical user interface. The amount of covered topics (usually one or two) per week depended on the complexity of the considered topic. Every week a learner had to pass up to four mandatory programming tasks and one quiz. Support mechanisms and requirements for course completion were the same as for AP.

IP2 was targeting learners who had previously participated in IP1 or in a similar course. The course also included topics such as two-dimensional arrays, double loops, data structures, etc. Mostly, a learner was expected to go through one topic per week. During the first six weeks, obtained skills and knowledge were checked with mandatory assessments. Learners solved up to four programming tasks and one quiz per week. The last two weeks were dedicated to the mandatory final project. During this period learners worked on their own projects and evaluated a project of a course mate. Learning materials provided in this course did not have stories presented in a belletristic way. Unlike the other two MOOCs this course did not provide support from a helpdesk.

3.3. Sample

In the current study different samples were used to answer the research questions. With regard to the first research question, the data from three above-mentioned MOOCs run in the 2016/2017 academic year were used. The sample comprised 1,536 AP participants who filled out the voluntary questionnaire (85.7% of enrollers), 1,282 IP1 participants (82.5% of enrollers), and 744 IP2 participants (78.7% of enrollers). The sample used to answer the second research questions comprised 292 non-completers and 773 completers from the AP run in September 2018 (Publication IV; Table 5).

Table 5. Description of the sample used in the second research question (based on the data from the Study Information System)

	AP non-completers	AP overall completers
Sample size	292	773
Female	53.4%	54.7%
Average age	36.4 (SD = 10.47)	33.2 (SD = 10.94)
Age range	10–60	10–70

The sample for the last two research questions comprised solely completers who successfully passed the AP course run in September 2018 and filled in both pre- and post-questionnaires (see sub-chapter 3.4. Data collection). An overview of this sample is provided in Table 6.

Table 6. Description of the sample used in the third and fourth research questions (based on the data from the Study Information System)

	AP completers who filled in questionnaires
Sample size	580
Female	56.7%
Average age	33.0 (SD = 11.08)
Age range	10–70

Among completers studied in the third and fourth research questions 61.7% had higher education, about half (50.2%) had no experience in learning programming and for about two thirds (60.9%) this was their first web-based course (Publication III).

It should be emphasised that the sample used to study performance by completers from identified clusters included data from 575 completers. The sample was slightly reduced because five completers did not consent to the use of their performance data.

3.4. Data collection

Quantitative data used in this study were collected from four different sources: the Study Information System, the Moodle learning management system and optional web-based pre- and post-questionnaires.

The Study Information System of the University of Tartu was used by the participants to sign up for the courses. The system collected data about participants' gender and age. These data were used to answer research questions 1, 3 and 4 (Table 7).

More data about participants' background was collected with the pre-questionnaire. The author joined the research group in September 2018, when the content of MOOCs and the web-based pre-questionnaire had already been developed and successfully used in several course runs. A link to the pre-questionnaire was provided in Moodle at the beginning of the course and filling it out was voluntary for participants. The pre-questionnaire included five question blocks. The first block about socio-demographic characteristics included questions related to participants' education level and employment status. The second block covered previous experience with programming education and web-based learning. The other three blocks were related to the participants' intention to complete the course, enrolment motivation and views on the work of a programmer but these topics were out of the thesis scope. Data from the first two blocks of the pre-questionnaire were used to answer research questions 1, 3 and 4 (Table 7).

After joining the research group, the author of the doctoral thesis started working on modifying a web-based post-questionnaire. It was intended to collect data about learners' engagement with course activities and help sources. While working on the post-questionnaire, the author of the thesis considered the following issues: relevance of the content, the wording and sequence of questions. This approach relied on suggestions by Sekaran and Bougie (2016) to keep statements short and simple. When the first version was created the author had an opportunity to discuss it within the research group. As a result, some irrelevant items were removed and the wording was changed to be more appropriate. The composed post-questionnaire was used in several course runs and a link to it was provided in Moodle at the end of the course. Filling out the post-questionnaire was voluntary for learners and had no influence on the course completion. The post-questionnaire included six question blocks. Keeping in mind the aim of the thesis, the author used two blocks. The block about behavioural engagement included a list of eleven activities provided in the course and, using a 5-point Likert scale, learners rated their engagement with each activity throughout the course (Appendix A). The block about cognitive engagement included a list of seven help sources and, using a 5-point Likert scale, learners rated their engagement with each of them in the case of facing a difficulty while solving a programming task (Appendix B). Completers' data from these two blocks were used to answer research questions 3 and 4 (Table 7).

In this study the author opted for a post-questionnaire instead of learning analytics since the latter provided data only about learners' behaviour in the Moodle learning management system. However, the author was also interested in investigating activities taken by learners off Moodle such as searching for additional materials on the Internet or asking somebody for help. Besides that, questionnaires are a common data collection method in research on behavioural (Phan et al., 2016) and cognitive (Deng et al., 2020b; Jung & Lee, 2018) engagement in MOOCs. With regard to cognitive engagement, questionnaires are recommended for capturing non-behavioural data (Samoilova et al., 2018). The response rates to both questionnaires within the studied groups were high (above 70%) which indicates a low likelihood of response errors (van de Oudeweetering & Agirdag, 2018).

Given that all mandatory assessments were implemented in Moodle, this learning management system was used to collect the following data about each learner: the number of attempts made per each mandatory programming task and quiz, the received scores per each quiz, and the course completion status. These data were used to answer research questions 2, 3 and 4 (Table 7).

This study followed instructions of personal data management developed by the Republic of Estonia Data Protection Inspectorate (Andmekaitse Inspektsioon, 2019) and The European Code of Conduct for Research Integrity developed by All European Academies (2017). In line with the standards, course participants were informed about the overall aim of the questionnaires, the purpose of data collection, the intended further usage of data, confidentiality of their responses and rights of the respondents, including the opportunity to with-draw their consent to participation in the research and the use of their data at any time. All collected data were stored in the internal database of the host institute and were managed only by one data specialist. The author of the doctoral thesis received data only in an anonymised format.

Table 7. An overview of data used in the study (ticks indicate variables that were used to answer research questions of the doctoral thesis)

Blocks	Variables	Data types	RQ1	RQ2	RQ3	RQ4
Socio-demographic characteristics	gender	categorical (binary)	✓		✓	✓
	age	numeric	✓		✓	✓
	education level	categorical (ordinal)	✓		✓	✓
	employment status	categorical (nominal)	✓		✓	✓
Previous experience	with programming education	categorical (ordinal)	✓		✓	✓
	with web-based education	categorical (ordinal)			✓	✓
Amount of engagement with each activity provided in the list (for behavioural engagement clusters)*	reading of learning materials	categorical (ordinal)			✓	
	visiting of the weblinks	categorical (ordinal)			✓	
	watching learning videos	categorical (ordinal)			✓	
	…	categorical (ordinal)			✓	
Frequency of engagement with each help source provided in the list (for cognitive engagement clusters)*	rereading of learning materials	categorical (ordinal)				✓
	searching on the Internet	categorical (ordinal)				✓
	using troubleshooters	categorical (ordinal)				✓
	…	categorical (ordinal)				✓
Performance	attempts per programming task	numeric		✓	✓	✓
	attempts per quiz	numeric		✓	✓	✓
	received quiz scores	numeric		✓	✓	✓

Note. * Using 5-point Likert scale learner's evaluation of the extent of agreement with each statement ranged from 0 = "not at all/never" to 4 = "all/always"

3.5. Data analysis

Collected quantitative data were analysed using software IBM SPSS Statistics 25.0 and 26.0. A preliminary analysis of the variables did not find missing data. Outliers were found only in the data about performance but were considered as a normal part of the data distribution (Publication IV) and were not excluded.

Descriptive statistics were used to calculate the average, mean, median, standard deviation and range of different variables. These were used to give an overview of study samples (Publications I–IV) and learners' performance (Publication IV), to describe some background variables of participants in the three programming MOOCs (Publication I) and of completers from clusters of behavioural engagement (Publication II) and cognitive engagement (Publication III).

Comparison tests (Field, 2009) were used to look for statistically significant differences within and between studied groups (Publications I-IV; Table 8).

Table 8. Comparison tests used to answer the research questions

<table>
<tr><th>Tests</th><th>RQ1</th><th>RQ3</th><th>RQ4</th><th>RQ2</th></tr>
<tr><td rowspan="2">ANOVA with Bonferroni post hoc test</td><td colspan="3">age</td><td></td></tr>
<tr><td></td><td>activities</td><td></td><td></td></tr>
<tr><td rowspan="2">Chi-square test</td><td colspan="3">gender, employment status, previous experience with programming education and web-based education</td><td></td></tr>
<tr><td></td><td colspan="2">education level</td><td></td></tr>
<tr><td rowspan="2">Kruskal-Wallis H-test</td><td>education level</td><td></td><td>help sources</td><td></td></tr>
<tr><td></td><td colspan="3">performance</td></tr>
<tr><td rowspan="2">Mann-Whitney U-test</td><td>education level</td><td></td><td>help sources</td><td></td></tr>
<tr><td></td><td colspan="2">performance</td><td></td></tr>
</table>

k-mean cluster analysis with the Euclidean distance was employed to identify different clusters of completers' behavioural and cognitive engagement. Given the *massiveness* of MOOCs, this exploratory technique has been suggested to identify clusters without knowing in advance what should be found (Kahan et al., 2017) by grouping exploring features. It has been actively employed in studying behavioural engagement (e.g., Arora et al., 2017; Khalil & Ebner, 2017) but remains under-explored with regard to cognitive engagement (Li & Baker, 2016). To distinguish behavioural engagement clusters, an analysis with a maximum of 10 iterations for models with three to six clusters was applied. In terms of cognitive engagement clusters, the analysis with a maximum of 20 iterations for models with three to five clusters was applied. The final number of clusters was chosen based on having enough members in each cluster and keeping in mind the educational perspective. Thus, four behavioural engagement clusters (Publication II) and five cognitive engagement clusters (Publication III) were identified among completers of AP.

4. FINDINGS

This chapter provides a summary of the main findings in accordance with research questions posed in the doctoral thesis. Sub-chapter 4.1. focuses on participants and completers in AP and their comparison to those in the other two programming MOOCs. Sub-chapter 4.2. provides an overview of the difference in performance between AP completers and non-completers. Sub-chapter 4.3. describes the identified behavioural engagement clusters, background variables and performance of the members of respective clusters. Sub-chapter 4.4. provides an overview of the identified cognitive engagement clusters, background variables and performance of the members of respective clusters.

4.1. Comparing participants and completers in AP with those in the other two programming MOOCs

The current sub-chapter answers the first research question. More detailed results were provided in Publication I.

4.1.1. AP participants versus participants in the other two programming MOOCs

Based on the pre-questionnaire data the background variables of MOOC participants were as follows. Slightly more than a half of the AP participants were female (Table 9) and their proportion in AP was significantly higher in comparison to IP1 (chi-square = 4.575, $p = .032$) and IP2 (chi-square = 49.185, $p < .001$). Most of the participants in all three MOOCs belonged to the age group 26-35. Meanwhile, there was no significant difference in age between the participants in AP and IP1 ($t = 1.192$, $p = .233$), but the AP participants were statistically younger than those in IP2 ($t = 5.327$, $p < .001$). With regard to education level, the AP participants had statistically lower highest obtained education level compared to the other MOOCs' participants (compared to IP1, $U = 786{,}655.0$, $p < .05$; compared to IP2, $U = 516{,}912.0$, $p < .001$). Like in the other studied programming MOOCs, most of the AP participants were employed. While a comparison of the AP and the IP1 results indicated that in AP the proportion of students was statistically higher (chi-square = 7.115, $p < .001$), the proportion of those who did not work or study was statistically lower (chi-square = 6.148, $p < .05$). No statistical difference was found between AP and IP2 in terms of employment status (in all cases $p > .05$). Among the AP participants there were statistically more of those who did not have any previous experience in programming (compared to IP1, chi-square = 326.137, $p < .0001$). Although less than a half of the AP participants had received any programming experience during their formal education, this proportion was statistically lower compared to the IP1 participants (chi-square = 336.719, $p < .0001$). Enrolment in IP2 required previous experience in programming and, hence, this question was not addressed to the IP2 participants. The proportion of non-starters was statistically higher in AP than in the other two MOOCs (compared to IP1, chi-square = 37.547, p < .001; compared to IP2, chi-square = 16.709, p < .001). While in IP1 younger people were more likely to enrol in the MOOC but not start (ANOVA with Bonferroni post hoc test F = 4.857, p < .01), no statistically significant difference in age between these three groups was found in AP and IP2 (in all cases p > .05).

