# Master Thesis

## What is the effect of a group discussion compared to individual reading on designing an outline for a MOOC?


Name: Anna Hinze,
Email: a.hinze@student.utwente.nl
Faculty: Behavioural Sciences
Master Educational Science and Technology

Supervisor: Bas Kollöffel
Email: b.j.kolloffel@utwente.nl

Organization: Technology Enhanced Learning and Teaching, University of Twente



Date: April 4, 2018


UNIVERSITY OF TWENTE.

## Acknowledgement

I would like to thank my first supervisor, Bas Kollöffel, for the support and advise in working on this project. The many meetings paid off and I am very pleased with how this project took shape. Further, I would like to thank Eduardo Hermsen for proposing the problem and trusting me to work on a solution. Without Eduardo, many aspects of this projects could not have been realised. I hope he will find this thesis to be useful for his future projects.

One cannot write in a thesis without proof readers. I would like to thank Jonas Hinze, Irene Reiling and Daisy Peel for taking the time to read my work and making helpful suggestions. I would also like to thank my mother and Tanja for listening to my complains when things did not go smoothly.

Finally, I want to thank my dogs Jive, May and Take for providing distraction and keeping me sane during this project.

WHAT IS THE EFFECT OF A GROUP DISCUSSION COMPARED TO INDIVIDUAL READING ON DESIGNING AN OUTLINE FOR A MOOC?


## Abstract

Designing massive open online courses (MOOCs) is fundamentally different from designing face-to-face courses. MOOC designers need to undergo a conceptual change, increase their self-efficacy and improve the quality of the MOOCs they design. After conducting interviews with several experienced MOOC designers, the need for training on MOOC design has become apparent. For this quasi-experimental study, two workshops on MOOC design were developed. While one workshop included a cooperative learning activity, the other incorporated an individual reading assignment. A sample of 42 participants took part in the study. By comparing participants' scores on questionnaires and the MOOC outlines that were designed, this study aims to answer the following question: What is the effect of a group discussion compared to individual reading on designing an outline for a MOOC? Specifically, what are the effects on quality, conceptual change and self-efficacy? Based on previous research, the group discussion was expected be more beneficial in increasing quality and self-efficacy and supporting conceptual change. To shed light onto these questions, changes in quality, conceptual change and self-efficacy were statistically analysed both within each group as well as between conditions. Participants' MOOC designs were evaluated by experts based on a scoring rubric. Results revealed a significant increase in conceptual change, quality and self-efficacy for both the individual reading as well as the group discussion intervention. Concerning self-efficacy, there was not sufficient evidence to claim either intervention as more effective. However, cooperative learning activities showed significant benefits over individual reading activities in stimulating conceptual change. Concerning quality, individual reading activities were shown to be more effective than cooperative learning activities. Therefore, a combination of both activities is recommended for workshops on MOOCs.



*Keywords:* MOOC, Conceptual Change, Self-efficacy, Quality

## Table of Contents

## List of Tables

## List of Figures

# WHAT IS THE EFFECT OF A GROUP DISCUSSION COMPARED TO INDIVIDUAL READING ON DESIGNING AN OUTLINE FOR A MOOC?

## Introduction

### The Problem is in the Concept

Technology has an ever increasing influence on modern education. Thus, educators should acquire a certain level of competence with regard to the design of online courses. This requires skills in transforming face-to-face classes taught on campus into infinitely scalable online courses (Northcote, Gosselin, Reynaud, Kilgour, & Anderson, 2015). Ideally, educators realise the fundamental differences between designing a course for a lecture room and designing a course to be taught online. Understanding the concepts of online course design optimally supports educators not only in feeling self-efficacious about designing massive open online courses (MOOCs), but also in being able to design MOOCs of high quality.

Research supports that designing a MOOC differs vastly from designing a face-to-face higher education course (Guàrdia, Maina, & Sangrà, 2013; MacLeod, Haywood, Woodgate & Sinclair, 2014). However, getting this message across to academic staff is not simple (Gosselin & Northcote, 2013; Northcote et al., 2015). Educators frequently not only lack self-efficacy when it comes to designing MOOCs, they also have not undergone the conceptual change necessary for designing high-quality online courses (Northcote et al., 2015). Consequently, as confirmed in an evaluation of 76 MOOCs, the instructional quality of MOOCs is generally low (Margaryan, Bianco, & Littlejohn, 2014).

Due to the recent popularity of MOOCs, many educators are faced with the task of designing a MOOC. Experience has shown that educators need to undergo a conceptual change from designing traditional instruction to developing MOOCs (Northcote et al., 2015). Educators often experience difficulties in realising the difference between course design for online courses compared to course design for face-to-face courses (Gosselin & Northcote, 2013). This constitutes one of the main problems for individuals involved in MOOC design. Interviews with educators in various stages of the MOOC design process have revealed that they often lack experience with MOOCs and are left to design them without any guidance. Consequently, they feel insecure and are not aware of fundamental differences in designing for MOOCs (see study 1). This results in them designing MOOC outlines of insufficient quality.

### Design of Online Courses

This section illustrates the challenges that online course designers face, the differences in course design for MOOCs versus face-to-face courses or regular online courses, the self-efficacious beliefs of online course designers and lastly introduces two instructional design models.

**Challenges in online course design.** Online courses have enjoyed increasing popularity with learners in recent years. However, educators do not always share the enthusiasm of the learners, as they are faced with certain challenges when designing online courses as opposed to face-to-face courses. Most obviously, educators face technological challenges. They may lack experience with certain technologies and consequently may not feel comfortable using them (Bali, 2014; Jasnani, 2013; Northcote et al., 2015; Shepherd et al., 2007). Additionally, teaching online requires a different set of skills from the educator. For example, educators need to engage a much bigger, more diverse group of learners. However, online courses are new to many educators, so they lack experience with this format of instruction and have not had a chance to develop those skills yet (Northcote et al., 2011; Shepherd et al., 2007). An online course further requires the designer to consider many practical issues, such as which platform to use, how to deal with copyright issues of the material studied and how to provide certification to successful learners (Kopp & Lackner, 2014). Another challenge is that learning is asynchronous and participants are often located in different time zones, making synchronous learning activities difficult to put into practice (Kopp & Lackner, 2014). Lastly, the prospective audience of an online course is rarely known when the designing phase starts, requiring educators to design for unknown learners (Kopp & Lackner, 2014; Scagnoli, 2012). Thus, educators should be prepared for a wide range of prior knowledge in their learners, which makes the design of an informative and appealing course challenging (Scagnoli, 2012).

**MOOCs are different.** Online courses which are directed at a large, global audience and are open for anyone to join are called MOOCs (Massive Open Online Courses). Often, thousands of individuals join a MOOC to learn about a specific topic (Kopp & Lackner, 2014). Research shows that designing a MOOC differs vastly from designing a face-to-face higher education course (Guàrdia, Maina, & Sangrà, 2013; MacLeod et al., 2014). While MOOCs tend to be learner-centred and interaction between peers is often highly encouraged, face-to-face courses are in many countries teacher-centred lectures, which less frequently encourage peer-interaction (Bali, 2014). If educators design MOOCs in the same way they design their face-to-face courses, they restrict learners to a teacher-centred, passive learning environment, keeping them from maximizing their learning potential (Bali, 2014). Not only does the design of a MOOC differ from that of a face-to-face course, it also differs from the design of regular, non-massive online courses which exists on a restricted university website and is limited to a relatively low number of learners (Jasnani, 2013; Kopp & Lackner, 2014). Thus, when designing a MOOC, certain aspects should be considered: First, the course content is of importance. A wide variety of learning material should be integrated in a MOOC (Glušac, Karuović., & Milanov., 2015). Second, interaction among learners is an indication of a high quality MOOC. This can take place through discussion, collaborative assignments or peer- assessment (Bali, 2014; Glušac et al., 2015; Varonis, 2014). Additionally, the structure of MOOCs is of interest when evaluating their quality. MOOCs should be structured in a flexible way and allow for self-directed learning (Adamopoulus, 2013; Glušac, 2015). Fourth, MOOCs should encourage active learning and the application of newly acquired knowledge. Real-world problems and examples should be used to transfer knowledge to learners (Bali, 2014; Karlsson, Godhe, Bradley, & Lindström, 2014). Finally, a MOOC should start with an informative and engaging introduction, so learners are encouraged to actively participate from the beginning of the course (Varonis, 2014). Due to all these aspects that should be considered when designing a MOOC, educators frequently experience problems during the design phase: They often lack confidence and skills when it comes to the design of online courses and thus experience self-doubts and feelings of low self-efficacy (Northcote et al., 2015).

