Computers & Education 161 (2021) 104054

Contents lists available at ScienceDirect

Computers & Education

journal homepage: http://www.elsevier.com/locate/compedu

Educational scalability in MOOCs: Analysing instructional designs to find best practices

Kasch Julia[a,*,1], Van Rosmalen Peter[b], Kalz Marco[c]

[a] Welten Institute, Open University of the Netherlands, P.O. Box 2960, 6401 DL, Heerlen, the Netherlands
[b] Maastricht University, the Netherlands
[c] Heidelberg University of Education, Germany

A R T I C L E I N F O

Keywords:
Distance education and online learning
Informal learning
Lifelong learning
Pedagogical issues
Teaching/learning strategies

A B S T R A C T

This study aims to reveal insights into the educational design of Massive Open Online Courses (MOOCs) in particular on their educational scalability: How do MOOCs provide interaction and formative feedback to high student numbers without being highly depending on the capacity of the teacher? We have applied a design analysis instrument that was specifically developed for large-scale online courses to analyse fifty MOOCs in a qualitative way. The goal of the analysis was to detect scalable best practices of formative feedback and interaction and focused on when, how and from whom students received formative feedback. To get more insight into the scalable best practices we also investigated on which complexity level they were provided. The analysis indicated scalable best practices on various complexity levels and across different learning activities. This shows that scalable formative feedback and interaction can be provided in MOOCs through different formats such quizzes, peer-feedback and simulations. The majority of the MOOCs in our sample provide student-content interaction during knowledge transfer activities ('knows'). A selection of design examples is discussed as potentially best practices for educational scalability, not only for MOOCs but also for online education in general. While the study shows examples of scalable design choices in (open) online education, it also indicates a need for more elaborate interactions and feedback in MOOCs in order to improve their educational value and quality.

1. Introduction

While the phenomenon of MOOCs has been announced already several times to be just a temporary hype without substance, the amount of (empirical) studies about learning in and with MOOCs has increased substantially (Rasheed, Kamsin, Abdullah, Zakari, & Haruna, 2019; Reich, 2015). In addition, the number of institutions and courses is still increasing which points to the fact that MOOCs will stay as a specific format for learning (ClassCentral, 2019). Besides studies on learning effectiveness, student retention and attrition and organisational questions, a rather neglected area of research has been on the design quality of these special types of open courses. Literature on the design of MOOCs focuses on how to design MOOCs (Conole, 2015; Rosewell & Jansen, 2014; Warburton & Mor, 2015) or evaluate to what degree MOOCs apply design principles (Margaryan, Bianco, & Littlejohn, 2015; Oh, Chang, & Park, 2019).

* Corresponding author.
E-mail addresses: julia.kasch@ou.nl, j.kasch@uu.nl (K. Julia).
[1] first author.

https://doi.org/10.1016/j.compedu.2020.104054
Received 21 April 2020; Received in revised form 26 August 2020; Accepted 19 October 2020
Available online 3 November 2020

Those studies build on known design principles such as Ten Principles of Instruction (Merril, 2013) or more specific Clark and Mayer's (2011) e-learning principles.

The study by Margaryan et al. (2015), to our knowledge the most-cited MOOC design study conducted, uses their CourseScan instrument to analyse the instructional design quality of a sample of 76 MOOCs. Coursescan is a checklist according to the 'Ten Principles of Instruction' (Merril, 2013), a classical instructional design approach with example dimensions like problem-centeredness, encouragement of collaborative learning and provision of expert feedback. The majority of MOOCs in their sample scored low on these ten fundamental design criteria of instructional quality confirming assumption of MOOC sceptics. In a similar approach Egloffstein, Koegler, and Ifenthaler (2019) have recently analysed the instructional quality of MOOCs in the domain of business and management with a slightly adapted CourseScan instrument and have concluded a "low overall instructional quality". In another recent study, Oh et al. (2019) used Clark and Mayer's (2011) e-learning principles to analyse a sample of 40 computer science xMOOCs on two different MOOC platforms. Also, this study reports a relatively low application of the principles evaluated. Within the context of the domain studied, the authors wonder why the strength of computer science itself regarding its technological innovation capacity is not exploited more. In another recent study, Hendriks, de Jong, Admiraal, and Reinders (2020), have conducted an analysis of MOOCs in the medical domain to analyse their instructional quality with respect to integration into campus teaching. The authors also apply the CourseScan instrument and their samples receives better scores with respect to the dimensions application, authentic resources, problem-centeredness, and goal-setting compared to earlier studies. The authors conclude that "more effective and efficient MOOC assessment methods are needed for the purpose of large-scale integration".

Jung, Kim, Yoon, Park, and Oakley (2019) have conducted a study about the impact of perceived instructional quality on learner control, sense of progress, and perceived effectiveness in MOOCs. In this study, authors have analysed instructional quality via the dimensions course content, course structure and organization, transactional interaction between student and content and assessment. The study shows that transactional interaction and course structure are significant predictors of perceived effectiveness. It is important to say, that this study only focuses on one transaction dimension. Authors conclude that "formative assessments and deep dialogues through instructors or experts are important". Askeroth and Richardson (2019) have conducted a qualitative study via semi-structured interviews in which they analyse the instructor perspective on learning quality in MOOCs. The instructors see scale aspects as a challenge while stressing that interaction is an important dimension to support learning for the diverse audience which MOOCs attract.

While the application of instructional principles can help to illustrate common design issues with MOOCs, such an approach fails to take into account that MOOCs provide a unique context for education (Ross, Sinclair, Knox, Bayne, & Macleod, 2014) and cannot be compared to a traditional education context for which those principles have been constructed. To compare the educational design between MOOCs and face-to-face courses, or even between face-to-face courses we would need a systematic way and realistic design standards which are commonly applied. It is not clear, if f2f-courses in higher education would ultimately be designed according to these standards and therefore, f2f-courses are not suited as a benchmark to compare against. Some authors of recent studies acknowledge the importance of scalability aspects for the design of MOOCs, but fail to apply scalability as a dimension of the instructional design components (Askeroth & Richardson, 2019). While classical instructional design instruments might work well for a typical formal learning context with a fixed amount of learners and a constant teacher/learner-ratio, the MOOC context requires a different approach that can cover the dynamic and broad range that MOOCs can have with respect to their size, teacher/learner ratio but also different levels of learning goals. Yet, it is unclear whether and to what extent classical design principles are applied in face-to-face courses. In addition, existing studies point out design issues, but fail to identify promising best practices. Best practices or design patterns may assist the designer of a course since they capture successful designs (Kasch, Van Rosmalen, & Kalz, 2017). Concrete examples and literature which illustrate educational best practices in MOOCs are absent. Given the increasing number of peer-reviewed MOOC literature (Liyanagunawardena, Adams, & Williams, 2013; Rasheed et al., 2019) more research is needed that starts analysing obstacles and weaknesses brought forward by earlier studies (Reich, 2015). Our own research efforts in particular the 'Educational Scalability Analysis Instrument' work towards bridging this gap in the literature since it enables an alternative way of looking at educational design in MOOCs. By zooming in on design principles regarding formative feedback and student interaction in MOOCs we provide insights into current design choices, issues and their potential. Questions regarding the scalability of the educational system as a whole are increasing. "Taking note that 414.2 million students will be enrolled in higher education around the world by 2030 – an increase from 99.4 million in 2000, and that online, open and flexible education is going mainstream, the importance of quality learning outcomes for learners cannot be overestimated." (Ossiannilsson, Williams, Camilleri, & Brown, 2015). The increasing enrolments in higher education worldwide require not only insight in the instructional design quality of MOOCs and online education but also best practices on how to create scalable teaching practices in campus education in which digital technologies are used. Moreover, the ongoing forced transition to online education due to Covid −19 did put considerable effort on staff to design and execute various online teaching formats. It is unlikely that after this experience education will return to 'business as usual' ignoring the gained experiences and practices. To build upon this successfully, best practices to guide scalable designs may give an important contribution.