Table 9. Participants' background variables in the three MOOCs on programming

		AP	IP1	IP2	Statistically significant difference in comparison to AP
Female		54.9%	51.2%	40.2%	AP > IP1 AP > IP2
Age range		8-76	12-73	13-73	N/A
Average age		33.7 (SD = 10.78)	34.1 (SD = 10.50)	36.0 (SD = 10.66)	AP < IP2
Education level	median	higher education	higher education	higher education	AP < IP1 AP < IP2
	mode	secondary education	higher education	higher education	
Employment status	employed	76.9%	77.6%	77.6%	
	student	14.6%	11.2%	11.8%	AP > IP1
	retired	1.3%	1.5%	1.6%	
	did not work or study	7.2%	9.8%	9.0%	AP < IP1
Lack of experience in programming		38.2%	8.7%	*previous experience is mandatory*	AP > IP1
Programming experience from formal education		42.1%	76.4%		AP < IP1
Registered, but did not start		18.6%	11.0%	12.5%	AP > IP1 AP > IP2

Note. N/A stands for "not applicable"

4.1.2. AP completers versus completers of the other two MOOCs on programming

Data about completers' background variables were collected from the pre-questionnaire and their course completion status was retrieved from Moodle. About two thirds of the AP participants successfully passed the MOOC. The proportion of completers in AP was statistically higher than in the other two programming MOOCs (compared to IP1, chi-square = 27.844, $p < .001$; compared to IP2, chi-square = 233.755, $p < .001$). No statistically significant difference between females and males was found in AP (chi-square = 1.492, $p = .474$) and the situation was same in IP1 (chi-square = .496, $p = .78$). However, male completers dominated in IP2 (chi-square = 11.083, $p = .004$). In both AP and IP1, master's degree holders were more likely to complete the MOOC compared to those with lower education level (with chi-square test, in all cases $p < .05$). However, in IP2 those with basic education level were more likely to complete the course (with chi-square test, in all cases $p < .05$). In both AP and IP1, no statistical difference in employment status was found (for AP chi-square = .625, $p = .891$; for IP1 chi-square = .415, $p = .937$). However, in IP2 employed people were less likely to complete the course (with chi-square test, $p < .05$). In both AP and IP1, completers with no previous experience in programming were less likely to complete the course than those who had studied programming before (for AP chi-square = 7.599, $p < .001$; for IP1 chi-square = 15.339, $p < .001$). Although in IP1 those who had obtained experience in programming from formal education were more likely to complete the MOOC (chi-square = 21.815, $p < .001$), no statistical difference was found in the case of AP (chi-square = .921, $p = .337$). For IP2, previous experience in programming was mandatory.

4.2. Performance by AP completers and non-completers

The current sub-chapter answers the second research question. More detailed results were provided in Publication IV.

Throughout the course, the AP learners were expected to solve six mandatory programming tasks (Table 4). On average, learners made 2.05 (SD = 3.931) attempts per task. While among completers the number of attempts per task varied from 1 to 121, among non-completers it varied from 1 to 55 (Table 10). Completers, on average, made statistically fewer attempts per task than non-completers ($H = 46.973$, $p < .001$).

Table 10. Attempts per programming task by completers and non-completers

	Completers	**Non-completers**
n	773	231
range	1–121	1–55
M (SD)	1.97 (3.902)	2.67 (4.107)

The mandatory assessments also included four quizzes (Table 4). On average, learners made 1.04 (SD = .218) attempts per quiz and received, on average, 9.73 (SD = 1.063) points. Completers made up to 4 attempts per quiz, while non-completers up to 3 (Table 11). Given the course completion requirements, among completers the quiz scores varied from 9 to 10, while among non-completers it ranged from 0 to 10. Although no significant difference in the average number of attempts per quiz between completers and non-completers was found (H = 1.130, $p > .05$), the difference was significant in terms of quiz scores (H = 85.037, $p < .001$).

Table 11. Attempts and received scores per quiz by completers and non-completers

		Completers	**Non-completers**
	n	773	288
Attempts per quiz	range	1-4	1-3
	M (SD)	1.04 (.211)	1.05 (.251)
Scores per quiz	range	9-10	0-10
	M (SD)	9.86 (.345)	9.06 (2.423)

4.3. AP completers' behavioural engagement clusters, background variables and performance of members

The current sub-chapter answers the third research question. More detailed results were provided in Publication II and Publication IV.

4.3.1. Identified behavioural engagement clusters among completers

Completers self-evaluated the amount of activities they engaged with throughout the course (Appendix A). The study results identified four behavioural engagement clusters among the AP completers (Figure 7). Members in each cluster more or less engaged with all eleven activities listed in the post-questionnaire. In all activities the difference between clusters was significant ($p < .01$). The difference was larger in activities such as reading additional materials (F = 183.335) and weekly encouraging emails (F = 144.535) but was smaller in activities such as answering self-assessment questions (F = 41.826), trying provided demos (F = 51.494), and reading forum posts (F = 52.282).

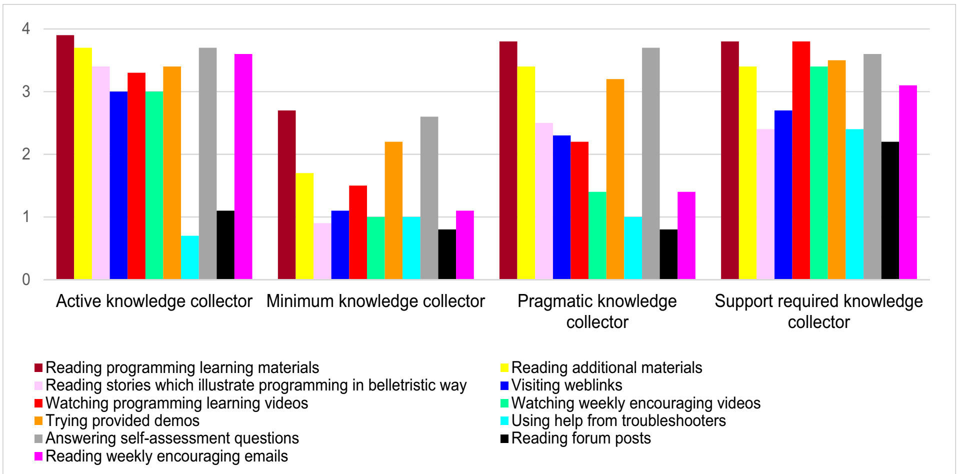

Figure 7. Identified behavioural engagement clusters among AP completers

The cluster of *active knowledge collectors* (n = 170; 29.3% of studied completers) consisted of completers who engaged with most of the activities. However, these completers read a few forum posts and used some troubleshooters. They engaged more with reading additional materials, belletristic stories and weekly encouraging emails than any others (in all cases $p < .01$).

The cluster of *minimum knowledge collectors* (n = 95; 16.4% of studied completers) consisted of completers who mainly engaged with reading learning materials, answering self-assessment questions and trying demos. Compared to members of other clusters, they engaged with less amount of activities.

The cluster of *pragmatic knowledge collectors* (n = 201; 34.6% of studied completers) consisted of completers who mainly engaged with reading learning materials and additional materials, trying provided demos and answering self-assessment questions.

The cluster of *support required knowledge collectors* (n = 114; 19.7% of studied completers) consisted of completers who engaged with all the provided activities. They engaged more with watching programming learning videos, using troubleshooters, and reading forum posts than members of any other cluster (in all cases $p < .01$).

The analysis of distances between cluster centres identified the greatest differences between the clusters of *active* and *minimum knowledge collectors* (5.595). The clusters of *active and support required knowledge collectors* were most similar (2.430). The identified behavioural engagement clusters were described in more detail in Publication II.

4.3.2. Member background variables

The collected pre-questionnaire data made it possible to investigate member background variables. Table 12 provides an overview of the findings regarding variables of the members of each identified behavioural engagement clusters. It should be noted that age was compared between clusters, while other variables were compared within clusters.

Table 12. Background variables of members of behavioural engagement clusters

	Active knowledge collector	**Minimum knowledge collector**	**Pragmatic knowledge collector**	**Support required knowledge collector**
Gender	NS	More males	NS	NS
Age	Older than *minimum* and *pragmatic* ones	Younger than *active* and *support required* ones	Younger than *active* ones	Older than *minimum* ones
Education level	More of those with higher education level	More of those with secondary or lower education level	NS	NS
Employment status	More working and inactive people compared to students	More students compared to working and inactive people	NS	More inactive people compared to students
Previous experience with programming education	NS	More of those with previous experience	NS	More of those with no previous experience
Previous experience with web-based learning	NS			

Note. NS stands for "no statistically significant"

More detailed results about the variety of member background variables of different behavioural engagement clusters were provided in Publication II.

4.3.3. Performance by cluster members

Submitted programming tasks and quizzes were used to study the performance by completers from different behavioural engagement clusters. As expected, in every cluster there were completers who passed the programming task on the first attempt. However, the maximum number of attempts made varied. In the cluster of *support required knowledge collectors* (hereinafter, *KC* refers to knowledge collectors) this number reached 121 attempts per task, while in all other clusters this number was below 20 (Table 13). A significant difference was found between behavioural engagement clusters in the average number of attempts per task ($H = 45.500$, $p < .001$). Completers from the cluster of *support required KC* made more attempts, on average, compared to completers from all other clusters (in all cases $p < .01$). *Minimum KC* made, on average, more attempts than *active* and *pragmatic KC* (in both cases $p < .05$).

Table 13. Attempts per programming task within completers' behavioural engagement clusters

	Active KC	Minimum KC	Pragmatic KC	Support required KC
n	169	95	197	114
range	1–19	1–18	1–15	1–121
M (SD)	1.62 (1.638)	1.77 (1.811)	1.63 (1.656)	2.91 (7.514)

With regard to quizzes, the number of attempts varied across the identified clusters from 1 to 4 (Table 14). Although the difference in the average number of attempts per quiz was quite small between behavioural engagement clusters, it was still statistically significant ($H = 17.662$, $p < .001$). Completers from the cluster of *active KC* made, on average, fewer attempts compared to completers from all other clusters (in all cases $p < .05$). *Pragmatic KC* made, on average, fewer attempts than *support required KC* ($p < .05$).

Table 14. Attempts and received scores per quiz within completers' behavioural engagement clusters

	Active KC	Minimum KC	Pragmatic KC	Support required KC
n	169	95	197	114
range	1–3	1–3	1–2	1–4
Attempts per quiz, M (SD)	1.02 (.138)	1.06 (.251)	1.03 (.175)	1.07 (.289)
Scores per quiz, M (SD)	9.92 (.267)	9.81 (.394)	9.90 (.294)	9.82 (.381)

A significant difference was found between clusters in terms of average received quiz scores ($H = 47.774$, $p < .001$). Average quiz scores received by *active KC* and *pragmatic KC* were higher than the ones received by *minimum*

KC and *support required KC* (in all cases p < .001). More detailed results about performance by behavioural engagement clusters' members were provided in Publication IV.

4.4. AP completers' cognitive engagement clusters, background variables and performance of members

The current sub-chapter answers the fourth research question. More detailed results were provided in Publication III and Publication IV.

4.4.1. Identified cognitive engagement clusters among completers

Completers self-evaluated their frequency in engaging with help sources in the case of encountering a difficulty while solving a programming task (Appendix B). The study results identified five cognitive engagement clusters among the AP completers (Figure 8). Members in each cluster more or less engaged with all seven help sources listed in the post-questionnaire. In all help sources the difference between clusters was significant ($p < .05$). The difference was larger in using troubleshooters ($F = 310.276$) and writing or reading forum posts ($F = 296.038$). For all clusters, rereading of learning materials was an undoubted important step in resolving difficulties. The analysis showed that the difference between all identified clusters was the smallest regarding this help source ($F = 3.030$).

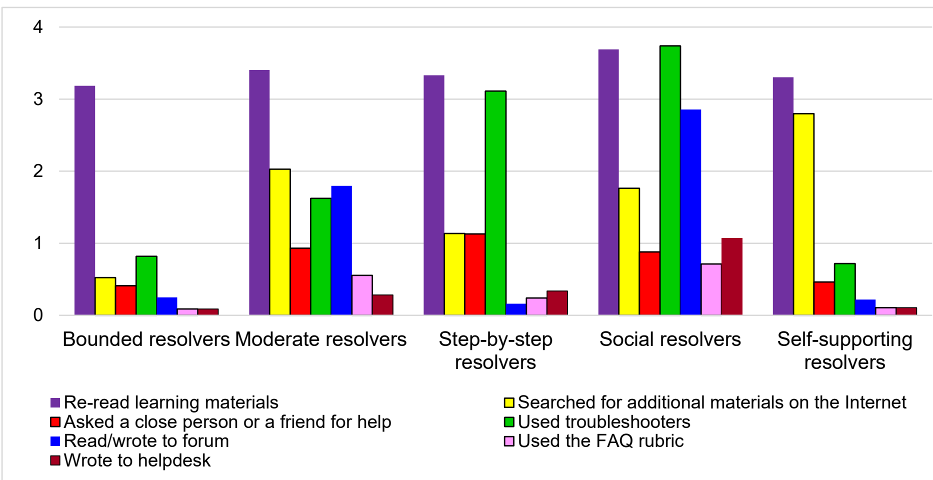

Figure 8. Identified cognitive engagement clusters among AP completers

The cluster of *bounded resolvers* (n = 172; 29.7% of studied completers) consisted of completers who, in the case of a difficulty, very often reread learning materials. They almost never used any communication channels to ask for help. *Bounded resolvers* were less likely to search for additional materials on the Internet than members of any other clusters ($p < .01$).