**Self-efficacy in online course designers.** According to Bandura, self-efficacy is defined as "one's beliefs in his or her ability to organise and execute the courses of action required to manage prospective situations." (Bandura, 1997, p. 3). Self-efficacy beliefs are domain-specific and formed by (1) mastery experiences, (2) vicarious experiences, (3) verbal persuasion and (4) physiological or affective state (Bandura, 1997, p. 50). Mastery experiences constitute the most powerful source of self-efficacy beliefs, as they pertain to similar situations someone has experienced before. If that situation been met successfully, self-efficacious beliefs increase. However, if the situation is viewed as a failure, a decrease in self-efficacy is likely, especially if there is no history of successful experiences (Bandura, 1997). Thus, one way to increase self-efficacy in learners is to ensure successful experiences when designing online courses. Vicarious experiences, like comparing one's own ability with that of a role model, allow one to judge one's own ability more accurately (Bandura, 1997). As individuals who are new to online course design were found to have fewer self-efficacious beliefs, it is important to support them. Studies showed that professional development programs can facilitate the development of self-efficacy in teachers of online courses (Gosselin & Northcote, 2013; Northcote et al., 2015). Self-efficacy is not only positively related to effectiveness, but also has a determining influence on task-effort and perseverance (Schunk & Pajares, 2005). Verbal persuasion refers to feedback to learners from observers. The more experienced the observer, in the eyes of learners, the stronger the influence on learners' self-efficacy beliefs (Bandura, 1997). Physiological and affective state refer to learners' emotions. When learning about the new skill of online course design, learners might exhibit physical and emotional symptoms of anxiety. As they become more comfortable with the task, their bodies relax. These physical and emotional changes can contribute to an increase in self-efficacious beliefs (Bandura, 1997). Following an instructional design can be useful in designing an online course effectively and thus develop more self-efficacy in one's own ability to teach online (Northcote et al., 2011). However, many design models take a linear approach to designing instruction, starting with analysis steps, followed by design and development steps and ending with an evaluation (Shelton & Saltsman, 2008). Yet, when it comes to MOOCs, a less linear approach might be advisable, as designing a MOOC is different from designing a regular online course and thus requires a different way of thinking (Kopp & Lackner, 2014). The process of conceptual change describes this mind shift.

## Conceptual Change

In contrast to the traditional view on learning, which assumes learning is the addition of more knowledge, the conceptual change view sees learning as the replacement of an old mental concept with a new, more accurate one. This is necessary whenever new information is inconsistent with the existing conceptual model. Consequently, the existing model needs to be adapted or replaced. This process is called accommodation (Posner, Strike, Hewson, & Gertzog, 1982). The following paragraphs explore conditions for conceptual change, how threshold concepts relate to online course design and how to facilitate conceptual change.

**Conditions for conceptual change.** Posner et al. (1982) stated four conditions which have to be met for conceptual change to occur.

1. Dissatisfaction with existing conditions: The learner must have experienced several instances in which the original concept failed to explain a phenomenon. Only if learners doubt the existing concept, are they susceptible to substantially altering or replacing their current mental concepts. However, considering to replace the old mental concept is a difficult approach. Consequently, a learner may choose a different option on how to deal with the discrepancy. Learners may reject the inconsistent information, evaluate them as irrelevant, mentally segregate the new information from the existing mental concept or attempt to alter existing concepts to assimilate the new information. In order to elicit dissatisfaction with the existing concept, learners need to (a) understand what exactly makes the information inconsistent with their current mental concept, (b) place importance on reconciling the conflicting information with the existing mental concept, (c) want to reduce inconsistencies in their set of mental concepts, and (d) fail at assimilating the conflicting information into existing mental concepts.
2. New concepts must be intelligible: The learner must understand the new concept and be able to create an internal representation of it. Often, using metaphors and analogies is helpful in supporting learners to grasp the new concept and stimulate further exploration.
3. New concepts must appear initially plausible: If a new mental concept does not provide a solution to the problems the old mental concept created, it is not likely to be adopted. Posner et al. (1982) state five ways which increase plausibility of a new concept: (a) the new concept is consistent with existing beliefs and assumptions, (b) the new concept is consistent with other knowledge, (c) the new concept is consistent with a learners' past experiences, (d) the learner can picture the new concept in a way that matches the real world as it is experienced, and (e) the new concept is able to solve problems the learner is aware of. Additionally, a new mental concept should be consistent with concepts in related fields. If a new concept solves the problem at hand, but contradicts other concepts in neighbouring fields, it is less likely to be adopted.
4. New concepts should be fruitful: A new mental concept should allow for new insights and be open to extensions, rather than solely be used for solving the problems created by the preceding mental concept (Posner et al., 1982).

Even though these conditions might suggest a linear order, Posner et al. (1982) clarify that this is not necessarily the case. Conceptual change is a complex process and requires a radical change in mental representations. Although the change it radical, it is not abrupt. Quite on the contrary, it seems to be a gradual process for most learners. They may only accept parts of a new mental concept first. However, these parts may well serve as the foundation for more adaptions at a later point. This process continues until the original concept has been replaced by a new concept entirely. Yet, the process of conceptual change is far from tidy and organized. It requires the learner to experiment, fail, reconsider and reflect. Sometimes new concepts are acquired, then learners regress to their old concepts until they encounter an inconsistency at a later point in time, due to which they review the new concept and ultimately adopt it (Posner et al., 1982).

**Conceptual change in online course design.** Online course design constitutes a new form of instruction for many educators and as such requires a conceptual change to their pedagogical beliefs (Northcote et al., 2015; Roehl, Reddy, & Shannon, 2013). In some literature such a substantial change is described as a threshold concept, comparable to a portal that opens up a novel way of thinking which was previously inaccessible (Meyer & Land, 2003). Threshold concepts are characterized as (a) transformative, as they entail a significant shift in perspective, (b) irreversible, as a threshold passed and consequently a perspective changed is difficult or impossible to unlearn, (c) integrated, as it underlines the interrelatedness of subjects, (d) frequently surrounded by new thresholds which open up once the previous threshold was passed, and (e) potentially troublesome, as they make conceptually difficult, counter-intuitive knowledge apparent (Meyer & Land, 2003). When designing online courses, educators may encounter theoretical and personal threshold concepts that are at odds with their personal and pedagogical beliefs, which may result in online courses of low quality and educators who experience low self-efficacy (Northcote et al., 2011). A number of threshold concepts identified by Northcote and colleagues (2011) are particularly relevant to this study:

1. Educators often do not realize the distinctive nature of the online learning environment, which does not imitate face-to-face education.
2. Different material is used for online courses to encourage interaction with and among learners.
3. Learning does not happen through passive absorption of knowledge, but through interaction and active knowledge construction
4. The content should be humanized by packaging it in a story and making the educator visible (Northcote et al., 2011).

Northcote and colleagues further found that educators worry most about pedagogical foundations for online teaching, even more so than they worry about technological issues. As many educators face the thresholds mentioned above, they should be targeted in all interventions designed to support educators in the design of online courses (Northcote et al., 2011). Northcote et al. (2015) accomplished that by administering professional development workshops, which significantly increased educators' self-efficacy and changed their threshold concepts over time. For the purpose of this research passing a threshold is considered a conceptual change. Effective strategies for stimulating conceptual change are introduced in the next paragraph.

**Stimulating conceptual change.** Over the years many pedagogical strategies have been developed to facilitate conceptual change in learners. Some of the most popular strategies include natural observation, simulations, models and analogies (Mills, Tomas, & Lewthwaite, 2016). However, for the purpose of this research, the focus will be on using cooperative methods to facilitate conceptual change. First, the educator should know which misconceptions learners are likely to possess (Bilgin, 2006). The misconceptions should then be used as a foundation for discussion, enabling the educator to moderate the discussion by asking questions pertaining to those misconceptions (Bilgin, 2006). In the case of MOOC design, one common misconception is that an online learning environment is a pile of material, so educators assume that uploading there course material is the main step in designing a MOOC (Northcote et al, 2015). In relation to this, many educators do not realise that learning in a MOOC happens through interaction and discussion (Northcote et al., 2015). Second, communication is an important factor for stimulating conceptual change

(Pea, 1993). Educators are thus encouraged to give learners opportunities for collaborative discourse, which facilitates conceptual change (Liu & Hmelo-Silver, 2009). Further, cooperative group work has been found to promote conceptual change (Bilgin, 2006). When working in a group, learners have the opportunity to share ideas and discuss their tasks. This allows them to make relations among concepts and detect misconceptions they themselves or other group members hold. Detecting differing concepts stimulates further discussion and learners are encouraged to argue concepts. Participating in this kind of cooperative activities facilitates conceptual restructuring (Bilgin, 2006). Detailed information about cooperative learning is presented in the next section.