Since existing design analysis instruments do not fit to the MOOC context and the purpose of our analysis, we have developed our own course analysis instrument following a qualitative approach (Appendix A). Unlike the CourseScan instrument by Margaryan et al. (2015), our 'Educational Scalability Analysis Instrument' was developed to map different design aspects regarding constructive alignment, types and complexities of learning tasks and the methods used to provide scalable interaction and formative feedback. Therefore, our instrument does not control whether and to what extent design principles are reflected in the MOOC but it analyses how design principles are used. It sheds light on the underlying course structure and is based on related design guidelines (Kasch et al., 2017). By using the instrument design choices become clear. The MOOCs we refer to in this study are courses to facilitate non-formal, open online courses that facilitate learning at scale (Joksimović, Poquet, Kovanović, Dowell, Mills, Gašević et al., 2018).With this study we examine the design of MOOCs and their educational scalability in search for best practices. The study is in line with the research tradition of learning design which is focusing on enabling teachers and other educational professionals to make pedagogically-informed decisions for designing technology-enhanced learning and teaching formats and the use of appropriate resources and methods in this design process (Bennett, Agostinho, & Lockyer, 2015; Conole, 2012). Due to the restrictions of earlier studies discussed we introduce a new framework into the study of design quality of MOOCs and focus therewith on the following guiding research question:

How is scalable formative feedback and student interaction realized in MOOCs?

To answer this research question, we have conducted a systematic analysis of 50 MOOCs from different platforms and domains to identify best practices in scalable MOOC design. Best practices are operationalized by applying an earlier developed instrument (Kasch, Van Rosmalen & Kalz, 2017). We focus on best practices instead of problems of current MOOC designs since we want to inform future MOOC designers about potentially scalable design options. The article is structured as follows: We introduce first our theoretical framework before we discuss the methodological aspects of our study. In the findings section we report about results of the analysis and discuss the results and their implications for theory and practice.

1.1. Theoretical framework

The underlying theoretical framework of this study is based on an earlier developed educational scalability framework (Kasch, Van Rosmalen & Kalz). With the educational scalability framework we introduced a new perspective on scalability in education by pointing out the inherent connection between scalable feedback and interaction practices, (indirect) costs via time investment by instructors, constructive alignment and quality in education. The framework is inspired by Ferguson and Sharples (2014) who state that teachers as well as students should gain benefit from scale in education as well as by Lane (2014) and Daniel and Uvalić-Trumbić (2011, pp. 1–12) who draw attention to the Iron Triangle (Kissick, 1994) and its overlap with (open) online education. The Iron Triangle states that there is a trade-off between three dimensions: scale, costs and quality when it comes to course design. A change in one dimension, for example an increase in scale, will always lead to a change in the other two dimensions such as increasing costs and decreasing quality. Translated to the course context, this would mean that an increase in student numbers (scale), will lead to an increase in staff/teachers (costs) in order to remain high educational quality (quality). Some researchers, however, argue that with the emergence of (open) online education, we are able to provide high quality courses, available to a large scale of learners with low costs (Miyazoe & Anderson, 2013). In this earlier study we have defined educational scalability in the following way: “Educational scalability is the capacity of an educational format to maintain high quality despite increasing or large numbers of learners at a stable level of total costs” (Kasch, Van Rosmalen & Kalz, 2017).

Following our previous work, the current study explores the educational design of MOOCs to analyse if we can identify best practices for scalable MOOC design. Since the provision of formative feedback (Mory, 2004) and interaction (Miyazoe & Anderson, 2013) are regarded as the bottleneck for scalability and a particularly challenging part of a course design special attention was payed to these two aspects. To get insights into the educational scalability of MOOCs the ‘Educational Scalability Analysis Instrument’ has been used. The instrument is divided into five sections with each a different focus: background information of the MOOC, student-teacher interaction, student-student interaction, student-content interaction, final general questions (Appendix A). For a better understanding of this study and our findings, the major underlying elements will be explained before we proceed with the research design of this study.

1) Interaction types

Interaction plays a dominant role in education. Learning always takes place in an interaction context, in which the learner is interacting with the teacher, other students and/or the content. Especially in distant and online education learners interact with content (learning environment) by e.g. receiving automated feedback, using tools and multimedia. Interaction is a vital factor influencing the educational scalability of a course design, it plays a role on all three dimensions of the Iron Triangle. Facilitating interaction in a MOOC on a large scale with hundreds or thousands of students asks for a course design that enables students to communicate and interact in a structured way. Additionally, to provide high quality interaction, interaction should be meaningful to students’ learning. This will influence the costs in terms of teacher time investment before and/or during the runtime of a MOOC but also time investment of fellow students.

We followed a proposal by Anderson (2003) who differentiates between student-teacher (S-T), student-student (S–S) and student-content (S–C) interaction. According to the author, student learning is supported as long as one of the three interaction types is provided at a high level. Incorporation of all three interaction types at a high level would increase the general learning value and experience, yet it might increase costs and time. We incorporated Miyazoe and Anderson’s (2013) interaction types in the ‘Educational Scalability Analysis Instrument’ to analyse, which types where offered in a MOOC, how they were offered and for what kind of learning activities (Kasch et al., 2017). This enabled us to find best practices on interaction type level and learning activity level.

2) Interaction richness

The richness of an interaction is the second major focus point of our instrument. To provide formative feedback on a large scale, teachers/designers have different options such as the use of video tools to reach high student numbers in one session, automation of the feedback response or peer-feedback (Siemens & Tittenberger, 2009). In this study we were interested in the choices that were made to provide formative feedback and with it the richness of that interaction. With ‘richness’ we refer to the meaningfulness of the formative feedback that a student receives. To analyse the richness, we refer to Mory's work on feedback levels (2004). Mory (2004) writes about different levels in the provision of feedback, distinguishing between the simple verification of whether a response was correct or incorrect and more elaborated forms of feedback in which learners receive an explanation of why a response is (in)correct and getting another try.