The cluster of *moderate resolvers* (n = 74; 12.8% of studied completers) consisted of completers who, besides rereading learning materials, often searched for materials on the Internet, used forums and troubleshooters.

The cluster of *step-by-step resolvers* (n = 124; 21.4% of studied completers) consisted of completers who mainly reread learning materials and used troubleshooters. Occasionally, they searched for additional materials on the Internet or asked a friend for help.

The cluster of *social resolvers* (n = 42; 7.2% of studied completers) consisted of completers who, compared to all others, engaged more frequently with rereading learning materials, using troubleshooters, and seeking help from forums and helpdesk (in all cases $p < .01$).

The cluster of *self-supporting resolvers* (n = 168; 29.0% of studied completers) consisted of completers who frequently reread learning materials and searched materials on the Internet. They were more often engaged with searching materials compared to all others (in all cases $p < .01$). Their frequency in engaging with other help sources was quite similar to *bounded resolvers*.

The analysis of distances between cluster centres identified the greatest differences between the clusters of *social* and *bounded resolvers* (4.322) and between *social* and *self-supporting resolvers* (4.331). The clusters of *self-supporting* and *moderate resolvers*, as well as the clusters of *self-supporting* and *bounded resolvers*, were most similar (respectively, 2.087 and 2.281). The identified cognitive engagement clusters were described in more detail in Publication III.

4.4.2. Member background variables

The data collected from the pre-questionnaire made it possible to investigate member background variables. Table 15 provides an overview of the findings regarding different background variables of members of each identified cognitive engagement clusters. It should be noted that age was compared between clusters, while other variables were compared within clusters.

Table 15. Background variables of members of cognitive engagement clusters

	Bounded resolvers	Moderate resolvers	Step-by-step resolvers	Social resolvers	Self-supporting resolvers
Gender	NS	NS	NS	More females	More males
Age	NS				
Education level	NS				
Employment status	NS				
Previous experience in learning programming	More of those with previous experience	NS	More of those with no previous experience	More of those with no previous experience	More of those with previous experience
Previous experience in web-based learning	NS	More of those with no previous experience	NS	More of those with no previous experience	More of those with no previous experience

Note. NS stands for "no statistically significant"

More detailed results about the variety of member background variables of different cognitive engagement clusters were provided in Publication III.

4.4.3. Performance by cluster members

Submitted programming tasks and quizzes were used to analyse performance by completers from different cognitive engagement clusters. Unsurprisingly, in every cluster there were completers who passed the programming task on the first attempt. Meanwhile, the maximum number of attempts made varied. Among *step-by-step resolvers* there was a completer who made 121 attempts per task, while a *social resolver* made 96 attempts. In all other clusters this number was below 40 (Table 16). A significant difference was found between cognitive engagement clusters in the average number of attempts per task ($H = 45.696$, $p < .001$). Completers from the cluster of *bounded resolvers* made, on average, fewer attempts than completers from all other clusters (in all cases $p < .05$). *Self-supporting resolvers* made, on average, fewer attempts per task than *moderate*, *step-by-step*, and *social resolvers* (in all cases $p < .05$).

Table 16. Attempts per programming task within completers' cognitive engagement clusters

	Bounded resolvers	Moderate resolvers	Step-by-step resolvers	Social resolvers	Self-supporting resolvers
n	170	72	123	42	168
range	1–33	1–36	1–121	1–96	1–19
M (SD)	1.56 (1.766)	2.07 (2.904)	2.44 (6.095)	2.46 (6.330)	1.65 (1.583)

Considering the quizzes, the number of attempts also varied from 1 to 4 across identified clusters (Table 17). A statistically significant difference was found in the average number of attempts per quiz between identified clusters ($H = 21.648$, $p < .001$). Completers from the cluster of *bounded resolvers* made, on average, fewer attempts compared to *step-by-step*, *social*, and *self-supporting resolvers* (in all cases $p < .01$). *Moderate resolvers* made, on average, fewer attempts per quiz than *step-by-step* and *social resolvers* ($p < .05$).

Table 17. Attempts and received scores per quiz within completers' cognitive engagement clusters

	Bounded resolvers	Moderate resolvers	Step-by-step resolvers	Social resolvers	Self-supporting resolvers
n	170	72	123	42	168
range	1–2	1–2	1–4	1–3	1–3
Attempts per quiz, M (SD)	1.01 (.114)	1.02 (.143)	1.06 (.279)	1.08 (.309)	1.04 (.211)
Scores per quiz, M (SD)	9.93 (.259)	9.88 (.327)	9.78 (.416)	9.87 (.338)	9.90 (.296)

A statistically significant difference was found in received quiz scores between identified clusters ($H = 65.572$, $p < .001$). The average quiz scores received by *bounded resolvers* were higher than the ones received by *moderate*, *step-by-step*, and *social resolvers* (in all cases $p < .05$). The average quiz scores received by *step-by-step resolvers* were lower than the ones received by *moderate*, *social*, and *self-supporting resolvers* (in all cases $p < .05$). More detailed results about performance by cognitive engagement clusters' members were provided in Publication IV.

5. DISCUSSION

This chapter provides a discussion of findings in the light of the proposed research model. A more detailed discussion of the findings is provided in Publications I–IV. Sub-chapter 5.1. focuses on the participants' background variables. Sub-chapter 5.2. discusses the completers' engagement clusters, which were identified based on their engagement with various course activities and help sources, and describes the clusters in terms of member background variables. Sub-chapter 5.3. considers participants' performance in the MOOC.

5.1. Participants' background variables

The study examined the background variables of participants in three programming MOOCs, and the impact of those variables on enrolment and completion probability. The analysed data included socio-demographic variables (i.e., gender, age, education and employment status) and data about previous experience. The results indicated that some background variables and their impact varied depending on the difficulty level of MOOC.

In both programming courses, "About Programming" and "Introduction to Programming I", which differed in terms of difficulty levels, more than half of the participants were female. Given that females tend to have less previous experience in course subject than males (Duran et al., 2020; Hennis et al., 2016), it seems logical that females enrol in courses that introduce the topic from the basics and that do not require programming experience (as in the cases of AP and IP1). However, the current study results are intriguing because in previous studies males have been in the majority in introductory programming MOOCs (Psathas et al., 2018), and females have been found to be less likely to enrol in STEM and computer science courses (de Souza & Perry, 2021; Jiang et al., 2018). It is a possibility that previous findings reflected the stereotype that the IT-field is not for females and, therefore, they may not dare to participate in programming courses. But it seems likely that this stereotype is starting to break down as the number of female participants in the programming field is slowly increasing (Duran et al., 2020). The high proportion of females who wanted to study programming in the courses examined seems to corroborate this trend. However, the percentage of females decreased as the difficulty level of course increased. The analysis showed that females were more likely to enrol in MOOCs like AP, which was short, easy and had a few tasks, than in any other studied programming course. Consequently, gender distribution among MOOC participants is related to difficulty level of the topic rather than to the course subject itself as has been argued previously (e.g., Alonso-Mencía et al., 2021; de Souza & Perry, 2021; Gil-Jaurena et al., 2017). Meanwhile, the study results indicated that in AP and IP1 gender did not influence completion. So, at least in introductory courses, such as AP and IP1, females can cope with the course as well as their male counterparts.

Not surprisingly, the lack of previous experience becomes one of the important arguments when choosing in which course to enrol. In AP, the easiest course, there were about four times more inexperienced participants than in slightly more difficult IP1. This could be related to the fact that AP was recommended to those who had never studied programming before, while IP1 was for those who had no or little experience in programming (e.g., previously completed AP). Obviously, inexperienced participants try to be sensible in their expectations and enrol in an easier course, like AP, first to get a general idea about programming. Just like in other studies (e.g., Kennedy et al., 2015; Semenova, 2020), experience played an important role in increasing course completion likelihood. But the current study also showed that it remains important for completion regardless of the difficulty level of the course (be it short and easy like AP, or longer and slightly more difficult like IP1). For IP2, previous experience in programming was mandatory.

In comparison to other MOOCs, the AP participants' education level was lower and in the course there tended to be a higher proportion of those whose employment status was marked as "student". This could be due to the fact that the AP participants' average age tended to be lower as well compared to other MOOCs. The other explanation might be related to the difficulty level of AP and the absence of the requirement of previous experience in the course field. Thus, AP could attract students who wanted to dip their toes into programming. Regarding completion, those with lower education level were less likely to complete not only AP but also IP1 (in comparison to master's degree holders). Perhaps, they just wanted to taste what programming was like or they might require more personalised support or course design to pass the course successfully. However, in IP2 (which had the highest difficulty level among the studied MOOCs, did not include helpdesk support and where learners were expected to work more on their own) those with basic education were more likely to complete. It is a novel finding in the MOOC context because in previous studies more educated people have been found to be more likely to complete a course (Morris et al., 2015; Pursel et al., 2016; Semenova, 2020). This rather surprising finding needs further investigation. Meanwhile, in IP2 employed people were less likely to complete the course. But it may be related to the fact that the highest proportion of IP2 completers was found among those with basic education level, who are likely not employed yet, and also the fact that some employed people may not be able to dedicate enough energy to learning in a course that expects more independent work.

5.2. Completers' engagement clusters

In this study, engagement was investigated from completers' perspective who had passed the MOOC "About Programming". Although they completed the course successfully, the results indicated that completers cannot be considered a homogeneous group as they differ in their engagement with course activities (such as learning materials, videos, demos, etc.) and help sources (such as helpdesk, friends, Internet, etc.). This is a novel result in the MOOC context because a more common approach has been to consider completers as one general group without identifying different sub-groups (e.g., Arora et al., 2017; Kang, 2020; Li and Baker, 2016). The study results also indicated that the identified clusters varied in terms of background variables of their members.

Engagement with a limited amount of AP activities as well as resolving a difficulty on their own without hoping for someone's help were more common for males and those experienced in programming. It could be related to the fact that males tend to have more previous experience in the course subject (Duran et al., 2020; Hennis et al., 2016), can be more confident in their own perceived abilities (Luik et al., 2020) and seldom ask other people for help (Nelimarkka & Hellas, 2018). Consequently, such completers could personalise learning by skipping the topics with which they are already familiar. According to the study results, the choice of help source (mostly learning materials alone or also the Internet) among those experienced in programming was to some extent related to their experience with web-based learning. Seeking for clues in both, learning materials and the Internet, was more likely among those AP completers who had never participated in a web-based course before. Thus, previous experience (be it with course subject or web-based learning) plays an important role in how people engage with a course.

Engagement with a large amount of all different types of course activities, including supportive ones, and seeking for help from numerous help sources was more likely among inexperienced in programming completers. The results support the assumption that novices need more different kinds of activities and help sources to understand a topic. Offering such possibilities is important regardless of the course's difficulty level because, as the study results indicated, learners with no experience in course field were less likely to complete both AP and IP1 (for IP2 previous experience with programming was mandatory). Meanwhile, more frequent utilisation of help sources where people answered a question (e.g., forum, helpdesk) was more common among those who had never participated in any web-based course and also among female completers. The latter aligns with the idea that females do not hesitate to ask other people for help (Kizilcec et al., 2017; Nelimarkka & Hellas, 2018).

Active engagement with almost all amount of activities, excluding supportive ones, was more likely among those with higher education level and those who were employed or inactive. This could be related to the fact that the age of completers belonging to this cluster tended to be higher (compared to other clusters). The study results support previous findings that more educated (Williams et al., 2018) and older (Torres & Beier, 2018) learners take part in more MOOC activities but contradict the findings by Kahan et al. (2017) that the majority of working learners engage with fewer activities. It is possible that in the current study completers with the mentioned background variables were interested in personal enrichment or thought about changing their career field and the course offered an opportunity to get introduced to programming.

According to the study results, in the MOOCs there is a cluster of completers who approach the learning process from a pragmatic point of view. They engage mostly with activities where the knowledge acquired provides the most direct route to course completion, but sometimes they can also seek help. Their engagement resembles the approach of *doing as little as possible and as much as necessary*, as they tend to optimise efforts to maximise their results (Arora et al., 2017). Although Arora et al. (2017), who focused on learners, assumed that such optimisation might be related to prior knowledge in the course subject, the current results in the context of solely completers indicated that, within the cluster of pragmatics, there were no distinctive differences neither in terms of previous experience in programming or web-based learning nor in socio-demographic characteristics. Thus, it can be concluded that everyone can optimise learning for oneself, and regardless of background variables it is possible to complete a MOOC successfully with less effort without engaging with all activities.