## Cooperative Learning

This section introduces the method of cooperative learning, including important components of cooperative learning, its relation to academic achievement and practical implications for implementing cooperative learning.

**Traditional instruction versus active learning.** The traditional lecture- and reading-based instruction is familiar to all university students. Typically, this type of instruction involves receiving information by reading articles and visiting lectures. Traditional instruction is teacher-centred, as students have little influence on the content and how they want to learn it. Students assume a passive role in this instructional method. In active learning, on the other hand, students are more engaged. They are not merely on the receiving end when it comes to learning information, but can actively discuss it with their peers. This allows them to exchange ideas, relate new information to existing knowledge and construct their own knowledge. Active learning facilitates deep learning through constructive processes by using learner-centred methods of instruction, as opposed to traditional instruction, which often uses rote learning, expecting students to memorise facts, not going beyond the surface (Ritchhart, Chruch & Morrison, 2011).

There is substantial evidence for the many benefits of active learning. In his review, Prince (2004) state that active student engagement improves students' attitudes and their thinking and writing skills. Moreover, recall of information can increase with more active learning. One form of active learning is cooperative learning (Zayapragassarazan & Kumar, 2012).

**Cooperative learning.** Multiple learners working in a small group to discuss and learn about a problem, that is what is called cooperative learning (Slavin, 1991). The theory behind cooperative learning may be traced back to the work on social constructivism of Lev Vygotsky (Doolittle, 1997). Vygotsky believed learning to be closely related to interactions within ones' culture. Upon being confronted with a new experience or idea, learners need to internalise this new information and connect it with their existing knowledge, views and attitudes. This includes evaluating, adjusting and integrating the information based on past experiences, as well as possibly changing the old way of thinking (Doolittle, 1997). In this way, learning is active construction and integration of new information. Vygotsky believed that learners' potential for cognitive development is limited on the lower end by what they can accomplish by themselves, and on the upper end by what they can accomplish with the help of a more capable peer or teacher (Doolittle, 1997). He used the term *zone of proximal development* to define this area of potential cognitive development between what learners can accomplish individually and what they can accomplish with the help of others (Vygotsky, 1980). As a dynamic construct, the zone of proximal development shifts towards the upper end in response to the increasing cognitive development of the learner (Doolittle, 1997). Consequently, learners who presently need support with a task can accomplish said task by themselves sooner. The task has moved from being in the zone of proximal development, and therefore needing support, to being in the zone of actual development. Vygotsky stresses three important aspects for teaching within the zone of proximal development (Moll, 1992):

1. Whole, authentic learning activities. Only whole activities include all the complex aspects and their relations, which are lost when they are being broken down into parts. Additionally, authentic activities are more relevant and meaningful to the learner.
2. Social interaction. Social interaction between the learner and a more experienced peer or teacher is a central part to cognitive development. Specific cooperative activities should include opportunity for both the less experienced learner and the more experienced peer or teacher to share their perspectives. This interdependence allows the learner to actively construct knowledge.
3. Individual change. The purpose of instructing learners is to stimulate change in their culturally relevant behaviour. This is accomplished by changing their zone of proximal development through cooperative interactions with more capable others.

The process of aiding learners in accomplishing a task they need support with is called scaffolding. Cultural interactions can serve as scaffolds to support learners in reaching the next step of development. This serves two purposes: First, learners receive support in accomplishing a task they cannot complete on their own. Second, learners are developing knowledge to successfully complete the task on their own in the future (Hmelo-Silver & Azevedo, 2006; Pea, 2004; Sharma & Hannafin, 2007). Several studies have found scaffolding to increase learning outcome (Lin et al., 2012; Lin & Liu, 2014) and self-efficacy (Lin & Liu, 2014; Wu & Looi, 2013). Cooperative learning, as a form of scaffolding, seems tailor-made to ensure experiences within the learners' zone of proximal development while more capable others are involved (Doolittle, 1997).

**Components of cooperative learning.** Although contributing factors to successful cooperative learning are under debate, there has long been consensus among researchers about five factors that determine the success of cooperative learning activities: (1) positive interdependence, (2) individual accountability, (3) face-to-face-interaction, (4) small-group and interpersonal skills, and (5) group processing (Johnson & Johnson, 1994). *Positive interdependence* is achieved when learning cannot be accomplished alone, but each group member understands their reliance on the group to achieve a goal. (Johnson & Johnson, 1994). Not only group members are interdependent, learners and teachers are also depending on each other to move forward in their development. Without other individuals in a society, one would not learn. Each individual is dependent on others to move forward in their cognitive development (Doolittle, 1997). *Individual accountability* also determines a group's success. Each group member should feel responsible for their contribution to the group's learning. Ideally, this would result in each group member's zone of proximal development to shift, so that they are able to execute a task, which they can only do in a group today, by themselves tomorrow (Doolittle, 1997). *Face-to-face interaction* refers to the time the group members should spend discussing ideas, challenging reasoning and encouraging, supporting and teaching each other (Johnson & Johnson, 1994). According to Vygotsky (1980) social interactions form the basis of learning, as they open up new zones of proximal development through which cognitive growth and learning are possible (Doolittle, 1997). *Small-group and interpersonal skills* are important for successful cooperative learning. Without social skills, there can be no effective communication in the group. Vygotsky referred to these skills as sociocultural signs and tools for interacting with others (Vygotsky, 1980), so they can be seen as a prerequisite for development (Doolittle, 1997). *Group processing* refers to the group's evaluation of their own actions and progress. It includes the group's reflection on which changes to make to reach the group's goal and aims to increase each members' productiveness (Johnson & Johnson, 1994). According to Vygotsky (1980), individuals are not solely responsible for their own development, but society, i.e. teachers and group members, bear responsibility for the learners' progress within their zones of proximal development as well (Doolittle, 1997). Group processing allows teachers, learners and group members to assess where the task or instruction fits into their zone of proximal development. If the task is below the lower end of the zone of proximal development, it has already been mastered and is likely to bore the learner. However, if a task is above the upper end of the zone of proximal development, it is too difficult for the learner, who will likely become frustrated. Ideally, the task should be within the zone of proximal development for each group member, to even the road for cognitive development (Doolittle, 1997).

Provided that above mentioned five factors are accounted for, cooperative learning has been proven to help learners to develop high-order thinking skills, improve their motivation as well as their interpersonal relations (Prince, 2004; Slavin, 1985). Further, cooperative learning is beneficial for learners' psychological wellbeing, self-esteem and self-efficacy (Li & Lam, 2013; Prince, 2004). Most importantly, cooperative learning is positively related to academic achievement.

**Cooperative learning and academic achievement.** Literature agrees on the positive effects of cooperative learning on academic achievement (da Costa & Galembeck, 2016; Li & Lam, 2013; Lou, Abrami, & d'Apollonia, 2001; Prince, 2004; Terenzini, Cabrera, Colbeck, Parente, & Bjorklund, 2001). Different perspectives have been put forward to explain this finding. The cognitive development perspective states that learners in heterogeneous learning groups, which include learners both of low and of high ability, receive and provide information. In these groups, learners of high ability are challenged by the questions of low ability learners. They have to explain their knowledge, possibly in different ways, to help their peers understand. This aids the construction and organisation of knowledge and helps high ability learners to relate the new information to existing knowledge (Lou et al., 2001) Additionally, with each instance that high ability learners explain information, they orally rehearse it, which helps retain it (Johnson & Johnson, 1994; Lou et al., 2001). Learners of low ability, on the other hand, benefit from cooperative learning groups as discussion is likely to raise cognitive conflicts in their understanding. They can ask questions and use the explanation by their peers to identify and correct misconceptions, restructuring their new knowledge and relate it to existing knowledge (Lou et al., 2001). Medium ability learners benefit least from heterogeneous learning groups. However, they still receive feedback from their peers and experience different perspectives while discussing the new information (Lou et al., 2001). Relating this reasoning to Vygotsky, the task carried out in a group of learners should be located on the upper limit of the zone of proximal development. Learners do not need to be able to perform the task alone, but be successful carrying it out in a group. The task will then move from the zone of proximal development into the actual development and the learner will be more likely to able to perform it individually in the future (Li & Lam, 2013).