3) Complexity and constructive alignment

The third dimension of the instrument deals with the complexity level of learning activities and the constructive alignment of learning goals, learning activities and learning assessment. There are numerous approaches to describe learning objectives or learning goals, the most widely accepted is probably the taxonomy of cognitive levels by Bloom (1956) which has been adapted by Krathwohl and Anderson (2009). For simplicity reasons and because we wanted to incorporate not only knowledge but also skills, we have used a more simplified taxonomy of learning objectives by Miller (1990). In this taxonomy, the different complexity levels of learning activities are differentiated by four levels: 1.) factual knowledge (knows), 2.) procedural knowledge and competence (knows how), 3.) performance (shows how) and last but not least 4.) application of knowledge and skills (action). These levels are complemented with the concept of constructive alignment (Biggs, 1996). In constructive alignment the alignment of learning goals, learning activities and learning assessment is analysed. For our current study this is an essential idea, since many examples of so called "learning at scale" are scalable because their objective is transfer of factual knowledge, which we regard also as a valid educational format. Taking into account the complexity level of learning activities helps us to understand the possibilities of scalable design. Is it possible to provide scalable formative feedback and interaction at large scale when dealing with high complexity learning? If so, how?

These aspects have been applied to a sample of MOOCS and the goal of the analysis is to collate these different dimensions into a systematic analysis to identify best practices. In the next chapter we introduce our methods.

2. Method

2.1. Research design

Given the focus of this study, a non-probability sampling strategy has been chosen in order to get a sample of MOOCs. For this case study, a sample size of 50 MOOCs was chosen (Fig. 1). By analysing one week of each selected MOOC with the 'Educational Scalability Analysis Instrument' we were able to analyse several key issues regarding scalable formative feedback and interaction that fits to our research goal in a time-efficient way (Creswell, 2007; Twining, Heller, Nussbaum, & Tsai, 2017). Since the instrument was specifically developed to analyse the design of feedback and interaction in MOOCs, we were able to analyse all cases and provide descriptive findings of scalable educational design choices. This allowed for a deeper and more detailed analysis of the educational design regarding scalable interaction and formative feedback in MOOCs in comparison with more general, quantitative rating instruments which have been applied in similar studies. In addition, the design was employed because we did not want to conduct a summative assessment about the quality of the MOOC design as such but rather were interested to describe and present their design and best practices in scalable MOOC design. The design of this study is based on the view that there is no standard way of designing a

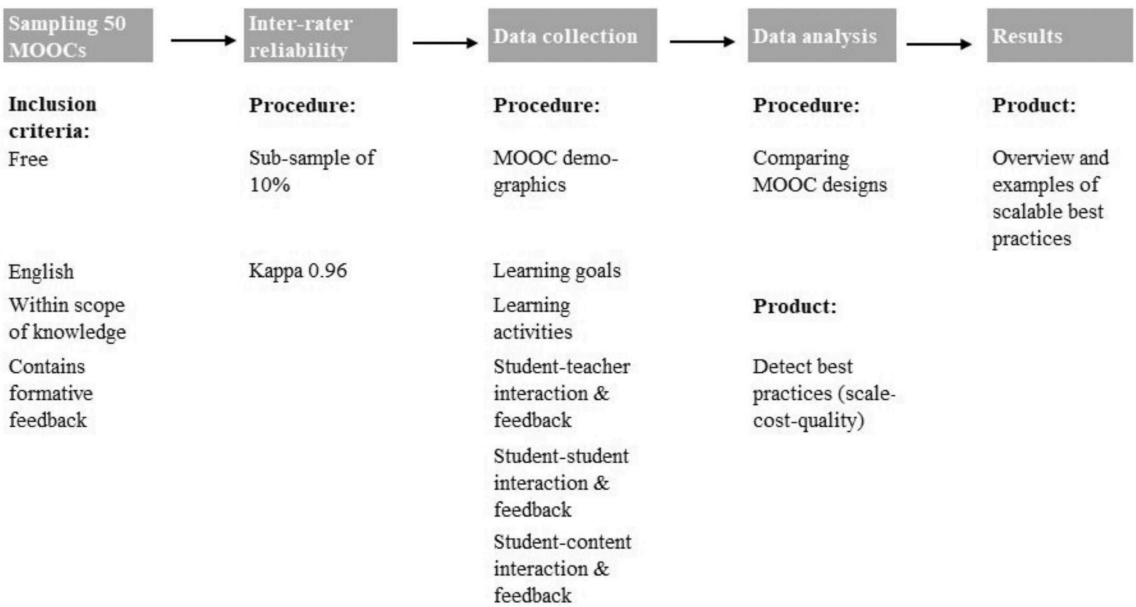

Fig. 1. *Overview of the study procedure.*

(large-scale) course. Simply applying general course design principles without taking into account the course context, learning goals and complexity will not result in a high quality (scalable) course. Additionally, the transfer of classical instructional design guidelines to courses with massive numbers of participants does not take into account the specific challenges of large-scale courses from our opinion.

With this study we strived to find and describe scalable best practices as they are currently applied. Although there are certain conditions that should be met when striving for high quality, such as constructive alignment, clear instructions and support, the implementation should meet the individual course context and requirements. Course goals, settings and (technical) possibilities have as many other aspects influence on the educational design of a course. Scalability and course design quality can be achieved in various ways and in different settings. By focusing on scalable best practices and examining a sample of MOOCs from different providers and on different platforms that contains formative feedback and interaction, we can illustrate the key issues of interest within each MOOC (Creswell, 2007).

2.2. Sample

In October 2018, 2068 MOOCs were available in English via ClassCentral (https://www.classcentral.com/), a search engine site for MOOCs. From the 2068 MOOCs, a sample of 50 MOOCs was selected, starting at the beginning of the list (Appendix B). MOOCs who met the following inclusion criteria were included in the sample of 50: (a) contain formative feedback, (b) free of charge and all elements were accessible without payment, (c) the researchers had enough content knowledge to understand the MOOC and (d) the middle week was available. For example, in a MOOC with a runtime of 7 weeks, the 3rd week would be available.

By analysing a sample of 50 MOOCs we intended to provide a description of a broad MOOC spectrum, in a way that was feasible within the project. During the sampling we saw that Coursera had by far the biggest MOOC offer. We first approached a random sampling strategy, but later decided to go for a stratified sampling due to the required coverage of different MOOC platforms in our analysis. We prioritized to include five MOOCs from edX, FutureLearn and Open2Study which had, next to Coursera, the biggest MOOC offer.