5.3. Performance in MOOC

In this study, learners' performance in the MOOC "About Programming" was studied through assessments, which were provided in the form of programming tasks and quizzes. The differences in performance were also explored in terms of completers' engagement clusters and background variables of the members of respective clusters. The results indicated that completers do not always achieve better results than non-completers, and completers' engagement patterns influence their performance.

Non-completers of AP made, on average, more attempts per programming task than completers. Non-completers may not understand the logic of programming code and attempt to solve a task by trial and error (Rõõm et al., 2020) or do not focus on finding the source of error (Karavirta, 2006). Therefore, their performance may suffer. However, the study results in terms of quizzes are intriguing. Not surprisingly, completers received, on average, higher quiz scores than non-completers. However, there was no significant difference between them in the average number of attempts per quiz. This rather interesting finding requires further research.

Although completers passed the MOOC successfully, they engaged with course activities and help sources differently, which in turn caused a difference in their performance. A tendency to make, on average, few attempts per assessment combined with high quiz scores was detected among those completers who actively engaged with a course, used a pragmatic approach or resolved difficulties by themselves. Previous studies (e.g., Kahan et al., 2017; Khalil & Ebner, 2017) have also found that those who visit more learning materials, watch more videos, write more comments in forums, etc., have overall better performance. However, those studies were focused on learners and behavioural engagement only. Nevertheless, the results of the current study indicated that the same tendency in performance can be also found among completers who put in lesser effort but focus on activities that are most relevant for completion. The only noteworthy difference, a small but still significant, was found in the average number of attempts per quiz, where pragmatic completers made more attempts than active engagers. In addition, according to the study results, few attempts per assessment and high quiz scores can be detected in those cognitive engagement clusters where members resolve difficulties on their own. Given that among them dominated experienced in programming and AP was an easy introductory course, it is obvious that their previous experience was probably sufficient for such performance.

A tendency to make, on average, many attempts per programming task and to receive low quiz scores was detected among those completers who engaged with numerous support mechanisms. It is quite logical that novices (as behavioural and cognitive engagement with support mechanisms was more common for completers inexperienced in programming) may find the course topic more challenging to understand immediately or may interpret received hints in a wrong way. However, the current study results are contrary to that of Carter et al. (2015) who found that better grades are achieved by those who receive more support.

The performance by members of the cluster in which completers had limited behavioural engagement with activities was intriguing. In terms of quizzes (the number of attempts and received scores) there was no significant difference found between them and support required completers. It is an interesting finding because the cluster of those with limited engagement was the only behavioural engagement cluster where experienced in programming dominated, while support required completers, for the most part, had never studied programming before. It is possible that those with limited engagement relied too much on their previous knowledge and started quizzes without reading all or part of materials properly or even skipping some topics. However, regarding the average number of attempts per programming tasks there was a significant difference found in favour of those with limited engagement. Thus, findings that completers with limited engagement in comparison to support required ones do not differ in mostly theoretical quizzes but make less attempts per programming task supports previous findings (Arora, 2017) that experienced in courses field might be more interested in solving assessments rather than improving their knowledge.

Generally, participants' several background variables were related to the difficulty level of programming MOOCs. Completers applied different behavioural and cognitive engagement patterns, which in turn caused differences in their performance. The study also indicated that completers with different engagement patterns varied in their background variables. The implications of the findings of the current study are presented below.

6. CONCLUSION AND IMPLICATIONS

This chapter presents the conclusions, including methodological and practical implications, limitations and suggestions for future research. In the current study, data from questionnaires and different systems were used to compose an overall picture of participants' background and their performance, as well as completers' behavioural and cognitive engagement. Several novel results in the MOOC context were found as the impact of completers' background variables on engagement, which in turn influences performance, is understudied. The study showed that in MOOCs, where participants have a diversity of backgrounds, engagement can be personalised more effectively if course instructors target certain groups of participants through providing appropriate activities and help sources. This, in turn, will support participants in working towards better quality of knowledge in a style suitable for them and, as a result, achieving good performance in assessment. In addition, the study results could be used for planning future MOOCs in a more cost-effective way.

6.1. Implications

Below are given practical implications to consider when developing a MOOC in a more cost-effective way and providing a personalised learning experience to enhance the quality of knowledge acquired.

- Although all completers have successfully passed the course, they should not be treated as one homogeneous group because they vary in their behavioural and cognitive engagement with the course. Instructors should provide a possibility for personalised engagement with various course activities and help sources to facilitate successful course completion by participants with different backgrounds.
- Course instructors could, before starting a MOOC, conduct a pre-course questionnaire to investigate participants' background and provide personalised course design in the light of the questionnaire results to enhance participants' learning experience. For example, in the case of a course with predominantly female participants, more attention should be paid to offering possibilities for personalised support and feedback from the teaching staff. Or in a course where experienced participants constitute the majority, instructors would be able to expend less effort on developing help sources that are costly (such as helpdesk) or require a lot of work to create (like troubleshooters) because such participants can cope with difficulties on their own.
- According to the current study, helpdesk was used only by a small group of MOOC completers. Given that this is a costly support mechanism and requires quick responses from a teaching staff, course instructors should think carefully about its employment. Troubleshooters and forums may

be a good alternative since the study results indicated their active utilisation by many completers.

- Some completers made numerous resubmissions before a programming task was passed. But MOOC instructors should be less concerned about the number of attempts, because individuals differ in their ability to understand a topic and cope with a difficulty. Allowing an unlimited number of attempts per assessment can help learners complete a MOOC successfully.
- In the studied MOOC “About Programming” an automatic assessment checking system was used to evaluate the concordance of a learner’s programming task solution with the task requirements. System feedback with detailed clues about the type or location of error could help learners better understand the error and, consequently, improve the quality of obtained knowledge.

Next, methodological implications are provided that could facilitate an understanding of the complex nature of behavioural and cognitive engagement.

- Researchers are still looking for ways to measure behavioural (e.g., Barthakur et al., 2021; Kang, 2020; Phan et al., 2016) and cognitive engagement (e.g., Deng et al., 2020b; Liu et al., 2018; Galikyan et al., 2021) in the context of MOOCs. Although data from questionnaires can be biased (e.g., Alonso-Mencía et al., 2021; Kovanović et al., 2019), these data can be used successfully together with data from systems to get a larger picture of how much and how frequently MOOC completers engage with course activities and help sources, including those provided outside the course (e.g., visiting weblinks, searching for materials on the Internet, asking a friend for help). This kind of data could not be obtained otherwise, for example, by using learning analytics or data logs from learning environments. The application of combined data collection methods can help course instructors to be more aware of the ways they can support learners’ engagement, because it enables them to identify the activities and help sources that are in more demand and those that may require improvement.
- The comprehension of the sophisticated nature of cognitive engagement can be improved through various approaches. For example, in the current study, cognitive engagement was observed through the lens of frequency in engaging with various help sources whereas previous research has measured cognitive engagement through video interaction events, like pausing or seeking (Li & Baker, 2018; Liu et al., 2018) or the content of forum posts (Galikyan et al., 2021; Wen et al., 2014b). It seems that utilisation of help sources can, in the light of the current study, be considered by future researchers as an indicator of cognitive engagement in the context of MOOCs.

- Due to using only data about MOOC completers, it was possible to investigate how behavioural and cognitive engagement patterns contributed to completers' performance. Categorising participants based on their engagement with the course activities and help sources can help understand how they actually engage with a MOOC, and facilitate possible improvements in course design and personalise course delivery.

6.2. Limitations and suggestions for future research

The current thesis has some limitations that should be kept in mind when generalising the findings to other MOOCs.

- The study was mainly focused on one MOOC on introductory programming in Python that was a relatively short course. The advantage of such courses is to enable participants of different gender, age, employment status, etc. to make their first steps in programming to decide if this field attracts them or not.
- The identified engagement clusters among completers were based on data from only those completers who answered both voluntary pre- and post-questionnaires. Even though participants were kindly asked to fill out both questionnaires and about three-quarters of completers did so, some differences are possible between this group and all MOOC completers.
- The data on completers' engagement were based on self-reporting. Although this data collection method is often employed to investigate learners' attitudes towards learning, they can under- or overestimate their actual engagement. In addition, evaluation of engagement was made by completers at the end of the course and some bias may occur. Guiding participants to more accurate self-evaluation and doing it on an ongoing basis would be helpful.
- The analysis of learners' performance was based on aggregated data (not separately per each programming task and quiz). This approach helped study performance in a more compact manner since the analysis of nine different clusters across ten assessments would lead to a more complex interpretation of the results. The performance could also be affected by factors such as the wording of assessments or learners being overloaded at work. Completers' performance was studied only in terms of engagement clusters.

The suggestions for future research are as follows:

- The study provides empirical evidence on the existence of various completers' clusters in introductory programming MOOC. Further work is required to examine whether the study results are still valid in MOOCs on other subjects and courses with different levels of difficulty. In addition, future research could continue to explore what behavioural engagement clusters may emerge within each cognitive engagement cluster and vice versa.

- MOOC learners form a heterogeneous community. It would be useful to know which engagement groups can be identified among learners based on their age, education, etc. This information could help instructors when designing a MOOC targeting a specific group. Based on a participant's background variables, the system could automatically suggest corresponding activities and help sources. This could help avoid information overload and reduce confusion about the different links usually provided in a course.
- Learners' engagement can change throughout the course. Monitoring of these dynamics on an ongoing basis would be a fruitful area for research. This would allow instructors to provide early intervention for supporting learners. In time, automatic notification with suggestions could include an overview of how this or that engagement pattern impacts on performance (e.g., how many attempts a learner with a particular behavioural and cognitive engagement pattern usually makes per task, what quiz scores may be achieved) and the likelihood of course completion.
- Future studies could investigate whether and if so, to what extent, differences in results may be related to different data collection methods (e.g., clickstream data vs questionnaire data). Further qualitative research could be undertaken to study in-depth learners' opinions about their engagement and learning in MOOCs. This could be quite beneficial for scientific literature in determining the suited methods for measuring and understanding behavioural and cognitive engagement in the MOOC context.

BIBLIOGRAPHY

Alexandron, G., Yoo, L. Y., Ruipérez-Valiente, J. A., Lee, S., & Pritchard, D. E. (2019). Are MOOC learning analytics results trustworthy? With fake learners, they might not be! *International Journal of Artificial Intelligence in Education*, *29*(4), 484–506. https://doi.org/10.1007/s40593-019-00183-1

Alonso-Mencía, M. E., Alario-Hoyos, C., Estévez-Ayres, I., & Delgado Kloos, C. (2021). Analysing self-regulated learning strategies of MOOC learners through self-reported data. *Australasian Journal of Educational Technology*, *37*(3), 56–70. https://doi.org/10.14742/ajet.6150

Andmekaitse Inspektsioon. (2019). *Isikuandmete töötleja üldjuhend*. Retrieved from https://www.aki.ee/sites/default/files/dokumendid/isikuandmete_tootleja_uldjuhend.pdf

All European Academies. (2017). *The European code of conduct for research integrity. Revised edition.* Retrieved from https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/european-code-of-conduct-for-research-integrity_horizon_en.pdf

Arora, S., Goel, M., Sabitha, A. S., & Mehrotra, D. (2017). Learner groups in massive open online courses. *American Journal of Distance Education*, *31*(2), 80–97. https://doi.org/10.1080/08923647.2017.1300461

Atapattu, T., Thilakaratne, M., Vivian, R., & Falkner, K. (2019). Detecting cognitive engagement using word embeddings within an online teacher professional development community. *Computers & Education*, *140*. https://doi.org/10.1016/j.compedu.2019.05.020

Barak, M., Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education*, *94*, 49–60. https://doi.org/10.1016/j.compedu.2015.11.010

Barthakur, A., Kovanovic, V., Joksimovic, S., Siemens, G., Richey, M., & Dawson, S. (2021). Assessing program-level learning strategies in MOOCs. *Computers in Human Behavior*, *117*. https://doi.org/10.1016/j.chb.2020.106674

Bates, A. W. (2019). *Variations in MOOC designs.* Retrieved from https://opentextbc.ca/teachinginadigitalage/chapter/section-7-4-design-models-for-moocs/

Bonafini, F. C. (2017). The effects of participants' engagement with videos and forums in a MOOC for teachers' professional development. *Open Praxis*, *9*(4), 433–447. http://doi.org/10.5944/openpraxis.9.4.637

Cabedo, R., Tovar, E., Martin, S., Llamas, M., Caeiro, M., Martinez, O., Strachan, R., Castro, M., & Reisman, S. (2018). Who are interested in open education? An analysis of the participants in the first MOOC of the IEEE education society. *Proceedings of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, *Japan*, *2*, 310–314. https://doi.org/10.1109/COMPSAC.2018.10248

Carter, J., Dewan, P., & Pichiliani, M. (2015). Towards incremental separation of surmountable and insurmountable programming difficulties. *Proceedings of the 46th ACM Technical Symposium on Computer Science Education (SIGCSE)*, *USA*, 241–246. https://doi.org/10.1145/2676723.2677294

Chen, C., Sonnert, G., Sadler, P. M., & Malan, D. J. (2020) Computational thinking and assignment resubmission predict persistence in a computer science MOOC. *Journal of Computer Assisted Learning*, *36*(5), 581–594. https://doi.org/10.1111/jcal.12427

Cobos, R., & Ruiz-Garcia, J. C. (2020). Improving learner engagement in MOOCs using a learning intervention system: A research study in engineering education.