In contrast, the cognitive elaboration perspective views cognitive restructuring as the single most important reason for academic achievement in cooperative learning. It assumes that learning groups elicit the need for elaboration in learners and thus advance the process of cognitive reconstruction (Li & Lam, 2013).

The social cohesion perspective states that cooperative learning is effective because it thrives off the social cohesion of the group. Academic achievement increases with social cohesion of the group (Li & Lam, 2013). Learners are said to place importance on the group and its members and thus help the group to reach its goal by supporting each group member. Learners identify themselves with the group and want the group to succeed in order to feel successful themselves (Li & Lam, 2013).

The motivational perspective assumes that all actions of individual learners are driven by self-interest to reach the group goal. This perspective views task motivation as the most powerful reason for group members to invest in their time and knowledge towards reaching the group goal (Li & Lam, 2013).

**Implementation of cooperative learning.** Not all cooperative learning is equally effective. For this reason, multiple aspects should be considered when implementing cooperative learning strategies. First, group size is of importance. The benefits of cooperative learning are improved in small groups. Groups of two have been found most effective when learning in front of a computer is involved. For classroom learning, groups of three to five learners are ideal (Lou et al., 2001). Second, more cooperative learning is not always better. It has been found that a medium amount of group work has the highest effect, while a low and high amount of group work are less effective (Prince, 2004). Third, group composition is a critical factor. Heterogeneous groups, which include learners of varying ability levels, have been found most effective for low-ability learners, as they benefit from receiving explanations from high-ability learners (Lou et al., 2001). Medium-ability learners learn most effectively in homogeneous groups with learners of equal ability. In heterogeneous groups, they may neither share nor receive explanation, while in homogeneous groups, they have similar expectations and group goals. Further, they do not need to accommodate the group's pace to the low-ability students (Lou et al., 2001). High-ability learners benefit from both homogeneous and heterogeneous groups. In heterogeneous groups they thrive on the opportunity to give explanations to learners of medium or low ability. In homogeneous groups they share similar expectations and can maintain a fast pace of learning and discussion without accommodating learners of lower ability (Lou et al., 2001).

Vast differences between face-to-face courses and online courses have become apparent. Left to their own devices, educators do not realise these differences and design MOOCs of insufficient quality. Further, they experience feelings of insecurity with regard to MOOC design (Northcote et al., 2015). Educators should be supported in acquiring the new concept of online course design before they are asked to design a MOOC. This conceptual change is required to ensure the design of high-quality online courses as well as a high level of self-efficacy in educators. In order to facilitate conceptual change professional development measures should be taken (Northcote et al., 2015). As cooperative learning has been proven to constitute an effective way of learning in a variety of domains, professional development activities will be based on this (da Costa & Galembeck, 2016; Prince, 2004). Typically, one of the first steps in designing a MOOC is to create an outline. For this reason, this research will focus on the design of MOOC outlines.

It is the intent of this research to investigate if cooperative learning will facilitate conceptual change in educators designing outlines for a MOOC. Further, it examines if cooperative learning increases the quality of the MOOC outlines designed, and the self-efficacy experienced by educators. This will be accomplished by comparing two different intervention workshops, one of which will incorporate cooperative learning activities, while the other will consist of individual work.

## Context

This section highlights the scientific and practical relevance of the study and states the research questions.

### Relevance

**Scientific relevance.** The content of this research is novel in several aspects. First, conceptual change is mostly studied from a student perspective. Very few studies investigate conceptual changes in educators. The same holds true for MOOCs. Most research examines MOOCs from a student perspective, while only a minority of research focuses on the experience of the instructors. Even fewer studies take a closer look at the design aspects of MOOCs with regard to the designer's feeling of self-efficacy. While the effect of cooperative learning on conceptual change has substantial support in literature, the idea of using cooperative learning strategies to increase the quality of MOOC outlines and the self-efficacy of educators is, as of yet, unexplored. Thus, this study breaches new terrain in applying a cooperative learning technique to (a) facilitate conceptual change in educators who design online courses, (b) increase the quality of the course they design and (c) increase their feeling of self-efficacy with regard to online course design.

**Practical relevance.** If the workshop would support educators' self-efficacy, process of conceptual change and the quality of their MOOC outline, this study would be relevant on multiple levels. First, an increase in self-efficacy would allow educators to be more confident in their designing skills. Second, the conceptual change would support educators in producing high-quality online course material, going above and beyond MOOC outlines, as they acquired the instructional concept of online course design. Third, the university's MOOC team would benefit from educators submitting MOOC outlines of higher quality. Less revision would be required and both educators and MOOC team members would be able to work more efficiently while experiencing less frustration. The time commitment required from the educators would decrease. This might increase the acceptance of MOOCs on a broad level and lowers the threshold for teachers contemplating about becoming involved in designing a MOOC. Fewer rounds of reviews required would also save financial resources of the university, while the overall instructional quality of its MOOCs would increase. Potential savings in the development of each single MOOC could be reinvested to develop more MOOCs. Overall, this experiment is justified by a variety of potential benefits.

### Research Question

**Research question.** As a first step, this study aims to investigate if there is a practical relevance for training in MOOC design, like literature suggests (Northcote et al., 2011). For this reason, a preliminary study was conducted, which explored the following research question:

- Is there a need for training on MOOC design in the eyes of educators who are involved in different stages of the process?

We hypothesize this question to be answered positively.

Consequently, this study aims to explore how such training should be designed. For this reason, it compares cooperative learning with individual learning. To accomplish this, elements of cooperative learning are used. In particular, this study compares learners who take part in a group discussion with learners who read an article. Specifically, this study investigates the quality of MOOC outlines they produce, their self-efficacy and their conceptual change. The following research questions are formulated:

WHAT IS THE EFFECT OF A GROUP DISCUSSION COMPARED TO INDIVIDUAL READING ON DESIGNING AN OUTLINE FOR A MOOC?

- What is the effect of a group discussion compared to individual reading on designing an outline for a MOOC?
    - What is the effect on the quality of the MOOC outline?
    - What is the effect on conceptual change in designers developing MOOC outlines?
    - What is the effect on self-efficacy in course designers developing MOOC outlines?

The following hypothesis is formulated:

Using a group discussion for designing MOOC outlines improves their quality, stimulates conceptual change, and increases self-efficacy in course designers compared to individual reading.

**Study 1**

The following section pertains to the preliminary study which preceded the main study. Its purpose was to establish the practical relevance for the main study. Further, it served as orientation and guidance for the researchers in developing the main study.

**Method**

**Research Design.** This research is comprised of interviews with individual educators. Therefore, it is of qualitative nature. Interview questions were chosen in collaboration with an expert on MOOC design. Afterwards, interviews and notes were analysed to evaluate if there is a need for training on MOOC design.

**Respondents.** A total of 20 possible respondents were approached via email and asked for a 30-minute interview. The contacts were provided by the university's MOOC expert. Ten educators agreed to be interviewed. Due to vacations and illness, two respondents had to cancel their interviews. The remaining eight educators were interviewed, four of which were male. All participants are employed by the University of Twente. In order to cover the topic from multiple perspectives, participants were chosen based on the stage of MOOC design they were at at the time of the interview. One participant was still in the planning stage and did not have a topic narrowed down yet, two participants were in the process of designing their MOOC, and the remaining five participants had finished the design stage and were running their MOOCs. Six different MOOCs were covered by interviewing these eight participants. Topics included Supply Chain Management, E-health, Geo-health, Ultrasound and Nanotechnology. Amongst the participants were lead educators (3), co-educators (2) and coordinators (3) of MOOCs, which allowed for different perspectives to be covered.

**Instrumentation.** The interview consisted of 13 questions. The first five questions were related to the design of online courses, i.e. *"How was your designing experience?"*. They aimed at gaining insights into educator's thoughts and feelings during the design process. Following, there were four questions about conceptual change, i.e. *"In your opinion, what is the main difference between classroom teaching and teaching an online course?"*. These questions were asked in order to assess whether educators experienced a conceptual change. Finally, there were four questions about educators' motivation, i.e. *"What was your motivation to design an online course?",* in order to estimate if educators are sufficiently motivated to invest their time in learning about online course design. After each question, educators had all the time they needed to answer it. The interviewer took notes in addition to recording the interview.

**Data analysis.** Recordings from the interviews were analysed with the help of the notes taken during the interview. Answers to each question were analysed separately according to the design stage educators were at. Responses were summarized and reduced to their key points. Finally, responses were compared.