2.3. Procedure and materials

Prior to the data collection, a subsample of 10% was analysed independently, by two of the researchers using our 'Educational Scalability Analysis Instrument'. The inter-rater reliability on the subsample was calculated resulting in a Cohen's kappa value of 0.96. Next, discrepancies were discussed and negotiated, until consensus was reached. Subsequently, the analysis of our 50 MOOCs sample was done by one of the two researchers using the online version of our instrument in a LimeSurvey environment. In addition, for

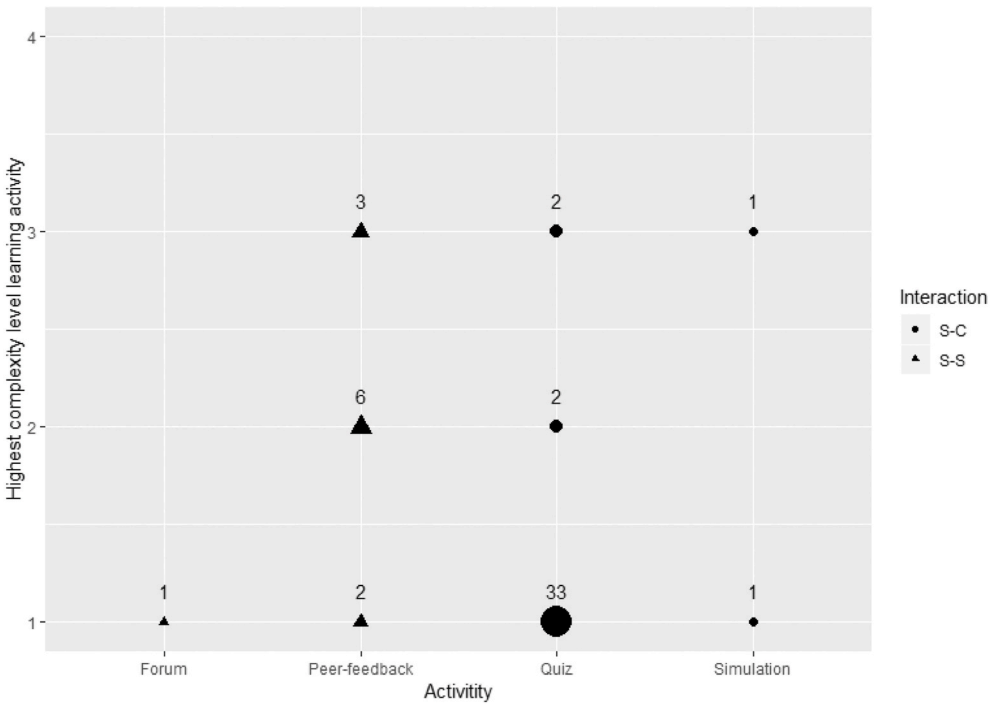

Fig. 2. Overview of the four most common learning activities (x-axis) per complexity level (y-as) in a sample of 50 MOOCs. The most common learning activities did not include S-T interaction.

documentation purposes, screenshots of interesting design choices are stored together with the analysis. During the analysis process results and questions were discussed within the research team. In this study, the choice was made to analyse one week of each MOOC. Starting at the middle week, we checked whether the week contained a form of formative feedback. If so, that week was analysed, if not, the upcoming week was looked at. For example: In the case of a MOOC with a runtime of eight weeks, the fourth week was analysed if it contained formative feedback. If not, the next week (week 5) was analysed. A MOOC that did not contain any form of formative feedback was not included in the sample. The first or last week of a MOOC since these often evolve around activities related to preparation, summative assessment or finalisation which would lead to a biased picture for our analysis. On average, it took 1 h to analyse one week of a MOOC. A MOOC analysis involved the researcher surveying the course description and analysing one week of the MOOC using the analysis instrument (Kasch et al., 2017). The learning goals, learning activities, with the corresponding instructions and materials and where possible student submissions of that week were analysed. To analyse the different leaning interactions (student-student, student-content, student-teacher) the researcher analysed all activities and materials that provided interaction such as quizzes, simulations, video's, fora and so forth. The analysis also included that the researcher participated in the learning activities such as reading papers, filling in quizzes, participating in peer-feedback activities. That way we were able to get a better understanding of the learning activities and the quality of the provided feedback and interaction. Whilst analysing a MOOC, the analysis instrument was filled in online within the LimeSurvey environment.

3. Findings

This study set out with the aim of investigating how scalable formative feedback and student interaction is realized in MOOCs. The analysis of 50 MOOCs revealed promising design choices regarding this question. Within this study, an educational design choice (i.e. learning activity, instruction, support) that provides rich/elaborate formative feedback in a scalable way through student interaction is analysed to see if and how it complies with scalable best practices. It is important to note that on each complexity level scalable best practices can be applied. In the following section, scalable best practices in the provision of formative feedback and interaction are summarized. For a detailed overview of the best practices see Appendix C (scalable S–C interactions), D (scalable S–S interactions) and E (scalable S-T interactions).

The findings are presented as follows: First an overview of the general findings is provided (see Figs. 1 and 2). Subsequently we zoom in to best practices and potential best practices, i.e. practices that with minimal changes have potential. In the current study, S–C interaction was by far the common form of student interaction. S–C interaction occurred 45 times, mostly during closed-ended multiple choice quizzes testing students' factual knowledge. Peer-feedback activities (S–S interaction) are the second most common way of receiving formative feedback. Examples of S–S interaction were found 13 times and in only 7 times students interacted with teachers (S-T interaction).

When looking at the richness of the provided formative feedback, we found that elaborate feedback was provided during all three interaction types. Rich, meaningful feedback was provided during quizzes (S–C) through highly elaborate automated feedback. Although the majority of quizzes provided superficial corrective feedback, a number of exceptions were found. Yet, receiving rich and personalized feedback (to some extent) was only available via fellow students during a peer-feedback activity with open text boxes or by office-hour videos/live hangout sessions with the teacher.

3.1. *Learning activities and complexity levels*

The analysis showed that in general learning goals were vague and did not clearly inform students about what to expect. In our sample of 50 MOOCs, 10 did not provide learning objectives at all. In 33 MOOCs the learning objectives were formulated in a non-measurable way, giving no or not enough information about the expected behaviour that students will perform, the specific content, the requirements under which the student will perform the behaviour and the needed minimum performance in order to succeed in the objective. However, we also found 7 MOOCs, where learning objectives were provided at a high level, being clear on some of the previously mentioned elements a learning goal should contain.

Fig. 2 gives an overview of the four most common learning activities in our sample, regardless of their complexity level. It shows the complexity level of the most common learning activities, including learning activities of all complexity levels. One week can contain different learning activities resulting in more than 50 cases. Sixteen learning activities could not be visualised due to missing information about their complexity level. As a result Fig. 2 visualises 51 cases in a sample of 50 MOOCs. The most common learning activities did not include S-T interaction which is why only S–S and S–C interaction are visualised.