Computer Applications in Engineering Education, *29*(4), 733–749. https://doi.org/10.1002/cae.22316

Cole, A. W., Lennon, L., & Weber, N. L. (2021). Student perceptions of online active learning practices and online learning climate predict online course engagement. *Interactive Learning Environments*, *29*(5), 866–880. https://doi.org/10.1080/10494820.2019.1619593

Cristea, A., Alamri, A., Kayama, M., Stewart, C., Alshehri, M., & Shi, L. (2018). Earliest Predictor of Dropout in MOOCs: A Longitudinal Study of FutureLearn courses. In B. Andersson, B. Johansson, S. Carlsson, C. Barry, M. Lang, H. Linger, & C. Schneider (Eds.), *Designing Digitalization (ISD2018 Proceedings).* Lund, Sweden: Lund University. http://aisel.aisnet.org/isd2014/proceedings2018/Education/5

Daniels, L. M., Adams, C., & McCaffrey, A. (2016). Emotional and social engagement in a Massive Open Online Course: An examination of Dino 101. In S. Y. Tettegah, & M. P. McCreery (Eds.), *Emotions, technology, and learning* (pp. 25–41). Academic Press. https://doi.org/10.1016/B978-0-12-800649-8.00004-3

de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning*, *32*(3), 218–231. https://doi.org/10.1111/jcal.12130

de Jong, P. G. M., Hendriks, R. A., Luk, F., Dos Santos Jr, A. C. S., & Reinders, M. E. J. (2021). Development and application of a massive open online course to deliver innovative transplant education. *Transplant Immunology*, *66*. https://doi.org/10.1016/j.trim.2020.101339

de Souza, N. S., & Perry, G. T. (2021). Women's participation in MOOCs in the IT area. *Computers & Education*, *173*, 1–11. https://doi.org/10.1016/j.compedu.2021.104270

Deng, R., Benckendorff, P., & Gannaway, D. (2019). Progress and new directions for teaching and learning in MOOCs. *Computers & Education*, *129*, 48–60. https://doi.org/10.1016/j.compedu.2018.10.019

Deng, R., Benckendorff, P., & Gannaway, D. (2020a). Learner engagement in MOOCs: Scale development and validation. *British Journal of Educational Technology*, *51*(1), 245–262. https://doi.org/10.1111/bjet.12810

Deng, R., Benckendorff, P., & Gannaway, D. (2020b). Linking learner factors, teaching context, and engagement patterns with MOOC learning outcomes. *Journal of Computer Assisted Learning*, *36*(5), 688–708. https://doi.org/10.1111/jcal.12437

Dissanayake, D., Perera, T., Elladeniya, C., Dissanayake, K., Herath, S., & Perera, I. (2018). Identifying the learning style of students in MOOCs using video interactions. *International Journal of Information and Education Technology*, *8*(3), 171–177. http://doi.org/10.18178/ijiet.2018.8.3.1029

Dodson, M. N., Kitburi, K., & Berge, Z. L. (2015). Possibilities for MOOCs in corporate training and development. *Performance Improvement*, *54*(10), 14–21. https://doi.org/10.1002/pfi.21532

Donitsa-Schmidt, S., & Topaz, B. (2018). Massive open online courses as a knowledge base for teachers. Journal of Education for Teaching, 44(5), 608–620. https://doi.org/10.1080/02607476.2018.1516350

Drew, C. (2019). *31 Major Learning Theories in Education, Explained!* Helpful Professor. Retrieved from https://helpfulprofessor.com/learning-theories/

Duran, R., Haaranen, L., & Hellas, A. (2020). Gender differences in introductory programming: Comparing MOOCs and local courses. *Proceedings of the 51st ACM*

Technical Symposium on Computer Science Education (SIGCSE), *USA*, 692–698. https://doi.org/10.1145/3328778.3366852
Feklistova, L., Lepp, M., & Luik, P. (2019). Completers' engagement clusters in programming MOOC: The case of Estonia. In L. Gómez Chova, A. López Martínez, I. Candel Torres (Eds.), *Proceedings of the 12th annual international conference of education, research and innovation (ICERI)* (pp. 1119–1126). IATED. http://doi.org/10.21125/iceri.2019.0341
Feklistova, L., Luik, P., & Lepp, M. (2020). Clusters of programming exercises difficulties resolvers in a MOOC. In C. Busch, M. Steinicke, T. Wendler (Eds.), *European Conference on e-Learning* (pp. 563–569). Academic Conferences International Limited.
Feklistova, L., Lepp, M., & Luik, P. (2021). Learners' performance in a MOOC on programming. *Education Sciences*, *11*(9). https://doi.org/10.3390/educsci11090521
Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). SAGE Publications.
Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59–109.
Galikyan, I., Admiraal, W., & Kester, L. (2021). MOOC discussion forums: The interplay of the cognitive and the social. *Computers & Education*, *165*. https://doi.org/10.1016/j.compedu.2021.104133
Gallego-Romero, J. M., Alario-Hoyos, C., Estévez-Ayres, I., & Delgado Kloos, C. (2020). Analyzing learners' engagement and behavior in MOOCs on programming with the Codeboard IDE. *Educational Technology Research and Development*, *68*(5), 2505–2528. https://doi.org/10.1007/s11423-020-09773-6
Gil-Jaurena, I., Callejo-Gallego, J., & Agudo, Y. (2017). Evaluation of the UNED MOOCs implementation: Demographics, learners' opinions and completion rates. *The International Review of Research in Open and Distributed Learning*, *18*(7), 141–168. https://doi.org/10.19173/irrodl.v18i7.3155
Goldberg, L. R., Bell, E., King, C., O'Mara, C., McInerney, F., Robinson, A., & Vickers, J. (2015). Relationship between participants' level of education and engagement in their completion of the Understanding Dementia Massive Open Online Course. *BMC medical education*, *15*(1), 1–7. https://doi.org/10.1186/s12909-015-0344-z
Gomez-Zermeno, M. G., & de La Garza, L. A. (2016). Research analysis on MOOC course dropout and retention rates. *Turkish online journal of distance education*, *17*(2), 3–14.
Graduate Student Instructor Teaching & Resource Center. (n.d.). *Cognitive Constructivism*. Retrieved from https://gsi.berkeley.edu/gsi-guide-contents/learning-theory-research/cognitive-constructivism/
Greene, J. A., Oswald, C. A., & Pomerantz, J. (2015). Predictors of retention and achievement in a massive open online course. *American Educational Research Journal*, *52*(5), 925–955. https://doi.org/10.3102/0002831215584621
Guajardo Leal, B.E., Valenzuela González, J.R., & Scott, J. (2019). Student engagement as a predictor of xMOOC completion: An analysis from five courses on energy sustainability. *Online Learning*, *23*(2), 105–123. http://doi.org/10.24059/olj.v23i2.1523
Hagan, D., & Markham, S. (2000). Does it help to have some programming experience before beginning a computing degree program? *Proceedings of the 5th annual SIGCSE/SIGCUE ITiCSE conference on Innovation and technology in computer science education*, *Finland*, 25–28. https://doi.org/10.1145/343048.343063

Hennis, T., Topolovec, S., Poquet, O., & de Vries, P. (2016). *Who is the learner: Profiling the engineering MOOC student*. SEFI 44th Annual Conference, Tampere, Finland.

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, *90*, 36–53. https://doi.org/10.1016/j.compedu.2015.09.005

Herrington, J., Oliver, R., & Reeves, T. C. (2003). Patterns of engagement in authentic online learning environments. *Australasian Journal of Educational Technology*, *19*(1). https://doi.org/10.14742/ajet.1701

Hew, K. F. (2014). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology*, *47*(2), 320–341. https://doi.org/10.1111/bjet.12235

Hone, K. S., & El Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers & Education*, *98*, 157–168. https://doi.org/10.1016/j.compedu.2016.03.016

Jiang, S., Schenke, K., Eccles, J. S., Xu, D., & Warschauer, M. (2018). Cross-national comparison of gender differences in the enrollment in and completion of science, technology, engineering, and mathematics Massive Open Online Courses. *PLOS ONE*, *13*(9), 1–15. https://doi.org/10.1371/journal.pone.0202463

Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A. C., & Brooks, C. (2018). How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*, *88*(1), 43-86. https://doi.org/10.3102/0034654317740335

Jung, Y., & Lee, J. (2018). Learning engagement and persistence in massive open online courses (MOOCS). *Computers & Education*, *122*, 9–22. https://doi.org/10.1016/j.compedu.2018.02.013

Kahan, T., Soffer, T., & Nachmias, R. (2017). Types of participant behavior in a massive open online course. T*he International Review of Research in Open and Distributed Learning, 18*(6), 1–18. https://doi.org/10.19173/irrodl.v18i6.3087

Kang, I. G. (2020). Heterogeneity of learners' behavioral patterns of watching videos and completing assessments in massive open online courses (MOOCs): A latent class analysis. *The International Review of Research in Open and Distributed Learning*, *21*(4), 222–237. https://doi.org/10.19173/irrodl.v21i4.4645

Karavirta, V., Korhonen, A., & Malmi, L. (2006). On the use of resubmissions in automatic assessment systems. *Computer science education*, *16*(3), 229–240. https://doi.org/10.1080/08993400600912426

Kearsley, G., & Shneiderman, B. (1998). Engagement theory: A framework for technology-based teaching and learning. Educational Technology, 38(5), 20–23. http://www.jstor.org/stable/44428478

Kennedy, G., Coffrin, C., de Barba, P., & Corrin, L. (2015). Predicting success: how learners' prior knowledge, skills and activities predict MOOC performance. *Proceedings of the fifth international conference on learning analytics and knowledge, USA*, 136–140. http://doi.org/10.1145/2723576.2723593

Kew, S. N., & Tasir, Z. (2021). Analysing students' cognitive engagement in e-learning discussion forums through content analysis. *Knowledge Management & E-Learning: An International Journal*, *13*(1), 39–57. https://doi.org/10.34105/j.kmel.2021.13.003

Khalil, M., Taraghi, B., & Ebner, M. (2016). Engaging learning analytics in MOOCs: The good, the bad, and the ugly. In M. Carmo (Ed.), *Proceedings of International Conference on Education and New Developments Slovenia*, 3–7.

Khalil, M., & Ebner, M. (2017). Clustering patterns of engagement in Massive Open Online Courses (MOOCs): The use of learning analytics to reveal student categories. *Journal of Computing in Higher Education*, *29*(1), 114–132. https://doi.org/10.1007/s12528-016-9126-9

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *Proceedings of the third international conference on learning analytics and knowledge*, *Belgium*, 170–179. https://doi.org/10.1145/2460296.2460330

Kizilcec, R. F., Pérez-Sanagustín, M. & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*, 18–33. https://doi.org/10.1016/j.compedu.2016.10.001

Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., de Vries, P., Hatala, M., Dawson, S., Siemens, G., & Gašević, D. (2019). Examining communities of inquiry in massive open online courses: The role of study strategies. *The Internet and Higher Education*, *40*, 20–43. https://doi.org/10.1016/j.iheduc.2018.09.001

Kuo, T. M., Tsai, C. C., & Wang, J. C. (2021). Linking web-based learning self-efficacy and learning engagement in MOOCs: The role of online academic hardiness. *The Internet and Higher Education*, *51*. https://doi.org/10.1016/j.iheduc.2021.100819

Lan, M., & Hew, K. F. (2020). Examining learning engagement in MOOCs: A self-determination theoretical perspective using mixed method. *International Journal of Educational Technology in Higher Education*, *17*(7), 1–24. https://doi.org/10.1186/s41239-020-0179-5

Lausa, S., Bringula, R., Catacutan-Bangit, A. E., & Santiago, C. (2021). Information-seeking behavior of computing students while programming: Educational learning materials usage, satisfaction of use, and inconveniences. In T. Klinger, C. Kollmitzer, A. Pester (Eds.), *Proceedings of the 2021 IEEE Global Engineering Education Conference (EDUCON),* 761–765. https://doi.org/10.1109/EDUCON46332.2021.9453905

Lawson, M. A., & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. *Review of Educational Research*, *83*(3), 432–479. https://doi.org/10.3102/0034654313480891

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: an international journal*, *46*(3), 517–528. https://doi.org/10.2224/sbp.7054

Lepp, M., Luik, P., Palts, T., Papli, K., Suviste, R., Säde, M., & Tõnisson, E. (2017a). MOOC in programming: A success story. In L. Campbel, R. Hartshorne (Eds.), *Proceedings of the International Conference on e-Learning* (pp. 138–147). Academic Publishing International.