**Procedure**

After introductions were made and permission to record the interview was granted, questioning commenced. Although the interview was scheduled for about 30 minutes, it took from 20 to 45 minutes, depending on how much the interviewee had to share. Interviews were conducted in an informal atmosphere at participant's offices or in meeting rooms. The researcher took notes during the interview. After answering the last question, the researcher gave a short description of the purpose of the interview and the planned study. If participants did not have any further questions, they were thanked and the interview was concluded.

**Results**

Results are divided into three categories of interest: Design experience, conceptual change and motivation to learn. The categories derived from the three blocks of questions asked.

**Design experience.** Participants of the preliminary study stated that they did not use design guidelines for the most part. Rather, they based their designs on intuition or personal experience. Half of the participants looked at other MOOCs to get ideas. The interviews revealed that most educators experienced the design phase as long and challenging. The roles and their tasks were often unclear and they underestimated the workload and time pressure:

*"I now understand why it takes months and months to make a movie. That was really an eye opener."* (supply chain management, educator),

*"We are kind of in a rushing phase … If we had more time we would have done it in a different way."* (Nanotechnology, lead educator).

Some participants did not feel confident and secure with the task of designing a MOOC. They experienced uncertainty and described the lack of requirements as troublesome:

*"Basically, we did not have an idea what was required … for us, it was a big black box."* (Supply chain management, educator).

However, there were also participants who enjoyed the experience and characterized it as *"fun, creative and collaborative"* (E-health, coordinator).

**Conceptual change.** When asked what they struggled with most in online teaching, educators mentioned the assessment of student understanding, the lack of discussion, the limitations in time and content depth and the time it takes to produce a course as the main problems. All participants found teaching online challenging and fundamentally different than teaching face-to-face courses. With regard to this conceptual change, four participants stated that they were aware of fundamental differences between MOOCs and face-to-face courses before they started designing their MOOC:

*"Before I began, I knew there was a difference, but it was difficult to translate this into the right materials."* (Geo-health, lead educator).

The other four participants experienced a conceptual change during the design process. Three of them described the conceptual change as a gradual process:

*"It is gradual. I mean, I knew it was different, but there were a few more aspects that I did not expect. I thought I could just use my course material; it would just be a matter of practicality to convert that to an online course. But I can tell you right away, that is not the way it works. It doesn't!"* (Nanotechnology, lead educator).

However, one participant characterized the conceptual change as a sudden realisation: *"It was quite sudden, yes."* (Nanotechnology, coordinator).

When asked about what triggered the conceptual change, participants mentioned time restrictions, level of content depths, organizing the structure, feedback and discussions with the team as important factors contributing to the changed perception.

*"Most of them (the educators) realized it when they started writing."* (Nanotechnology, coordinator).

**Motivation to learn.** When asked if they would have been willing to participate in a course on MOOC design, all eight participants answered positively. However, five participants stressed the limited time at their availability. For them, any course would have to be short and easily accessible:

*"I would find it very interesting, but I am not sure if I would personally find the time to do that… it could be helpful though."* (idea phase, lead educator).

Three participants stated that they would prefer a very practical approach, that does not include passive listening to lecture, but active experience with designing. One of them reasoned that many professors think of themselves as experts not only in their domain of expertise, but also in how to teach in their field, which might not always be accurate. These professors would, according to the participant, benefit from learning-by-doing:

*"Best thing is to just do it … You learn virtually most by just doing it."* (supply chain management, educator).

## Discussion

Results from the interviews confirm the need for training on MOOC design for several reasons:

1. Participants did not use any guidelines in designing their courses. Rather, they relied on intuition and experience. However, basing a design on intuition is potentially problematic as intuitions can change with the conceptual change that half of the participants experienced during the design process. Additionally, feeling experienced in designing face-to-face courses might be confused for experience in MOOC design, which are two very different skills. On the other hand, some participants described how challenging and unclear the design process was and that they were not confident in their course design skills. As confidence and self-efficacy seem to be an issue for educators, it was decided to include this construct in the main study. Training on MOOC design should aim to increase self-efficacy in educators, as well as provide guidelines to base the design on.

2. Participants listed a variety of difficulties they experience when teaching online. Considering these difficulties during the design process allows educators to set up the MOOC in a way to counteract these problems to a certain degree. For example, the problem of lack of discussion could be reduced by implementing more discussion activities in the MOOC. Knowing about these potential problems

during the design process would help prevent them and make running the course more enjoyable and efficient for educators. Training on MOOC design should incorporate information which helps to prevent common difficulties in teaching MOOCs.

3. Half of the participants did not realise the difference between designing face-to-face courses and MOOCs until they were in the design process. If this conceptual change occurred before the design process was started, participants would not be as surprised by how much time the design process takes. They would understand that designing a MOOC does not mean simply uploading the existing material and making it available to a broader audience. Knowing this from the beginning would save participants frustration and also help to guide their designs into the right direction from the beginning. Training on MOOC design should include activities that facilitate conceptual change.

Overall, training would clearly benefit educators when designing a MOOC. However, results from the interview also revealed some aspects that should be considered when implementing such training. First, educators are often involved in different research projects and teaching activities. They struggle to find the time to produce content for a MOOC in their busy agendas. Additionally, participants expressed the wish for learning-by-doing and practical experiences. They want to apply their new knowledge right away. Consequently, training on MOOC design should be concise, while at the same time providing opportunity for practical experiences and promoting the application of new knowledge. Ideally, it would provide educators with a product which they can use for their actual MOOC, so it is perceived as time well spent.

In summary, training on MOOC design should

- Increase self-efficacy in educators
- Provide guidelines to base the design on
- Include information which helps to prevent common difficulties in teaching MOOCs
- Include activities that facilitate conceptual change
- Be as concise as possible
- Provide as much opportunity to practice as possible
- Lead to a product for educators to keep working with
- Lead to MOOCs of higher quality

For the main study, a workshop was developed that incorporates all of the above design elements. In order to make the workshop easily accessible, all elements are incorporated in to one workshop session. The workshop includes three activities of hands-on practice, a short presentation, and either a reading activity (article group) or a discussion activity (discussion group). See table 2 for details on how each design requirement was translated into elements of the workshop. A detailed description is provided under procedure.

Table 1
*Design Requirements and Consideration in Workshop*

| Design requirement | Consideration in workshop |
|---|---|
| Increase self-efficacy | Provide experience though practice activities (design MOOC outline and revise MOOC outline) |
| Provide design guidelines | Presentation, information in article/discussion |
| Information to prevent difficulties in teaching MOOC | Information in article/discussion |
| Facilitate conceptual change | Information in article/discussion, changing face-to-face outline in MOOC outline |
| Be concise | One-instance workshop |
| Provide opportunity to practice | Three practice activities of designing MOOC outline |
| Provide product to work with | Final MOOC outline can be used for future MOOC |
| Lead to MOOC of higher quality | All workshop activities are aimed at increasing quality of MOOC |

## Study 2

After establishing the need for training and the requirements of such training on MOOC design in the preliminary study, the main study investigated the research question of what the effect of a group discussion is compared to individual reading on designing an outline for a MOOC. In order to study this question, two different workshop interventions were designed. The workshops were carried out at the University of Twente, Netherlands, in cooperation with the Technology Enhanced Learning and Teaching department.

### Method

In this section the research design of the second study is elaborated, followed by information about the respondents and instrumentation.

**Research Design.** This research is a quasi-experimental study, which aims to determine the effects of two different interventions. Both interventions are workshops. A group discussion is part of one workshop, while the other uses a written article. Participants fill in questionnaires and create three MOOC outlines during the workshop, which are then compared using a pre-test post-test design. If results show higher quality MOOC outlines for either condition, this serves as an indicator that that intervention is indeed effective in the design of MOOC outlines. An increase in self-efficacy of teachers in either condition serves as an indication for the effectiveness of that intervention with regard to self-efficacy. Quantitative data is being collected for self-efficacy to allow for objective comparisons between the two conditions. MOOC quality is evaluated qualitatively by scoring the MOOC outlines based on a scoring rubric. Scores of participants in both conditions are compared. Comparisons of quality scores and self-efficacy scores in both conditions allow answering each of the research questions. Further, a number of open-question questionnaires are administered to gain insight in participants' concepts of MOOCs and to determine if they experienced conceptual change.