The most common learning activity used in our sample is the quiz which was provided in 37 MOOCs. Peer-feedback activities are the second most common activity which were provided in 11 out of the 50 MOOCs. In 2 cases simulations were provided and in one case the forum was used as a learning activity. All MOOCs of our sample contained a forum, however, the forums we analysed were not used as an activity but a place for informal communication and interchange without any tutor support or a specific task or role directive. Students were given questions they could reflect on but did not receive clear instructions on how to use the forum for discussion e.g. to how many comments they should react; why they should/could discuss with others.

When looking at the complexity levels of these learning activities, Fig. 2 shows that from the 37 quizzes, 33 were provided on a low complexity (level 1), two quizzes were on level 2 (knows how) and 2 quizzes on level 3 (shows how). This shows that most of the time quizzes are used to assess factual knowledge, yet there are exceptions where quizzes are used on a higher level. We found in our sample peer-feedback activities on three of the four complexity levels. In two cases peer-feedback activities were provide at level 1 (knows), in three cases on level 2 and in another three cases on level 3. It is important to note that the distinction in complexity levels in this study is made to get an impression of the diversity of learning activities. A level 1 complexity level is not seen as something negative or low of quality. Acquitting factual knowledge is often seen as a first step in learning and in some cases needed as a basis in order to continue learning on higher complexity levels.

3.2. Interaction and formative feedback

As shown in Fig. 3, the most common interaction type in our sample is the S–C interaction which appeared in 39 out of the 50 MOOCs. In contrast to Figs. 2 and 3 visualises only the learning activities with the highest complexity. In our sample there was no case in which a learning activities with the highest complexity level provided S-T interaction. Therefore, only S–S and S–C interaction are visualised in Fig. 3. We adapted Mory's (2004) work to our MOOC research context and used the following four instead of five levels: Level 1 'no feedback' indicates that there is no or limited feedback (e.g. correct or not-correct or a simple answer model), level 2 'simple verification feedback' is used if the student receives information that not only points out that an answer or performance was correct or incorrect but also provides a short explanation why it was correct or incorrect or what should be contained and why. The richness of the interaction increases if the students receive information on why an answer was correct and the other options incorrect (level 3 'elaborated feedback') and, finally, level 4 is applied if the feedback also addresses where to focus on and what to improve.

3.2.1. Student-content interaction and formative feedback

The most interaction was provided through student-content interaction. Examples of S–C interaction are quizzes or activities such as simulated learning tasks and games. In 34 cases students interacted with the learning environment/content during low complexity learning activities (Fig. 1). There were only 5 cases where S–C interaction took place on higher complexity levels (level 2 and 3) in the form of a quiz and a simulation. The level of richness of an interaction is defined by the quality of the feedback. The richness of the S–C interaction was in 25 out of the 39 cases low (level 1). A simple verification through automated feedback from the learning environment on students' performance was provided. A simple multiple choice quiz where the automated feedback consists of a cross in case of a wrong answer and a checkmark in case of a correct answer would be an example of a low complexity learning activity (level 1) where the richness/value of the interaction with the student was poor (level 1). Multiple choice quizzes are by far the most chosen form to provide students with feedback. It was interesting to see the differences between quizzes regarding the elaborateness of the feedback and the modality of answer options. The analysis showed that best practices in S–C can be found in the form of 'automated elaborate formative feedback' during quizzes, simulations and games. The educational quality of a quiz can be influenced by the modality of the answer options (i.e. multiple choice, open-ended, written answer options, audio recordings etc.). Other forms of S–C interaction were a checklist, a wordcloud and a chatbot.

Table 1 gives an overview of the detected best practices for S–C interaction which are 'automated elaborate formative feedback,

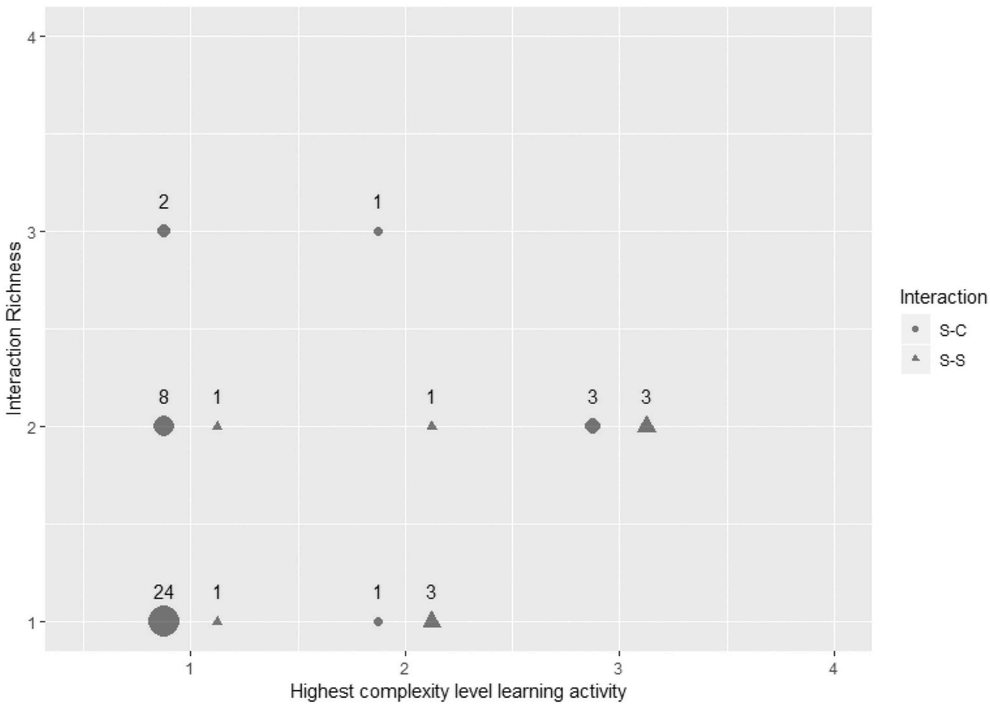

Fig. 3. *An overview of the number of S–C and S–S interactions identified on the highest complexity level in the analysed week of each MOOC per complexity level and per interaction richness level (n = 48, 2 interactions were ambiguous; the sample contained no S-T interaction on the week's highest level).*

'simulated learning tasks' and 'automated process and content hints'. The best practices summarized in Table 1 provide insight into scalable design options when using S–C interaction. Table 1 provides additional information to Fig. 1 where the four most common learning activities are presented. The detected best practices can be applied on various complexity levels (low as well as high). Examples and detailed descriptions can be found in Appendix C.