Lepp, M., Luik, P., Palts, T., Papli, K., Suviste, R., Säde, M., Hollo, K., Vaherpuu, V. & Tõnisson, E. (2017b). Self- and automated assessment in programming MOOCs. In Joosten-ten Brinke, D., Laanpere, M. (Eds.), *International Computer Assisted Assessment Conference: Vol. 653. Technology Enhanced Assessment* (pp. 72–85). Springer. https://doi.org/10.1007/978-3-319-57744-9_7

Lepp, M., Palts, T., Luik, P., Papli, K., Suviste, R., Säde, M., Hollo, K., Vaherpuu, V., & Tõnisson, E. (2018). Troubleshooters for tasks of introductory programming MOOCs. *The International Review of Research in Open and Distributed Learning*, *19*(4), 56–75. https://doi.org/10.19173/irrodl.v19i4.3639

Li, Q., & Baker, R. (2016). Understanding Engagement in MOOCs. In T. Barnes, M. Chi, M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining* (pp. 605–606).

Li, Q., & Baker, R. (2018). The different relationships between engagement and outcomes across participant subgroups in massive open online courses. *Computers & Education*, *127*, 41–65. https://doi.org/10.1016/j.compedu.2018.08.005

Liu, M. C., Yu, C. H., Wu, J., Liu, A. C., & Chen, H. M. (2018). Applying learning analytics to deconstruct user engagement by using log data of MOOCs. *Journal of Information Science & Engineering*, *34*(5), 1175–1186. https://doi.org/10.6688/JISE.201809_34(5).0004

Luik, P., Feklistova, L., Lepp, M., Tõnisson, E., Suviste, R., Gaiduk, M., Säde, M., & Palts, T. (2019). Participants and completers in programming MOOCs. *Education and Information Technologies*, *24*(6), 3689–3706. https://doi.org/10.1007/s10639-019-09954-8

Luik, P., Lepp, M., Feklistova, L., Säde, M., Rõõm, M., Palts, T., Suviste, R., & Tõnisson, E. (2020) Programming MOOCs – different learners and different motivation. *International Journal of Lifelong Education*, *39*(3), 305–318. https://doi.org/10.1080/02601370.2020.1780329

Manwaring, K. C., Larsen, R., Graham, C. R., Henrie, C. R., & Halverson, L. R. (2017). Investigating student engagement in blended learning settings using experience sampling and structural equation modeling. *The Internet and Higher Education*, *35*, 21–33. https://doi.org/10.1016/j.iheduc.2017.06.002

Morris, N. P., Swinnerton, B. J., & Hotchkiss, S. (2015). Can demographic information predict MOOC learner outcomes? *Proceedings of the European MOOC stakeholder summit 2015,* 199–207.

Morrison, M., & Newman, T. S. (2001). A study of the impact of student background and preparedness on outcomes in CS I. *ACM SIGCSE Bulletin*, *33*(1), 179–183. https://doi.org/10.1145/366413.364580

Mubarak, A. A., Cao, H., Zhang, W., & Zhang, W. (2021). Visual analytics of video-clickstream data and prediction of learners' performance using deep learning models in MOOCs' courses. *Computer Applications in Engineering Education*, *29*(4), 710-732. http://doi.org/10.1002/cae.22328

Nelimarkka, M., & Hellas, A. (2018). Social help-seeking strategies in a programming MOOC. In T. Barnes, D. D. Garcia, E. K. Hawthorne, M. A. Pérez-Quiñones (Eds.), *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 116–121. https://doi.org/10.1145/3159450.3159495

Overstreet, M. (2017). Principled/Digital: Composition's "Ethics of Attunement" and the Writing MOOC. In E. Monske, & K. Blair (Eds.), *Handbook of Research on Writing and Composing in the Age of MOOCs* (pp. 75–86). IGI Global. https://doi.org/10.4018/978-1-5225-1718-4.ch005

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

Phan, T., McNeil, S. G., & Robin, B. R. (2016). Students' patterns of engagement and course performance in a massive open online course. *Computers & Education*, *95*, 36–44. https://doi.org/10.1016/j.compedu.2015.11.015

Pietarinen, J., Soini, T., & Pyhältö, K. (2014). Students' emotional and cognitive engagement as the determinants of well-being and achievement in school. *Inter-*

national Journal of Educational Research, *67*, 40–51. https://doi.org/10.1016/j.ijer.2014.05.001
Poellhuber, B., Roy, N., & Bouchoucha, I. (2019). Understanding participant's behaviour in massively open online courses. *The International Review of Research in Open and Distributed Learning*, *20*(1), 221–242. https://doi.org/10.19173/irrodl.v20i1.3709
Pradubwate, R., Pheeraphan, N., Sirawong, N., & Trirat, N. (2020). Characteristics and learning behavior of active learners on SWU-MOOC. *Proceedings of the 2020 11th International Conference on E-Education, E-Business, E-Management, and E-Learning*, 158–162. https://doi.org/10.1145/3377571.3377603
Psathas, G., Chalki, P., Demetriadis, S., & Tsiara, A. (2018). Profiles and motivations of participants in Greek MOOC for Python programming. *Proceedings of the 2018 Learning With MOOCS (LWMOOCS)*, 70–73. https://doi.org/10.1109/LWMOOCS.2018.8534636
Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2016). Understanding MOOC students: Motivations and behaviours indicative of MOOC completion. *Journal of Computer Assisted Learning*, *32*(3), 202–217. https://doi.org/10.1111/jcal.12131
Ramesh, A., Goldwasser, D., Huang, B., Daumé III, H., & Getoor, L. (2013). Modeling learner engagement in MOOCs using probabilistic soft logic. In *NIPS workshop on data driven education*, *USA*, *2*.
Rizvi, S., Rienties, B., Rogaten, J., & Kizilcec, R. F. (2020). Investigating variation in learning processes in a FutureLearn MOOC. *Journal of computing in higher education*, *32*(1), 162–181. https://doi.org/10.1007/s12528-019-09231-0
Rodríguez-del-Pino, J. C., Rubio Royo, E., & Hernández Figueroa, Z. (2012). A virtual programming lab for Moodle with automatic assessment and anti-plagiarism features. In H. R. Arabnia, *Proceedings of the 2012 International Conference on e-Learning, e-Business, Enterprise Information Systems, & e-Government*.
Rõõm, M., Luik, P., Lepp, M. (2020). Learners' use of time in MOOCs about programming. In L. Gómez Chova, A. López Martínez, I. Candel Torres (Eds.), *Proceedings of the 12th International Conference on Education and New Learning Technologies* (pp. 4380–4387). IATED. http://doi.org/10.21125/edulearn.2020.1162
Samoilova, E., Keusch, F., & Kreuter, F. (2018). Integrating survey and learning analytics data for a better understanding of engagement in MOOCs. In H. Jiao, R. W. Lissitz, A. V. Wie (Eds.), *Data analytics and psychometrics: informing assessment practices* (pp. 247–261). Information Age Publishing.
Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach* (7th ed.). John Wiley & Sons.
Semenova, T. (2020). The role of learners' motivation in MOOC completion. *Open Learning: The Journal of Open, Distance and e-Learning*, 1–15. https://doi.org/10.1080/02680513.2020.1766434
Semenova, T. V., & Rudakova, L. M. (2016). Barriers to taking massive open online courses (MOOCs). *Russian Education & Society*, *58*(3), 228–245. https://doi.org/10.1080/10609393.2016.1242992
Shah, D. (2021). *By The Numbers: MOOCs in 2021*. Retrieved from https://www.classcentral.com/report/mooc-stats-2021
Shah, D. (2022). *The 250 Most Popular Online Courses of All Time*. Retrieved from https://www.classcentral.com/report/most-popular-online-courses

Soffer, T., & Cohen, A. (2019). Students' engagement characteristics predict success and completion of online courses. *Journal of computer assisted learning*, *35*(3), 378–389. https://doi.org/10.1111/jcal.12340

Sun, Y., Ni, L., Zhao, Y., Shen, X. L., & Wang, N. (2019). Understanding students' engagement in MOOCs: An integration of self-determination theory and theory of relationship quality. *British Journal of Educational Technology*, *50*(6), 3156–3174. https://doi.org/10.1111/bjet.12724

Teng, Y., & Wang, X. (2021). The effect of two educational technology tools on student engagement in Chinese EFL courses. *International Journal of Educational Technology in Higher Education*, *18*(1), 1–15. https://doi.org/10.1186/s41239-021-00263-0

Tseng, S. F., Tsao, Y. W., Yu, L. C., Chan, C. L., & Lai, K. R. (2016). Who will pass? Analyzing learner behaviors in MOOCs. *Research and practice in technology enhanced learning*, *11*(1), 1–11. https://doi.org/10.1186/s41039-016-0033-5

Torres, W. J., & Beier, M. E. (2018). Adult development in the wild: The determinants of autonomous learning in a massive open online course. *Learning and Individual Differences*, *65*, 207–217. https://doi.org/10.1016/j.lindif.2018.06.003

van de Oudeweetering, K., & Agirdag, O. (2018). Demographic data of MOOC learners: Can alternative survey deliveries improve current understandings? *Computers & Education*, *122*, 169–178. https://doi.org/10.1016/j.compedu.2018.03.017

Walji, S., Deacon, A., Small, J., & Czerniewicz, L. (2016). Learning through engagement: MOOCs as an emergent form of provision. *Distance Education*, *37*(2), 208–223. https://doi.org/10.1080/01587919.2016.1184400

Wang, L., & Sui, J. (2020). Correlation analysis between emotional engagement and achievement of university students in online learning based on an elective course. In Z. Pan, A. Cheok, W. Müller, M. Zhang (Eds.), *Lecture Notes in Computer Science: Vol. 11782. Transactions on Edutainment XVI* (pp. 13–24). Springer. https://doi.org/10.1007/978-3-662-61510-2_2

Wang, M. T., Willett, J. B., & Eccles, J. S. (2011). The assessment of school engagement: Examining dimensionality and measurement invariance by gender and race/ethnicity. *Journal of School Psychology*, *49*(4), 465–480. https://doi.org/10.1016/j.jsp.2011.04.001

Watt, H. M., Carmichael, C., & Callingham, R. (2017). Students' engagement profiles in mathematics according to learning environment dimensions: Developing an evidence base for best practice in mathematics education. *School Psychology International*, *38*(2), 166–183. https://doi.org/10.1177/0143034316688373

Watted, A., & Barak, M. (2018). Motivating factors of MOOC completers: Comparing between university-affiliated students and general participants. *The Internet and Higher Education*, *37*, 11–20. https://doi.org/10.1016/j.iheduc.2017.12.001

Wen, M., & Rosé, C. P. (2014a). Identifying latent study habits by mining learner behavior patterns in massive open online courses. *Proceedings of the 23rd ACM international conference on information and knowledge management*, *China*, 1983–1986. http://doi.org/10.1145/2661829.2662033

Wen, M., Yang, D., & Rosé, C. (2014b). Linguistic reflections of student engagement in massive open online courses. *Proceedings of the 8th International AAAI Conference on Web and Social Media*, *USA*, *8*(1), 525–534.

Whitmer, J., Schiorring, E., James, P., & Miley, S. (2015). How students engage with a remedial English writing MOOC: A case study in learning analytics with big data. *EDUCAUSE Learning Initiative (ELI)*, 1–6. Retrieved from https://library.educause.edu/~/media/files/library/2015/3/elib1502-pdf.pdf

Williams, K. M., Stafford, R. E., Corliss, S. B., & Reilly, E. D. (2018). Examining student characteristics, goals, and engagement in Massive Open Online Courses. *Computers & Education*, *126*, 433–442. https://doi.org/10.1016/j.compedu.2018.08.014

Wysocki, C. D. (2007). *A study of cognitive engagement in online learning* [Unpublished PhD dissertation]. Washington State University, USA. Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.428.3312&rep=rep1&type=pdf

Zafras, I., Kostas, A., & Sofos, A. (2020). MOOCs & participation inequalities in distance education: A systematic literature review 2009-2019. *European Journal of Open Education and E-learning Studies*, *5*(1), 68–89. http://doi.org/10.46827/ejoe.v5i1.3260

World Economic Forum. (2020). *The future of jobs report 2020*. Retrieved from http://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf

Appendix A. POST-QUESTIONNAIRE BLOCK ABOUT BEHAVIOURAL ENGAGEMENT

How would you describe your actions at the time when you were learning to program?