The group discussion and article both had the same focus. Out of the many differences between MOOCs and face-to-face courses, a few aspects have been selected based on their expected relevance for conceptual change. This decision was made in close collaboration with experts in MOOC design and based on misconceptions with regard to MOOCs which are commonly held by educators, such as that uploading course material is the main part of creating a MOOC (Northcote et al., 2015). The following differences were the focus of both interventions: Social learning in MOOCs, structure of MOOCs, big questions and storylines in MOOCs, learning activities in MOOCs (Adamopoulus, 2014; Bali, 2014; Glušac et al., 2015; Karlsson et al., 2014; Varonis, 2014).

**Respondents.** Convenience sampling with random assignments to conditions was used. Students of the UT were approached through an invitation to a workshop on online course design posted to the SONA website. Additionally, flyers and social media were used for advertising. Participants were picked at random. As there were multiple timeslots for workshops offered, the date that participants signed up for determined the condition they were assigned to.

Although this study focussed on designing MOOCs for higher education, the population of focus was students, as it proved impossible to acquire the necessary number of teachers to conduct this experiment. Sampling criteria was purposefully kept broad, in order to allow a wide variety of students to participate. The only requirement was that participants are able to understand, read and write English.

A total of 46 students signed up for the study, from which 42 actually participated. Four students did not show up. There were 22 participants in the article condition and 20 participants in the discussion condition. Discussion group size ranged from four to six students. All participants (55% males) were students at the Universty of Twente, most of them studied psychology or communication science. The mean age was 20 (age range from 18 to 26). All participants were naïve to the purposes of the study. They received three hours of test subject credit for participating. All of them were novices in the field of course design. Three participants had teaching experience, four participants had designed courses before and three had participated in MOOCs. Only one participants had experience in developing instructional videos.

**Instrumentation.** An experiment was conducted as part of this research. The setup involved two groups, one of which used a group discussion while the other used an article. Quality of MOOC outlines, conceptual change and self-efficacy were measured. The instruments used are elaborated below. Questionnaires are presented in the appendix.

***MOOC Concept.*** In order to measure the perceived difference between MOOCs and face-to-face courses, participants were given an empty table which asked them to fill in three main differences and elaborate on them. The table consists of three columns with the headlines "Difference", "MOOC/online course" and "Face-to-face/traditional course" and three empty rows for participants to fill in.

***MOOC Questionnaire.*** A three-item questionnaire was administered to measure participants' perception of a MOOC. The open question items asked for participants' opinions on the main purpose of a MOOC, the structure of a MOOC and videos in MOOCs.

***Design Experience Questionnaire.*** This five-item open-question questionnaire was administered at the end of the workshop and asked participants for their experience while designing the MOOC outline. Specifically, participants were asked to evaluate if they have created a good outline, if they encountered any difficulties in transferring the face-to-face outline into a MOOC outline, if they realized a fundamental difference between face-to-face courses and MOOCs and what activity during the workshop they found particularly helpful.

***Change Reflection Questionnaire.*** This questionnaire contained two open questions. Participants were asked to reflect on the changes they have made to their MOOC outlines and on what inspired them to make those changes.

The MOOC Concept, the MOOC questionnaire, the design experience questionnaire and the change reflection questionnaire were all administered to evaluate conceptual change in how participants perceive MOOCs. All these questionnaires had been developed based on information gathered in preliminary interviews with teachers who were in different phases of the MOOC design process. Some of them had finished the process and already ran their MOOCs several times, others were in the middle of the design process, while some were still in the planning phase. Their experiences and feedback went into developing these questionnaires. After that, all questionnaires were checked by an expert in the field of MOOC development, and further adapted based on his feedback. Pilot studies were run with the final questionnaires, before they were used in the experiment.

***Quality.*** The quality of MOOC outlines was assessed using a scoring rubric that has been developed specifically for this experiment. After reviewing a number of existing instruments, it has been concluded that none fit the purpose of evaluating a MOOC outline in such an early state of development. For this reason, it was decided to design a new rubric based on Merrill's 5 Star Instructional Design Rating (Merrill, 2001) and the Distance Education Learning Environment Survey (DELES) instrument (Walker & Fraser, 2005). Both these instruments have been chosen because they are grounded in educational theory and have been applied successfully in the evaluation of online courses (Cropper, Bentley, & Schroder, 2009; Walker & Fraser, 2005). Merrill's 5 Star Instructional Design Rating consists of five scales: Problem, activation, demonstration, application and integration. Each scale contains three items, which can be awarded bronze, silver or gold level, depending on the level of depths in which they are covered (Merrill, 2001). The DELES consists of six scales with a total of 34 items. The scales are instructor support, student interaction & collaboration, personal relevance, authentic learning, active learning and student autonomy. Items are answered on a 5-point Likert scale ranging from "*Never*" to "*Always*" (Walker & Fraser, 2005). The authentic learning scale and the student interaction & collaboration scale were most influential in the development of this rating rubric. However, these two instruments only served as a starting point, as they did not cover aspects related to introduction and structure. Development of the rubric used in this study mainly took place in close collaboration with an expert in the field of MOOC development. A pilot study was conducted to gain further information. The final version included 21 items in total, which have been divided into five categories. Introduction (four items), content (six items), interaction (three items), transfer of knowledge (four items) and structure (four items) were assessed. Additionally, eight of those items (from all of the categories) served as indicators of conceptual change by judgement of a MOOC expert. A separate score was calculated for this. The categories *content*, *interaction*, *transfer of knowledge* and *conceptual change* of the presented rating rubric were most heavily influenced by the 5 Star Instructional Design Rating and the DELES. Consequently, this instrument was grounded in theory, as well as designed based on experience of an expert in the field.

In the following, scoring categories are described in more detail.

*Introduction.* This category included four items. Points were awarded for (1) an introduction to the course, (2) mentioning learning objectives, (3) asking a Big Question and (4) mentioning an introduction of the educators. The item "The outline mentions a Big Question" was considered an indicator of conceptual change.

*Content.* This category consisted of six items. Points were awarded if the outline includes (1) relevant media, (2) examples, (3) a variety of learning activities, (4) opportunities for discussion, (5) opportunities for reflection and (6) a form of assessment. The items about "a variety of learning activities" and "opportunities for discussion" were considered indicators of conceptual change.

*Interaction.* This category included three items. Points were awarded if the outline (1) includes opportunity to ask questions, (2) includes opportunity for feedback to the learner and (3) promotes social learning. The items about "feedback to the learner" and "social learning" were considered indicators of conceptual change.

*Transfer of Knowledge.* This category included four items. Points were awarded if the outline (1) includes an opportunity for learners to apply their new knowledge in a different situation, (2) includes activation of prior knowledge, (3) includes a real world problem and (4) provides an opportunity to practice. The item "The outline includes a real world problem" was considered an indicator of conceptual change.

*Structure.* This category consisted of four items. Points were awarded if (1) the outline provides a clear structure, (2) the outline includes a storyline or central theme, (3) the duration of each activity is appropriate and (4) the outline matches the task (MOOC or traditional course). The items "the outline includes a storyline or central theme" and "the outline matches the task" were considered indicators of conceptual change.

*Indicators of conceptual change.* The eight items mentioned above were scored for indicators of conceptual change.

***Self-efficacy.*** The short form of the Occupational Self-Efficacy Scale (OSS-SF) (Rigotti, Schyns & Mohr, 2008) was used to assess participants' self-efficacy at two different times during the experiment. This instrument measured self-efficacy on an ordinal level. Wording of items was adapted to "…in course design" instead of "…in my job". The scale contained six items. This instrument proofed good reliability (Cronbach's α = .90) and construct validity (Rigotti et al., 2008). Figure 1 represents a sample item.

**Whatever comes my way in course design, I can usually handle it.**

| Not at all true | | | | Very true |
|---|---|---|---|---|
| o | o | o | o | o |

*Figure 1.* Sample item of Occupational Self-efficacy Scale.

To ensure uniformity, the scale of this questionnaire was adapted to 5-point Likert scale ranging from 1 (*Not at all true*) to 5 (*Very true*). Instructions and wording of items were adapted to fit the context of this study. All instruments were chosen based on their extent. Length of questionnaires was kept to a minimum to alleviate the stress and time burden on participants. All measurements were at an individual level.

**Procedure**

This section describes in detail how the intervention was carried out. First, the general procedure of the workshop, which is identical for both conditions, is described. Following, the activities which happened only in the article condition are illustrated. Next, the procedure of the discussion condition is described.