3.2.2. Student-student interaction and formative feedback

The second most common interaction type in our sample was the S–S interaction where students interact with each other in the MOOC. In the 50 MOOCs, students interacted in 8 cases through peer-feedback and in 1 case through structured forum discussion. In those 9 cases, only 2 cases used S–S interaction for low complex activities but 7 cases for higher complex activities. However, the richness of S–S interaction never reached higher than level 2 indicating that in none of the cases students were asked to provide each other with information on how to improve their work. When interacting with each other students had the task to provide simple verification feedback in the form of correct/incorrect (level 1), yet they were also asked to provide additional information on why a performance was correct or incorrect (in 4 cases). Table 2 gives an overview of the detected scalable best practices for S–S interaction which are 'peer-feedback instructions' and 'discussion prompts'. Examples and detailed descriptions can be found in Appendix D

Additional note: Despite these promising findings, none of the MOOCs of our sample mentioned the educational goal of peer-feedback. In none of the MOOCs, students were informed about the benefits of providing and receiving peer-feedback. Why should they engage in peer-feedback? Despite the fact that it is a scalable solution, no attention was given to the student perspective and their needs. Additionally, students were not informed what to do with the received peer-feedback herewith limiting the value of peer-feedback from a student perspective.

3.2.3. Student-teacher interaction and formative feedback

Table 3 summarized two forms of scalable best practices that were applied during S-T interaction. S-T interaction was provided in 7 out of 50 MOOCs. As mentioned earlier, S-T interaction is not visualised in Figs. 1 and 2 because it was not provided during a high complexity learning activity nor during one of the most common learning activities. Nevertheless, scalable best practices were found in the form of 'personalized feedback' and 'guest speakers'. Scalable S-T interaction entails that the teacher is able to interact meaningfully with high student numbers and provide them with formative feedback in a time efficient way. Implementing S-T interaction in an open online setting can be both complex and resource intensive. Examples and detailed descriptions can be found in Appendix E. Appendix D

4. Discussion

In reviewing the literature on MOOC design, no information was found on scalable formative feedback and student interaction. There are previous studies that have noted the importance of high quality design in MOOCs, yet concrete well-described examples how to implement quality designs are missing. Especially, formative feedback and peer-feedback/peer-review at scale are topics, teachers and MOOC providers are interested in (Ferguson & Sharples, 2014; Nicol & MacFarlane-Dick, 2006; Warburton & Mor, 2015).

This study set out with the aim of finding scalable best practices of formative feedback and interaction. The most important finding of this study were design examples that showed that the Iron Triangle (Daniel & Uvalić-Trumbić, 2011, pp. 1–12) can indeed be optimized in the context of formative feedback and interaction in MOOCs. We found best practices of the provision of high-quality feedback and interaction at large scale without increasing on-demand teacher involvement. However, optimizing all three aspects of the Iron Triangle (cost, scale, quality) is not *per* definition achieved by MOOCs or the benefits they entail such as automated feedback or online peer-feedback.

The results of this study match on a high level those of a previous MOOC analysis study by Margaryan et al. (2015). We could confirm that the majority of learning activities was of low complexity, focusing mostly on factual knowledge, rather than providing a range of learning activities on different complexity levels and incorporating more authentic learning activities. This was also reported

Table 1
Overview detected best practices for student-content interaction.

Identified best practices	
Automated elaborate formative feedback	Provide quizzes, simulations and/or games for students to check, practice and apply their knowledge. Design elaborated formative feedback for these learning activities. Elaborate feedback can contain information and explanations about why an answer is correct or incorrect. Depending on the learning activity it also informs students about how to improve their work/performance and where to look in the course materials for additional information on that topic.
Simulated (authentic) learning tasks	Enable students to apply their theoretical knowledge and practice skills in a controlled environment. Through S–C interaction in the form of quizzes, simulations or games students can be applied at large scale online. Authentic learning activities are not only relevant within the MOOC context but also outside of it. Practicing in a controlled online environment also makes it possible to support students with automated formative feedback. Learning activities can be provided with increasing complexity, enabling students to practice in different scenarios.
Automated content and process hints	Hints provided during or after the learning activity in the form of process (which steps should be taken next) or contents hints (what is the correct answer). Provide short recaps in the form of very short video clips of texts that shortly summarize the main take-home message from that exercise. Personalized automated feedback through chat-bots enables students to interact with a virtual agent and provides an interactive online learning experience.

Table 2
Overview detected best practices for student-student interaction.

Identified best practices	
Peer-feedback instructions	Give students, prior to a peer-feedback activity the opportunity to get familiar with the online peer-feedback process and, if applicable tools (e.g. via rubric). Provide technical information on how the peer-feedback process is organized in the MOOC (i.e. how to upload an assignment, how to download a peers' assignment, how to submit feedback). Additionally, inform students about when (i.e. deadline) and how to provide feedback (i.e. the order of tasks that need to be done). Provide review criteria prior to the peer-feedback activity and supplemented with critical questions that can function as leading questions during the review process. Instructions also inform students about what to expect from peer-feedback: how much time they should invest in providing feedback and how to be constructive and supportive.
Discussion prompts	Enable structured and guided online discussions among students by providing prompts, questions dilemma's. Stimulate students to discuss, reply and share their ideas. Inform them about some ground rules such as how long their comments should be and how to reply to peers.

Table 3
Overview detected best practices for student-teacher interaction.

Identified best practices	
Personalized feedback	Develop live hangout sessions or pre-recorded 'office hour' videos in which the teacher answers most asked questions or comments' on frequently made mistakes. The feedback is not only personal but also timely and gives the students the possibility to interact with the teacher and learn from each other's' questions.
Guest speakers	Invite an expert to answer student questions or give an online lecture. The expert can be invited for a live hangout session with the teacher or can record a video session. Guest speakers can interact with students, answer questions from their own background and provide personalized feedback.

by Bali (2014) who conducted a small MOOC pedagogy study. We also found a lack of clarity and instruction when it comes to collaborative learning. Students are often not informed (enough) about their roles and what is expected from them. As was reported in previous studies (Bali, 2014; Margaryan et al., 2015), a lack of expert feedback in the form of personal teacher feedback (S-T interaction), was also the case in our sample. This, however, is not surprising in a MOOC context were more teacher-independent learning is promoted. In MOOCs, teachers do not have a different role but will also have to invest more time in course development and pre-planning of learning activities (Ross et al., 2014). Instruction can be provided through text, video and visualisations enabling student support through different channels (Anderson & Dron, 2011). That way, expert feedback can still be provided in an automated way during quizzes, simulations, games or other tools. The sequencing and alignment of learning activities, the course structure and positioning of instructions, tasks, examples and tips need to be planned and applied in a transparent way so that students can easily navigate through it. Not being able to guide students on a daily basis through the course materials, teachers have to foresee to some extent the study behaviour and needs of students and support them through the course. In line with previous studies, we found that learning objectives, in the vast majority, were either not measurable or not mentioned. Without informing students in a clear way what they can achieve after having participated in a course the overall quality of a course design decreases. This also made it difficult for us to analyse the constructive alignment in MOOCs. Without understanding what the goals are, it is not possible to say something about the alignment between the goals, learning activities and (formative) assessment. Ideally students should be informed about the learning goals prior to enrolling.