Please choose the option (all, about ¾, about half, about ¼, none) that best fills the gap in your case.

	all	**about ¾**	**about half**	**about ¼**	**none**
Read … of the programming learning materials					
Read … of the additional materials					
Read … of the stories which illustrate programming in a belletristic way					
Visited … of the provided weblinks					
Watched … of the programming learning videos					
Watched … of the weekly encouraging videos					
Tried … of the provided demos					
Solved … of the mandatory assessments using help from troubleshooters					
Answered … of the self-assessment questions					
Read … of the forum posts					
Read … of the weekly encouraging emails					

Appendix B. POST-QUESTIONNAIRE BLOCK ABOUT COGNITIVE ENGAGEMENT

What did you usually do when you encountered a difficulty while solving a programming task?

Please select the most appropriate answer.

	Always	**Most of the time**	**Some-times**	**On a few occasions**	**Never**
Reread learning materials one more time					
Searched for additional materials on the Internet					
Asked a close person or a friend for help					
Used troubleshooters					
Read/write to forum					
Used FAQ					
Wrote to helpdesk					

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, Associate Professors Piret Luik and Marina Lepp, for guiding me throughout my doctoral study. Their patience in replying to my questions and support encouraged me to complete the thesis. I would also like to thank opponents Valentina Dagienė, Tobias Ley and Elena Sügis for their valuable feedback.

I really appreciate the great support from my family, friends and colleagues that I was getting during my study.

The degree study has been financially supported by the Institute of Computer Science at the University of Tartu and IT Academy.

SUMMARY IN ESTONIAN

Sissejuhatava programmeerimise MOOCi õppijad: taustamuutujad, kaasatuse mustrid ja õpisooritus

Tänapäeval täiendavad inimesed pidevalt enda teadmisi ja oskusi selleks, et olla edukad. Üks võimalus personaalseks ja professionaalseks arenguks on osalemine vaba juurdepääsuga e-kursusel (ingl *massive open online course*, MOOC). MOOCid annavad õppijale võimaluse kursusel osaleda internetiühenduse olemasolul õppijale sobival ajal ning sõltumata tema asukohast. Arvestades suurt osalejate hulka ja nende erinevat tausta, on kõikide õppijate kaasatus (ingl *engagement*) õppeprotsessis MOOCide korraldajatele väljakutseks. Paremate kursuste korraldamiseks ja õppijate kaasatuseks õppeprotsessis on oluline aru saada mitte ainult sellest, kes kursusel õpivad ja kes selle lõpetavad, vaid ka sellest, milliseid kaasatuse mustreid leidub lõpetajate seas ja milline on mustrite seos õpisooritusega.

Aasta 2012 kuulutati MOOCi aastaks (Pappano, 2012). MOOCide uurijad analüüsivad jätkuvalt õppijate sotsiaal-demograafilisi näitajaid, õpisooritust, kaasatust, jne. (nt. Alonso-Mencía et al., 2021; de Jong et al., 2021; de Souza & Perry, 2021; Deng et al., 2019). Uurijad on arvamusel, et õppijate taust muutub läbi aja ja seetõttu on uued uuringud vajalikud (Duran et al., 2020). Kuigi on leitud, et mida rohkem õppija on õppeprotsessis kaasatud, seda suurema tõenäosusega saab ta MOOCil hakkama (Goldberg et al., 2015), jääb õppijate kaasatus ikkagi ebaselgeks (Deng et al., 2019; Kuo et al., 2021; Lan & Hew, 2020). Kaasatuse uurimine võimaldab aru saada, miks MOOCidel tekivad lõpetajate ja mittelõpetajate rühmad (Pursel et al., 2016; Sun et al., 2019). Kaasatus on mitmedimensionaalne nähtus (Deng et al., 2020a; Lan & Hew, 2020). Käitumuslik kaasatus (ingl *behavioural engagement*) on seotud parema edukusega (Deng et al., 2019) ja eelnevalt on teadlased tuvastanud erinevaid MOOCidel õppijate rühmi, lähtudes käitumuslikust kaasatusest (Arora et al., 2017; Khalil & Ebner, 2017; Kizilcec et al., 2013). Käesolevas doktoritöös on käitumuslik kaasatus käsitletud MOOCi lõpetaja seisukohalt ja on defineeritud kui tegevuste hulk, mida lõpetaja kursuse jooksul tegi. Kognitiivne kaasatus (ingl *cognitive engagement*) on väga oluline *online* õppimisel (Kuo et al., 2021) ning kriitiline teadmiste omandamisel (Li & Baker, 2018) ja soovitud tulemuste saavutamisel (Li & Baker, 2016). Vaatamata selle olulisusele on kognitiivset kaasatust vähe uuritud (Deng et al., 2019; Kuo et al., 2021). Käesolevas doktoritöös on kognitiivset kaasatust vaadeldud MOOCi lõpetaja seisukohalt ja on defineeritud kui erinevate abiallikatega kaasatussagedus, mida lõpetaja kasutas, juhul kui tal tekkisid probleemid programmeerimisülesande lahendamisel. Selleks, et MOOCide korraldajad saaksid paremini aru, kuidas just lõpetajad tegutsesid eduka kursuse läbimise nimel, on kasulik uurida seda õppijate rühma. Doktoritöö autor ei leidnud teaduslikest andmebaasidest artikleid, mis käsitleksid kaasatust lõpetajate seisukohalt.

Doktoritöö eesmärk oli uurida osalejate taustamuutujaid ja nende mõju kursusele registreerumisele ning lõpetamise tõenäosusele, tuvastada lõpetajate seas käitumuslikke ja kognitiivseid kaasatuse rühmasid ning uurida neid taustamuutujate ja õpisoorituse osas. Taustamuutujad (ingl *background variables*) hõlmasid endas sotsiaal-demograafiliste näitajaid (st. vanus, sugu, haridustase, tööhõive staatus) ning eelnevat kogemust programmeerimise õppimise ja veebipõhise õppega. Uurimuse fookuses oli MOOC "Programmeerimisest maalähedaselt". Selle kursuse töötas välja ja korraldas Tartu Ülikooli arvutiteaduse instituudi informaatika didaktika töörühm. Alguses uuriti doktoritöös MOOCi osalejate ja lõpetajate taustamuutujaid (Publikatsioon I). Andmed koguti MOOCile registreerimise vormi ja küsitluse abil. Uurimuse käigus selgus, et võrreldes kahe teise programmeerimise MOOCiga (mida pakkus sama töörühm), oli kursusel "Programmeerimisest maalähedaselt" statistiliselt olulisel määral rohkem naisi ja neid, kelle haridustase oli madalam. Lõpetajate osas selgus, et uuritud MOOCil, sarnaselt MOOCiga "Programmeerimise alused I", ei olnud statistiliselt olulist erinevust nais- ja meeslõpetajate osakaalu ning erinevate tööhõive staatuste vahel. Suurem lõpetajate osakaal oli magistrikraadiga lõpetajate hulgas. Väiksem lõpetajate osakaal oli nende õppijate puhul, kes ei ole varem programmeerimist õppinud.

Edasi võrreldi MOOCi "Programmeerimisest maalähedaselt" mittelõpetajate ja lõpetajate õpisooritust (Publikatsioon IV). Selleks kasutati programmeerimisülesannete lahenduste esitamise ja testide sooritamise katsete arve ning testide tulemusi. Uurimuse käigus selgus, et mittelõpetajad ja lõpetajad vajasid testi sooritamiseks keskmiselt sama palju katseid. Mittelõpetajatel oli programmeerimisülesannete lahenduste esitamiskordade arv suurem ja neil olid testipunktid madalamad.

Järgmisena uuriti lõpetajate käitumuslikku (Publikatsioon II) ja kognitiivset (Publikatsioon III) kaasatust ning nende seoseid lõpetajate taustamuutujate (Publikatsioonid II ja III) ja õpisooritusega (Publikatsioon IV). Tulemused näitasid, et lõpetajad ei ole üks homogeenne rühm. Käitumusliku kaasatuse puhul eristati lõpetajate seas nelja teadmiste kogujate (ingl *knowledge collectors*) rühma, kes kursuse jooksul olid erinevalt kaasatud:

- *aktiivsed teadmiste kogujad* osalesid peaaegu kõikides tegevustes, välja arvatud foorumipostituste lugemine ja murelahendajate kasutamine ülesannete lahendamisel. Võrreldes kõikide teiste klastritega lugesid nad rohkem lisamaterjale (so. silmaringi materjalid ja jutustused programmeerimisest) ja iganädalasi e-kirju. Selles rühmas domineerisid kõrgharidusega õppijad. Testide sooritamisel tegid nad keskmiselt vähem katseid võrreldes teiste rühmadega;
- *minimaalsete teadmiste kogujate* jaoks olid peamisteks teadmiste saamise allikateks tekstilised põhimaterjalid, enesekontrolli testid ja materjalides toodud näited. Võrreldes kõikide teiste rühmadega lugesid nad vähem põhi- ja lisamaterjale, vaatasid vähem videoid ning külastasid ka vähem

veebilinke. Selles rühmas domineerisid meessoost ja kesk- või põhiharidusega lõpetajad, kes olid varem programmeerimist õppinud;

- *pragmaatilised teadmiste kogujad* keskendusid nendele tegevustele, mis olid vajalikud uute teadmiste omandamiseks. Võrreldes *aktiivsete teadmiste kogujatega* lugesid nad vähem lisamaterjale ning olid vähem huvitatud foorumipostituste ja iganädalaste e-kirjade lugemisest. Samas kasutasid nad aga murelahendajaid võrreldes *aktiivsete teadmiste kogujatega* palju rohkem. Taustamuutujate osas tähenduslikke erinevusi ei leitud. Pragmaatikute õpisooritus ei erinenud *aktiivsete teadmiste kogujate* tulemustest, v.a testide sooritamise katsete arv;
- *toetust vajavad teadmiste kogujad* kasutasid kõiki võimalusi, mida kursusel pakuti. Võrreldes teiste rühmadega kasutasid nad palju erinevaid abiallikaid. Selles rühmas domineerisid õppijad, kes ei ole varem programmeerimist õppinud. Suurim katsete arv (121) programmeerimisülesande lahenduse esitamisel kuulus selle rühma liikmele. Võrreldes kõikide teiste rühmadega esitasid nad programmeerimisülesande lahendusi keskmiselt rohkem arv kordi.

Uurimuse tulemused näitasid, et MOOCil võivad olla lõpetajad, kes teevad kõiki tegevusi, aga ka need, kes kursuse jooksul teevad vaid mõnda tegevust.

Kognitiivse kaasatuse puhul eristati lõpetajate seast viis lahendajate (ingl *resolvers*) rühma, kes uuritud MOOCi läbimisel olid erinevalt kaasatud:

- *piiritletud lahendajad* lugesid probleemi korral tihti korraldajate koostatud õppematerjalid uuesti läbi. Võrreldes kõikide teiste rühmadega otsisid nad harva abimaterjale internetist. Selles rühmas domineerisid need, kes on varem programmeerimist õppinud. Neil oli võrreldes teistega keskmiselt kõige väiksem programmeerimisülesannete lahenduste esitamise katsete arv;
- *mõõdukad lahendajad* lugesid tihti õppematerjalid uuesti läbi, otsisid infot internetist või foorumist ja kasutasid murelahendajaid. Selles rühmas domineerisid õppijad, kelle jaoks see kursus oli esimene veebipõhine kursus;
- *üksikasjalikud lahendajad* eelistasid ressursse, mis sisaldasid detailseid juhiseid. Nendeks olid korraldajate koostatud õppematerjalid ja murelahendajad. Viimaste kasutamise sagedus oli *üksikasjalikel lahendajatel* kõrgem võrreldes enamike teiste rühmadega. Suurem osa sellesse rühmasse kuuluvatest lõpetajatest ei olnud varem programmeerimist õppinud. Suurim katsete arv (121) programmeerimisülesande lahenduse esitamisel kuulus selle rühma liikmele;
- *sotsiaalsed lahendajad* olid väga aktiivsed abi otsimisel. Võrreldes kõikide teiste rühmadega lugesid nad sagedamini õppematerjale, kasutasid murelahendajaid, otsisid abi teistelt foorumites ja kirjutasid abiliinile. Selles rühmas domineerisid naised ning õppijad, kes ei ole varem

programmeerimist õppinud ja kelle jaoks see oli esimene veebipõhise õppimise kogemus;

- *enesetoetavad lahendajad* otsisid võrreldes kõikide teiste rühmadega kõige sagedamini abimaterjale internetist. Õppematerjalide läbilugemine oli samuti neile abiks probleemi lahendamisel. Selles rühmas domineerisid mehed ja need, kes osalesid veebipõhisel kursusel esimest korda, aga on varem programmeerimist õppinud.

Tulemused näitasid, et lõpetajate erinevat sagedust erinevate abiallikate kasutamisel võib pidada märgiks püsivast soovist MOOC edukalt läbida. Samuti selgus, et abiallikate kasutamise võib võtta aluseks kognitiivse kaasatuse tuvastamiseks ja mõõtmiseks.