**General Procedure.** Data was collected during a workshop on MOOC design. First, participants filled in demographic questions. After this, they were asked to design an outline for a face-to-face lesson on a topic of their choice (outline 1). For this, they were provided a template file. Ten minutes into the task they handed in their outlines and listened to a 5-minute presentation on MOOCs, which contained a walk-through of a FutureLearn MOOC. After the presentation, the first round of questionnaires was filled in: MOOC Concept, MOOC Questionnaire and OSS-SF. Next, participants were asked to change their face-to-face lesson outline into a MOOC outline, using the template from outline 1. After 10 minutes they were asked to hand in their work (outline 2) and start either the group discussion or working with the article. See below for details on the tasks in both conditions. After 20 minutes spend either working on the article or discussing, participants were instructed to revise their MOOC outlines and hand in the final version (outline 3). The second round of questionnaires concluded the experiment. Participants filled in MOOC Concept, MOOC Questionnaire and OSS-SF a second time. Additionally, they answered the design experience questionnaire and the change reflection questionnaire. Finally, participants were thanked and debriefed. The duration of the whole workshop was approximately 90 minutes. Table 2 provides a chronological overview of what participants handed in.

Table 2

*Chronological Overview of Workshop Procedure Including Material and Data Collected*

| Content of workshop | Handed in |
|---|---|
| Demographic questions | Demographic data |
| Face-to-face lesson outline | Outline 1 |
| Presentation | - |
| Questionnaires round 1 | MOOC Concept, MOOC Questionnaire and OSS-SF |
| MOOC outline | Outline 2 |
| Group discussion / Article | - / Notes article |
| Revised MOOC outline | Outline 3 |
| Questionnaires round 2 | MOOC Concept 2, MOOC Questionnaire 2, OSS-SF 2, Design Experience Questionnaires, Change Reflection Questionnaire |

As both workshops included exercises in which participants design their own MOOC outlines, there were whole, authentic learning activities that are meaningful included in each condition (Moll, 1992).

During the workshop only the group of participants and the researcher were present. Participants signed an informed consent form. They were informed that their data is used for research purposes, saved anonymously and treated confidentially. Further, they were informed about their right to withdraw from the experiment at any moment without giving reasons. Approval from the Ethical Committee had been granted.

In order to meet high quality and validity standards, both workshops were as similar as possible with regard to the location, time of day, presentation and procedural aspects. However, due to the social character of a workshop, differences in conditions were unavoidable. The researcher realized this threat to construct validity, but felt that a workshop was still the best way to simulate the process of designing a MOOC, as it included the social component integral to the design process.

**Article Intervention.** At the beginning of this 20-minute section, participants were send a link to an article about MOOC design. The article was 1750 words long and contained design guidelines published by FutureLearn, a UK-based MOOC platform. The article is presented in the appendix. Participants were instructed to take notes using a template, which they sent in at the end of the task. They were made aware that they may not have enough time to read the whole article, but should skim it for useful information for their MOOC outlines. After 20 minutes, participants handed in their notes and moved on to reviewing the MOOC outlines.

**Group Discussion Intervention.** Discussion groups included four to six participants. The principles of cooperative learning were followed as much as possible (Johnson & Johnson, 1994; Moll, 1992). The limited time frame and ad hoc composition of groups did not allow for covering all principles to the fullest extent.

According to the distinction made by Panitz (1996), discussion groups worked cooperatively, not collaboratively, as the task was content-specific and participants designed individual outlines. The instructor, or in this case the researcher, was in charge of the discussion and remained in control of the task, which is a distinction of cooperative learning (Panitz, 1996).

For the group discussion every participant was given a statement about MOOC design. One by one, participants then read out the statement and argued why they agreed or disagreed with it. Other participants were asked for their opinion on the topic, so that a discussion developed. The researcher guided the discussion by asking provoking questions if it slowed down or keeping it on topic if it threatened to go off track. She also monitored the time, so that the discussion was concluded after 20 minutes and all important topics had been covered. Below is the list of statements that have been provided as discussion prompts. The discussion prompts are based on common misconceptions that educators may hold (Bilgin, 2006; Northcote et al., 2015).

*A video in a MOOC should be at least 5 minutes long to go into enough detail*
*A MOOC should focus around a big question*
*A MOOC should primarily help individuals to learn by themselves*
*A video of a lecture is a good foundation for a MOOC*
*A MOOC should follow a linear structure*
*A MOOC should primarily facilitate the exchange of ideas*

**Data Analysis**

All data analysis was conducted in close collaboration with an expert on online course design, who provided feedback and reviewed coding schemes and quality rubrics. The data analysis for this research was generated using the Real Statistics Resource Pack software (Release 5.1) for Microsoft Excel.

**MOOC Concept.** Answers from all participants were analysed and reduced to a keyword for each of the three rows in the table separately. This lead to up to three keywords mentioned by each participant. Keywords were then assigned to one of six categories derived from data, which are *convenience*, *interaction*, *social aspects*, *organization*, *content* and *methods* (see table 3). It was counted how frequently each category was mentioned for each condition separately for both pre-test and post-test.

Table 3

*Categories and Associated Keywords for Main Difference*

| Category | Keywords |
|---|---|
| Convenience | Flexibility, availability, requirements, self-directedness |
| Interaction | Interaction, staff, relationship with educator |
| Social aspects | Social learning, global community |
| Organization | Structure, pace, scale |
| Content | Exercises, media, content, assessment |
| Methods | Motivation, way of learning |

**MOOC Questionnaire.** Answers from all participants were analysed and reduced to one or more keywords for each question separately. Participants mentioned up to four keywords for the item *main purpose*, up to six keywords for the item *structure* and up to three keywords for the item *good video*. If a participant mentioned multiple keywords, all of them were scored. For the *main purpose* item, frequency of keywords was counted and each keyword was treated as a stand-alone category, which were *knowledge*, *availability*, *flexibility*, *self-directedness*, *exercises* and *social learning*.

For the *structure* item, keywords were assigned to one of three categories derived from data, which were *content*, *extent* and *progression* (see table 4). Categories were created in close collaboration with an expert in the field of online course design. It was counted how frequently each category was mentioned for each condition separately for both pre-test and post-test.

Table 4

*Categories and Associated Keywords for Structure*

| Category | Keywords |
|---|---|
| Content | Information |
| | Exercises |
| | Reflection |
| | Discussion |
| | Evaluation |

| | |
|---|---|
| | Variation |
| | Narrative |
| Extent | Strict |
| | Concise |
| Progression | Visible progress<br>Linear progress<br>Flexible progress |
| | Easy to difficult |
| | Broad to specific |

For the *video* item, keywords were assigned to one of three categories derived from data, which were *length*, *content* and *presentation* (see table 5). Categories were created in close collaboration with an expert in the field of online course design. It was counted how frequently each category was mentioned for each condition separately for both pre-test and post-test. If a time range was given for length, the lower length was scored.

Table 5

*Categories and Associated Keywords for Video Conception*

| Category | Keywords |
|---|---|
| Length | 5 minutes |
| | 10 minutes<br>15 minutes<br>> 20 minutes |
| Content | Limited |
| | Differentiated |
| Presentation | Engaging<br>Understandable<br>Powerpoint |

**Design Experience Questionnaire.** For the first item, which asked for satisfaction with the outline, it was analysed if the participant was satisfied or not. The frequency of *yes* and *no* answers were counted. The same analysis was used for the second item, which asked participants if they found the transfer from face-to-face outline to MOOC outline difficult. The third item asked for the most difficult part in transferring the outline. The frequency of each difficulty mentioned was counted. If multiple difficulties were mentioned, only the first one was counted. If participants did not fill in this question, it was scored that they did not experience any difficulties. The fourth item asked for the perceived point of conceptual change. Answers were analysed for keywords like *article*, *discussion* or *design process*. If participants stated that they did not see a difference between online and face-to-face courses or did not answer the question, it was scored as *no conceptual change*. The fifth item was scored similarly. It asked for the most helpful activity. Answers were analysed for the activities mentioned. If multiple activities were mentioned, only the first one was scored. If no activity was mentioned, it was scored as *nothing*. It was counted how frequently specific activities were mentioned.

**Change Reflection Questionnaire.** For the first item, each change that participants made was treated as a stand-alone category. If multiple changes were mentioned, they were all scored. It was counted how often specific changes were mentioned. The same procedure was conducted for item two, which asked for the inspiration to make changes.

**Quality.** The three outlines that each participant turned in were scored using the quality rubric. For each outline separately, one point per item was awarded if it existed in the outline. For example, if an outline mentioned a learning objective, the item *the outline mentions a learning objective* received one point. Points were added up for each of the five categories separately, as well as for the total score. Additionally, the conceptual change score was calculated by adding up the eight of the 21 items which were associated with conceptual change. These eight items had been selected in collaboration with an expert in online course design. The coding scheme is presented in the appendix.