Contrary to the study by Margaryan et al. (2015) who concluded that the instructional design quality of MOOCs in general is low, the results of this study give a more nuanced picture. MOOCs with high instructional design quality are available and show scalable best practices regarding the provision of formative feedback and student interaction. This also shows that claims about instructions design quality in MOOCs highly depends on the approach that is used when analysing MOOCs. Compared to the study Margaryan et al. (2015) we analysed common design principles regarding formative feedback, interaction, constructive alignment and complexity levels in a qualitative way. This enabled us to look into the possibilities of current instructional design in MOOCs.

Something that was not the focus of this study but that surprisingly stood out during the analysis, were the design differences between the MOOC platforms in our sample (Coursera, edX, Canvas, FutureLearn, Open2Study). Differences could be seen regarding the type and amount of support students receivied during learning activities and the diversity in learning activities. This shows that the choice of platform and thus the common practice in a platform or the possibilities that come with that platform can influence the instructional design of a MOOC. However, Bali (2014) found that even MOOCs of the same platform (Coursera) show differences in pedagogy. This makes it difficult to talk about MOOCs in general terms since difference can occur between as well as within platforms. One aspect that can influence the educational scalability is whether a MOOC is available to study right away or whether each week of the MOOC will be made available on a weekly basis. If the entire material of a MOOC is available at once, as we have seen in all Coursera courses, students are free to choose how to proceed through the course. However, this can also mean that there are not enough students around when it comes to student-student interaction such as a peer-feedback activity. Although peer-feedback can be a scalable way of providing formative feedback, it still depends on students' availability.

MOOCs that provide elaborate formative feedback through automatization or even by teachers are not the majority, yet they are available. Clear, structured student interaction in the form of peer-feedback can and is achieved in some cases just like the possibility to engage in learning activities of various complexity ranging from factual knowledge quizzes to more complex quizzes or simulations. By focusing on how scalable formative feedback and interaction are realized in MOOCs and using design quality criteria that are related to these two aspects (Merrill, 2013; Mory, 2004; Siemens & Tittenberger, 2009) we were able to provide insights and examples on MOOC design. When analysing the educational design quality of formative feedback and interaction we did not choose to compare MOOC design with face-to-face course design, since it does not add anything to the discussion about design quality as long as face-to-face courses are not analysed and assessed by the same principles in a standardized way. Yet, we found that common peer-feedback and course design recommendations are applied in MOOCs enabling scalable formative feedback and interaction. The relevance of providing elaborated formative feedback is mentioned and promoted in several studies (Nicol & Macfarlance, 2006) as is the provision of real-world problems/authentic tasks (Merrill, 2013).

Based on identified scalable best practices for interaction and formative feedback in MOOCs and the literature we support the following design recommendations:

- Constructive alignment regarding learning goals and learning activities (Biggs, 2003; Conole, 2015)
- Automated (personalized) formative feedback, elaborating on correct and incorrect answers (Hattie & Timperley, 2007; Mory, 2004)
- Supported and trained peer-feedback activities (Boud & Molloy, 2013; Carless & Boud, 2018; Gielen, Peeters, Dochy, Onghena, & Struyven, 2010)
- Authentic practice through simulations and games (Siemens & Tittenberger, 2009; Warburton & Mor, 2015)
- Content and process hints; examples (Mory, 2004)
- Structured online interaction and communication (Gamage, Fernando, & Perera, 2015; Siemens & Tittenberger, 2009; Warburton & Mor, 2015)
- Use of external materials and content experts (Siemens & Tittenberger, 2009; Warburton & Mor, 2015)

4.1. Significance and limitations

The purpose of this study, to explore course design choices in current MOOCs and provide insight into (potential) best practices, is valid given the lack of knowledge regarding the educational design of MOOCs. This study contributes to the literature and course design practice, for three reasons: (1) it offers an approach on how to analyse the educational scalability of a course design; (2) it provides examples of (potential) scalable best practices regarding formative feedback and student interaction in MOOCs and (3) it introduces a new perspective on educational scalability and design analysis.

Limitations of this study are connected to our sample size and our inability to analyse all aspects of a course design. Given the high number of online MOOCs, it is not feasible to analyse all of them in such a detail which means that the study itself has a scalability issue. Conducting this study has showed us that educational scalability can be analysed in several ways using different perspectives and approaches. Within this study we have focused on specific parts of a course and analysed courses from an outsider perspective. In spite of the limitations, transferability of our findings is possible as they are applicable in the context of online courses and offline, on-campus or flipped classrooms. The design analysis instrument enables a systematic analysis of educational design and is based on common design guidelines and principles. That way multiple courses can be analysed and compared in a systematic way providing the user (i.e. teachers, course designers, students) a detailed description of their course design. Our findings can inspire teachers, course designers and students to analyse and improve the educational scalability of their courses and/or to analyse others as a source of inspiration (Van Rosmalen, Kasch, Kalz, Firssova, & Brouns, 2017). Given the increasing student numbers and the growing importance of digital technologies in campus education, currently further augmented by Covid-19 epidemic, this study is highly relevant for the provision of MOOCs as well as online or blended education as part of campus education.

5. Conclusion

In this qualitative paper, we investigated how scalable formative feedback and student interaction is realized in MOOCs. By analysing the educational design quality of MOOCs using our 'Educational Scalability Analysis Instrument' we found scalable best practices in MOOC design and got more insight into educational design choices that can enhance quality and in the end the educational scalability of a course. The analysis shows that providing a list with scalable best practices or guidelines is a complex aim since the scalability of an educational design depends on the context and the way it is applied. Our systematic analysis using our design analysis instrument enabled us to do exactly that. It showed that design guidelines and principles for face-to-face education such as constructive alignment, elaborate formative feedback, transparent and clear instructions are also appropriate for (open) online education and can be realized at scale in MOOCs.

It became clear that, although the majority of MOOCs in our sample only provide superficial formative feedback through student-content interaction, good examples of high quality formative feedback can be found too. The approach and perspective on MOOCs and educational design can have an impact on the outcome of a design study. Overall, this study strengthens the idea that the Iron Triangle can be stretched in MOOCs (Daniel & Uvalić-Trumbić, 2011, pp. 1–12). To achieve educational scalability (high quality with low costs on high scale) it is not about simply offering formative feedback through scalable ways such as peer-feedback or automated quizzes but rather the educational quality that is provided when using peer-feedback or quizzes. These findings have implications for design-theories for (large-scale) online-learning. It became evident, that examples of scalable course design for MOOCs exist. These examples can be used as a starting point for the development of design guidelines for MOOCs, but also for future blended learning formats in which the increase of student number will require to rethink the intelligent use of digital technologies for campus education.

The findings of this study will be of interest to everyone who is interested in scalable formative feedback and student interaction in an online as well as offline context. Those who are looking for examples on how to provide scalable feedback and moreover, how to determine quality aspects of an educational design, will get insights and examples. The findings show that MOOCs hold high potential to provide elaborate formative feedback and interaction. Further research on this topic might explore how students value different levels of feedback. Another interesting study could investigate how teachers and MOOC designers go through the design process when aiming for educational scalability. Which quality aspects do they value in educational scalability and to what extent do they succeed in realizing them in practice. There can be a gap between scalable course design ideas in theory and practice, depending on the barriers teachers and designers face regarding their knowledge about a MOOC platform and its features.