Selles doktoritöös esitatud tulemused võivad täiendada teaduskirjandust, et paremini aru saada MOOCi fenomenist, kus õppijatel on võrreldes traditsiooniliste klassiruumides toimuvate tundidega suurem autonoomia. Uurimistulemustest võib järeldada, et MOOCide korraldajad peavad pakkuma erinevaid tegevusi ja abiallikaid, mis oleksid suunatud konkreetsetele sihtrühmadele. See võib hõlbustada personaliseeritud õppimist ja õppijate tõhusat kaasatust õppeprotsessis.

PUBLICATIONS

ELULOOKIRJELDUS

Nimi:	Lidia Feklistova
Sünniaeg:	17. juuli 1984
Kodakondsus:	Eesti
Kontakt:	lidia.feklistova@ut.ee

Haridus

2018–(2022)	Tartu Ülikool, Arvutiteaduse instituut, PhD informaatika erialal
2005–2007	Tartu Ülikool, tehnikateaduse magistri kraad (infotehnoloogia)
2005–2007	Tartu Ülikool, loodusteaduse magistri kraad (informaatikaõpetaja)
2002–2005	Tartu Ülikool, tehnikateaduse bakalaureuse kraad (infotehnoloogia)
1990–2002	Tartu Annelinna gümnaasium

Teenistuskäik

2012–…	Tartu Ülikool, Arvutiteaduse instituut, informaatika nooremlektor
2007–2016	Tartu Rakenduslik kolledž, arvutiõpetaja

Teadustegevus

Peamised uurimisvaldkonnad:

- Arvuti kasutamine hariduses
- Informaatika didaktika

DISSERTATIONES INFORMATICAE PREVIOUSLY PUBLISHED IN DISSERTATIONES MATHEMATICAE UNIVERSITATIS TARTUENSIS

19. **Helger Lipmaa.** Secure and efficient time-stamping systems. Tartu, 1999, 56 p.
22. **Kaili Müürisep.** Eesti keele arvutigrammatika: süntaks. Tartu, 2000, 107 lk.
23. **Varmo Vene.** Categorical programming with inductive and coinductive types. Tartu, 2000, 116 p.
24. **Olga Sokratova.** Ω-rings, their flat and projective acts with some applications. Tartu, 2000, 120 p.
27. **Tiina Puolakainen.** Eesti keele arvutigrammatika: morfoloogiline ühestamine. Tartu, 2001, 138 lk.
29. **Jan Villemson.** Size-efficient interval time stamps. Tartu, 2002, 82 p.
45. **Kristo Heero.** Path planning and learning strategies for mobile robots in dynamic partially unknown environments. Tartu 2006, 123 p.
49. **Härmel Nestra.** Iteratively defined transfinite trace semantics and program slicing with respect to them. Tartu 2006, 116 p.
53. **Marina Issakova.** Solving of linear equations, linear inequalities and systems of linear equations in interactive learning environment. Tartu 2007, 170 p.
55. **Kaarel Kaljurand.** Attempto controlled English as a Semantic Web language. Tartu 2007, 162 p.
56. **Mart Anton.** Mechanical modeling of IPMC actuators at large deformations. Tartu 2008, 123 p.
59. **Reimo Palm.** Numerical Comparison of Regularization Algorithms for Solving Ill-Posed Problems. Tartu 2010, 105 p.
61. **Jüri Reimand.** Functional analysis of gene lists, networks and regulatory systems. Tartu 2010, 153 p.
62. **Ahti Peder.** Superpositional Graphs and Finding the Description of Structure by Counting Method. Tartu 2010, 87 p.
64. **Vesal Vojdani.** Static Data Race Analysis of Heap-Manipulating C Programs. Tartu 2010, 137 p.
66. **Mark Fišel.** Optimizing Statistical Machine Translation via Input Modification. Tartu 2011, 104 p.
67. **Margus Niitsoo**. Black-box Oracle Separation Techniques with Applications in Time-stamping. Tartu 2011, 174 p.
71. **Siim Karus.** Maintainability of XML Transformations. Tartu 2011, 142 p.
72. **Margus Treumuth.** A Framework for Asynchronous Dialogue Systems: Concepts, Issues and Design Aspects. Tartu 2011, 95 p.
73. **Dmitri Lepp.** Solving simplification problems in the domain of exponents, monomials and polynomials in interactive learning environment T-algebra. Tartu 2011, 202 p.

74. **Meelis Kull.** Statistical enrichment analysis in algorithms for studying gene regulation. Tartu 2011, 151 p.
77. **Bingsheng Zhang.** Efficient cryptographic protocols for secure and private remote databases. Tartu 2011, 206 p.
78. **Reina Uba.** Merging business process models. Tartu 2011, 166 p.
79. **Uuno Puus.** Structural performance as a success factor in software development projects – Estonian experience. Tartu 2012, 106 p.
81. **Georg Singer.** Web search engines and complex information needs. Tartu 2012, 218 p.
83. **Dan Bogdanov.** Sharemind: programmable secure computations with practical applications. Tartu 2013, 191 p.
84. **Jevgeni Kabanov.** Towards a more productive Java EE ecosystem. Tartu 2013, 151 p.
87. **Margus Freudenthal.** Simpl: A toolkit for Domain-Specific Language development in enterprise information systems. Tartu, 2013, 151 p.
90. **Raivo Kolde.** Methods for re-using public gene expression data. Tartu, 2014, 121 p.
91. **Vladimir Šor.** Statistical Approach for Memory Leak Detection in Java Applications. Tartu, 2014, 155 p.
92. **Naved Ahmed.** Deriving Security Requirements from Business Process Models. Tartu, 2014, 171 p.
94. **Liina Kamm.** Privacy-preserving statistical analysis using secure multiparty computation. Tartu, 2015, 201 p.
100. **Abel Armas Cervantes.** Diagnosing Behavioral Differences between Business Process Models. Tartu, 2015, 193 p.
101. **Fredrik Milani.** On Sub-Processes, Process Variation and their Interplay: An Integrated Divide-and-Conquer Method for Modeling Business Processes with Variation. Tartu, 2015, 164 p.
102. **Huber Raul Flores Macario.** Service-Oriented and Evidence-aware Mobile Cloud Computing. Tartu, 2015, 163 p.
103. **Tauno Metsalu.** Statistical analysis of multivariate data in bioinformatics. Tartu, 2016, 197 p.
104. **Riivo Talviste.** Applying Secure Multi-party Computation in Practice. Tartu, 2016, 144 p.
108. **Siim Orasmaa.** Explorations of the Problem of Broad-coverage and General Domain Event Analysis: The Estonian Experience. Tartu, 2016, 186 p.
109. **Prastudy Mungkas Fauzi.** Efficient Non-interactive Zero-knowledge Protocols in the CRS Model. Tartu, 2017, 193 p.
110. **Pelle Jakovits.** Adapting Scientific Computing Algorithms to Distributed Computing Frameworks. Tartu, 2017, 168 p.
111. **Anna Leontjeva.** Using Generative Models to Combine Static and Sequential Features for Classification. Tartu, 2017, 167 p.
112. **Mozhgan Pourmoradnasseri.** Some Problems Related to Extensions of Polytopes. Tartu, 2017, 168 p.

113. **Jaak Randmets.** Programming Languages for Secure Multi-party Computation Application Development. Tartu, 2017, 172 p.
114. **Alisa Pankova.** Efficient Multiparty Computation Secure against Covert and Active Adversaries. Tartu, 2017, 316 p.
116. **Toomas Saarsen.** On the Structure and Use of Process Models and Their Interplay. Tartu, 2017, 123 p.
121. **Kristjan Korjus.** Analyzing EEG Data and Improving Data Partitioning for Machine Learning Algorithms. Tartu, 2017, 106 p.
122. **Eno Tõnisson.** Differences between Expected Answers and the Answers Offered by Computer Algebra Systems to School Mathematics Equations. Tartu, 2017, 195 p.

DISSERTATIONES INFORMATICAE UNIVERSITATIS TARTUENSIS

1. **Abdullah Makkeh**. Applications of Optimization in Some Complex Systems. Tartu 2018, 179 p.
2. **Riivo Kikas**. Analysis of Issue and Dependency Management in Open-Source Software Projects. Tartu 2018, 115 p.
3. **Ehsan Ebrahimi**. Post-Quantum Security in the Presence of Superposition Queries. Tartu 2018, 200 p.
4. **Ilya Verenich**. Explainable Predictive Monitoring of Temporal Measures of Business Processes. Tartu 2019, 151 p.
5. **Yauhen Yakimenka**. Failure Structures of Message-Passing Algorithms in Erasure Decoding and Compressed Sensing. Tartu 2019, 134 p.
6. **Irene Teinemaa**. Predictive and Prescriptive Monitoring of Business Process Outcomes. Tartu 2019, 196 p.
7. **Mohan Liyanage.** A Framework for Mobile Web of Things. Tartu 2019, 131 p.
8. **Toomas Krips.** Improving performance of secure real-number operations. Tartu 2019, 146 p.
9. **Vijayachitra Modhukur.** Profiling of DNA methylation patterns as biomarkers of human disease. Tartu 2019, 134 p.
10. **Elena Sügis.** Integration Methods for Heterogeneous Biological Data. Tartu 2019, 250 p.
11. **Tõnis Tasa.** Bioinformatics Approaches in Personalised Pharmacotherapy. Tartu 2019, 150 p.
12. **Sulev Reisberg.** Developing Computational Solutions for Personalized Medicine. Tartu 2019, 126 p.
13. **Huishi Yin.** Using a Kano-like Model to Facilitate Open Innovation in Requirements Engineering. Tartu 2019, 129 p.
14. **Faiz Ali Shah.** Extracting Information from App Reviews to Facilitate Software Development Activities. Tartu 2020, 149 p.
15. **Adriano Augusto**. Accurate and Efficient Discovery of Process Models from Event Logs. Tartu 2020, 194 p.
16. **Karim Baghery.** Reducing Trust and Improving Security in zk-SNARKs and Commitments. Tartu 2020, 245 p.
17. **Behzad Abdolmaleki.** On Succinct Non-Interactive Zero-Knowledge Protocols Under Weaker Trust Assumptions. Tartu 2020, 209 p.
18. **Janno Siim.** Non-Interactive Shuffle Arguments. Tartu 2020, 154 p.
19. **Ilya Kuzovkin.** Understanding Information Processing in Human Brain by Interpreting Machine Learning Models. Tartu 2020, 149 p.
20. **Orlenys López Pintado.** Collaborative Business Process Execution on the Blockchain: The Caterpillar System. Tartu 2020, 170 p.
21. **Ardi Tampuu.** Neural Networks for Analyzing Biological Data. Tartu 2020, 152 p.

22. **Madis Vasser.** Testing a Computational Theory of Brain Functioning with Virtual Reality. Tartu 2020, 106 p.
23. **Ljubov Jaanuska.** Haar Wavelet Method for Vibration Analysis of Beams and Parameter Quantification. Tartu 2021, 192 p.
24. **Arnis Parsovs.** Estonian Electronic Identity Card and its Security Challenges. Tartu 2021, 214 p.
25. **Kaido Lepik.** Inferring causality between transcriptome and complex traits. Tartu 2021, 224 p.
26. **Tauno Palts.** A Model for Assessing Computational Thinking Skills. Tartu 2021, 134 p.
27. **Liis Kolberg.** Developing and applying bioinformatics tools for gene expression data interpretation. Tartu 2021, 195 p.
28. **Dmytro Fishman.** Developing a data analysis pipeline for automated protein profiling in immunology. Tartu 2021, 155 p.
29. **Ivo Kubjas.** Algebraic Approaches to Problems Arising in Decentralized Systems. Tartu 2021, 120 p.
30. **Hina Anwar.** Towards Greener Software Engineering Using Software Analytics. Tartu 2021, 186 p.
31. **Veronika Plotnikova.** FIN-DM: A Data Mining Process for the Financial Services. Tartu 2021, 197 p.
32. **Manuel Camargo.** Automated Discovery of Business Process Simulation Models From Event Logs: A Hybrid Process Mining and Deep Learning Approach. Tartu 2021, 130 p.
33. **Volodymyr Leno.** Robotic Process Mining: Accelerating the Adoption of Robotic Process Automation. Tartu 2021, 119 p.
34. **Kristjan Krips.** Privacy and Coercion-Resistance in Voting. Tartu 2022, 173 p.
35. **Elizaveta Yankovskaya.** Quality Estimation through Attention. Tartu 2022, 115 p.
36. **Mubashar Iqbal.** Reference Framework for Managing Security Risks Using Blockchain. Tartu 2022, 203 p.
37. **Jakob Mass.** Process Management for Internet of Mobile Things. Tartu 2022, 151 p.
38. **Gamal Elkoumy.** Privacy-Enhancing Technologies for Business Process Mining. Tartu 2022, 135 p.