Credit author statement

Julia Kasch: Conceptualization, Methodology, Validation, Formal analysis, Writing - Original Draft, Writing - Review & Editing; **Peter van Rosmalen:** Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing, Supervision, Project administration, Funding acquisition**; Marco Kalz:** Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Acknowledgement

This work is financed via a grant by the Dutch National Initiative for Education Research (NRO)/The Netherlands Organisation for Scientific Research (NWO) and the Dutch Ministry of Education, Culture and Science under the grant nr. 405-15-705 (SOONER/http://sooner.nu).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2020.104054.

References

Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *International Review of Research in Open and Distance Learning, 4*(2), 1–14.
Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *International Review of Research in Open and Distance Learning, 12*(3), 80–97.
Askeroth, J. H., & Richardson, J. C. (2019). Instructor perceptions of quality learning in MOOCs they teach. *Online Learning, 23*(4).
Bali, M. (2014). MOOC pedagogy: Gleaning good practice from existing MOOCs. *Journal of Online Learning and Teaching, 10*(1), 44.
Bennett, S., Agostinho, S., & Lockyer, L. (2015). Technology tools to support learning design: Implications derived from an investigation of university teachers' design practices. *Computers & Education, 81*, 211–220.
Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347–364.
Biggs, J. (2003). Aligning teaching for constructing learning. *Higher Education Academy, 1*(4).
Bloom, B. S. (1956). *Taxonomy of educational objectives* (Vol. 1, pp. 20–24). New York: McKay: Cognitive domain.
Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education, 38*(6), 698–712.
Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325.
ClassCentral. Online Degrees Slowdown. (2019). A review of MOOC stats and trends in 2019. https://www.classcentral.com/report/moocs-stats-and-trends-2019/. (Accessed August 2020).
Conole, G. (2012). *Designing for learning in an open world* (Vol. 4). Springer Science & Business Media.
Conole, G. (2015). Designing effective MOOCs. *Educational Media International, 52*(4), 239–252.
Creswell, J. W. (2007). Five qualitative approaches to inquiry. *Qualitative inquiry and research design: Choosing among five approaches, 2*, 53–80.
Daniel, J., & Uvalić-Trumbić, S. (2011). *The impact of new business models for higher education on student financing.* Commonwealth of Learning.
Egloffstein, M., Koegler, K., & Ifenthaler, D. (2019). Instructional quality of business MOOCs: Indicators and initial findings. *Online Learning: OLJ= Online Learning Journal, 23*(4), 85–105.
Ferguson, R., & Sharples, M. (2014). September). Innovative pedagogy at massive scale: Teaching and learning in MOOCs. In *European conference on technology enhanced learning* (pp. 98–111). Cham: Springer.
Gamage, D., Fernando, S., & Perera, I. (2015). Quality of MOOCs: A review of literature on effectiveness and quality aspects. August. In *2015 8th international conference on ubi-media computing (UMEDIA)* (pp. 224–229). IEEE.
Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304–315.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.
Hendriks, R. A., de Jong, P. G., Admiraal, W. F., & Reinders, M. E. (2020). Instructional design quality in medical massive open online courses for integration into campus education. *Medical Teacher, 42*(2), 156–163.
Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., et al. (2018). How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research, 88*(1), 43–86.
Jung, E., Kim, D., Yoon, M., Park, S., & Oakley, B. (2019). The influence of instructional design on learner control, sense of achievement, and perceived effectiveness in a supersize MOOC course. *Computers & Education, 128*, 377–388.
Kasch, J., Van Rosmalen, P., & Kalz, M. (2017). A framework towards educational scalability of open online courses. *Journal of Universal Computer Science, 23*(9), 845–867, 2017.
Kissick, W. L. (1994). *Medicine's dilemmas: Infinite needs versus finite resources.* Yale University Press.
Krathwohl, D. R., & Anderson, L. W. (2009). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
Lane, A. (2014). Placing students at the heart of the iron triangle and the interaction equivalence theorem models. *Journal of Interactive Media in Education, 2*(5), 1–8.

Liyanagunawardena, T., Adams, A., & Williams, S. (2013). MOOCs: A systematic study of the published literature 2008-2012. *International Review of Research in Open and Distance Learning, 14*(3), 202–227.
Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education, 80*, 77–83.
Merrill, M. D. (2013). *First principles of instruction: Identifying and designing effective, efficient and engaging instruction*. Hoboken, NJ: Pfeiffer.
Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*(9), S63–S67.
Miyazoe, T., & Anderson, T. (2013). Interaction equivalency in an OER, MOOCS and informal learning era. *Journal of Interactive Media in Education*, 1–15.
Mory, E. H. (2004). Feedback research revisited. *Handbook of research on educational communications and technology, 2*, 745–783.
Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.
Oh, E. G., Chang, Y., & Park, S. W. (2019). Design review of MOOCs: Application of e-learning design principles. *Journal of Computing in Higher Education*, 1–21.
Ossiannilsson, E., Williams, K., Camilleri, A. F., & Brown, M. (2015). *Quality models in online and open education around the globe. State of the art and recommendations*. Oslo: International Council for Open and Distance Education.
Rasheed, R. A., Kamsin, A., Abdullah, N. A., Zakari, A., & Haruna, K. (2019). A systematic mapping study of the empirical MOOC literature. *IEEE Access, 7*, 124809–124827.
Reich, J. (2015). Rebooting MOOC research. *Science, 347*(6217), 34–35.
Rosewell, J., & Jansen, D. (2014). The OpenupEd quality label: Benchmarks for MOOCs. *INNOQUAL: The International Journal for Innovation and Quality in Learning, 2*(3), 88–100.
Ross, J., Sinclair, C., Knox, J., Bayne, S., & Macleod, H. (2014). Teacher experiences and academic identity: The missing components of MOOC pedagogy. *Journal of Online Learning and Teaching, 10*(1), 57.
Siemens, G., & Tittenberger, P. (2009). *Handbook of emerging technologies for learning*. Canada: University of Manitoba.
Twining, P., Heller, R. S., Nussbaum, M., & Tsai, C. C. (2017). *Some guidance on conducting and reporting qualitative studies*.
Van Rosmalen, P., Kasch, J., Kalz, M., Firssova, O., & Brouns, F. (2017). Towards 'MOOCs with a Purpose': Crowdsourcing and analyzing scalable design solutions with MOOC learners. September. In *European conference on technology enhanced learning* (pp. 486–491). Cham: Springer.
Warburton, S., & Mor, Y. (2015). A set of patterns for the structured design of MOOCs. *Open Learning: The Journal of Open, Distance and e-Learning, 30*(3), 206–220.