**Self-efficacy.** Participants' answers were scored and analysed using the Real Statistics Resource Pack software (Release 5.1) for Microsoft Excel.

## Results

This section presents results from the various questionnaires and exercises, as well as from the analysis of the outlines that participants designed. First, results from all questionnaires relating to conceptual change are presented. Second, the quality of outlines is compared. At the end of the second section, indicators of conceptual change derived from the outlines are presented. The third section describes results for self-efficacy from the OSS-SF.

**Conceptual Change.** The MOOC Concept exercise, the MOOC Questionnaire, the Design Experience Questionnaire and the Change Reflection Questionnaire were analysed to gain information about participants' perceptions and conceptual changes. Results are presented below.

***MOOC Concept.*** The differences between MOOCs and face-to-face courses listed by participants were analysed. The questionnaire provided space for up to three differences. Keywords were extracted and assigned to one of six categories. It was counted how often specific differences were mentioned (see table 6).

Table 6

*Summary of Categories Mentioned in Article and Discussion Condition*

| | | Article condition | | Discussion condition | |
|---|---|---|---|---|---|
| Category | Keywords | Pre-test | Post-test | Pre-test | Post-test |
| Convenience | Total | 22 | 15 | 29 | 21 |
| | - Flexibility | 14 | 8 | 13 | 11 |
| | - Availability | 7 | 5 | 12 | 8 |
| | - Technological requirements | 0 | 1 | 0 | 0 |
| | - Self-directedness | 1 | 1 | 4 | 2 |
| Interaction | Total | 16 | 11 | 18 | 7 |
| | - Interaction | 13 | 9 | 15 | 6 |
| | - Lack of staff | 1 | 0 | 1 | 1 |
| | - Relationship with educator | 2 | 2 | 2 | 0 |
| Social aspects | Total | 3 | 10 | 0 | 9 |
| | - Social learning | 1 | 10 | 0 | 5 |
| | - Global community | 2 | 0 | 0 | 4 |
| Organization | Total | 12 | 13 | 3 | 8 |
| | - Structure | 5 | 9 | 2 | 6 |
| | - Pace | 6 | 4 | 1 | 2 |
| | - Scale | 1 | 0 | 0 | 0 |
| Content | Total | 1 | 5 | 1 | 5 |
| | - Exercises | 0 | 2 | 0 | 0 |
| | - Media | 0 | 1 | 0 | 1 |
| | - Content | 0 | 1 | 1 | 4 |
| | - Assessment | 1 | 1 | 0 | 0 |
| Methods | Total | 3 | 6 | 6 | 3 |
| | - Motivation | 2 | 2 | 3 | 1 |
| | - Way of learning | 1 | 4 | 3 | 2 |

Analysis of the concept that participants held at the time of the pre-test revealed a focus on surface characteristics of MOOCs. For example, the flexibility that a MOOC offers was mentioned by 27 participants from both groups. The difference in interaction was noted by 28 participants. However, differences in content and social aspects were rarely mentioned. In the post-test, participants focused on social aspects and content more than on surface characteristics like convenience. In the post-test, social aspects were mentioned by 19 participants, while differences in content were elaborated by 10 participants. One participant listed *"storytelling, real life examples and context"* as well as *"big questions"* as important features of MOOCs. Another participant mentioned how questions are dealt with as a main difference. In the pre-test, that participant stated that *"questions in MOOCs cannot be answered"* while they *"can be answered face-to-face"*. However, in the post-test the same participant noted that *"questions in a MOOC can be asked to a very broad community"* while *"questions in a traditional course can be asked to a smaller group"*.

***MOOC Questionnaire.*** Reported answers for each of the three items were analysed and coded. Seven different main purposes of MOOCs were mentioned for the first question about the main purpose of a MOOC (see table 7). If a participant mentioned multiple purposes, all of them were scored.

Table 7

*Summary of Perceived Main Purpose of a MOOC*

| | Article condition | | Discussion condition | |
|---|---|---|---|---|
| Main purpose | Pre-test | Post-test | Pre-test | Post-test |
| Knowledge | 12 | 12 | 7 | 12 |
| Availability | 7 | 6 | 10 | 7 |
| Flexibility | 5 | 4 | 4 | 2 |
| Self-directedness | 4 | 7 | 6 | 3 |
| Exercises | 2 | 4 | 0 | 0 |
| Inspiration | 0 | 0 | 0 | 3 |
| Social learning | 2 | 4 | 1 | 9 |

To compare the number of main purposes mentioned in both conditions, paired sample *t*-tests were conducted for each condition separately. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 8 and 9. In the article condition, there was no significant difference in the number of categories mentioned between the pre-test and the post-test; $t(21) = 2.02$, $p = .057$. In the discussion condition, the number of categories mentioned in the post-test was significantly higher than the number of categories mentioned in the pre-test; $t(19) = 2.63$, $p < .05$. Thus, the number of categories participants mentioned increased significantly after the discussion intervention, but not after the article intervention.

Table 8

*Descriptive Statistics Associated With the Number of Categories Mentioned in the Article Condition*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Pre-test | 22 | 1.45 | 0.74 | 1.98 | 5.73 |
| Post-test | 22 | 1.68 | 0.78 | 1.31 | 2.37 |

Table 9

*Descriptive Statistics Associated With the Number of Categories Mentioned in the Discussion Condition*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Pre-test | 20 | 1.4 | 0.50 | 0.44 | -2.02 |
| Post-test | 20 | 1.8 | 0.41 | -1.62 | 0.70 |

Some participants changed their answers from the pre-test to the post-test, while others perceived the main purpose as unchanged by the intervention. One participant in the article group stated in the pre-test: *"The main purpose of a MOOC is to educate at a larger scale."* and in the post-test: *"The main purpose of a MOOC is to make learning easily accessible and attainable."*. A characteristic response in the discussion group was given by one participant in the pre-test: *"(The main purpose is to) learn wherever* *and whenever you are by interacting online with videos, tests and peers and teachers.".* This answer changed after the intervention to: *"(The main purpose is to) get to think about a big question by storytelling lectures that inspire and make you want to discuss it with other peers."* However, an independent samples *t*-test did not show a significant difference in the number of participants who changed their answer after the intervention between the article and the discussion condition; $t(40) = 2.02$, $p = .132$. Differences in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 10. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 41)= 2.37$, $p = .644$.

Table 10

*Descriptive Statistics Associated With the Change of Main Purpose from Pre-test to Post-test*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 0.36 | 0.49 | 0.61 | -1.80 |
| Discussion condition | 20 | 0.60 | 0.50 | -0.44 | -2.02 |

A repeated measures ANOVA was conducted to compare the effect of the type of intervention on perceiving social learning as the main purpose of a MOOC before and after the intervention. There was a significant effect for time, $F\ (1, 38) = 8.96$, $p < .005$. There was no significant effect for type of intervention, $F\ (1, 38) = 1.36$, $p = .251$. The interaction was not statistically significant, $F\ (1, 38) = 3.23$, $p = .08$ (see figure 2).

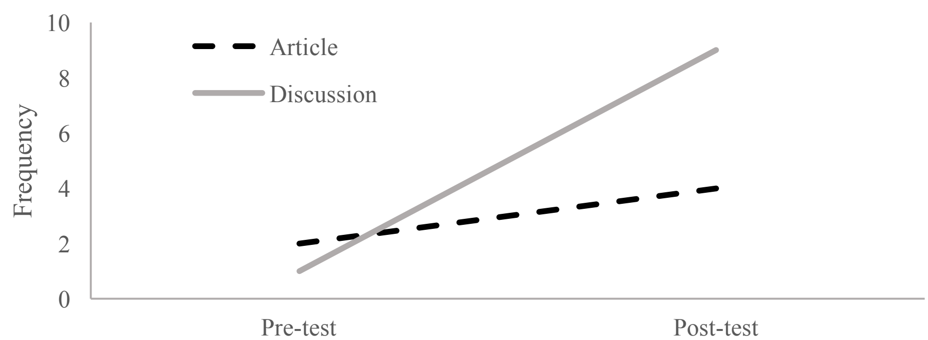


*Figure 2.* Frequency of how often social learning was mentioned as the main purpose of a MOOC before and after the intervention.

For the second question about the structure of a MOOC, 13 aspects were mentioned. Seven aspects have been assigned to the category "content", two aspects to the category "extent" and four aspects to the category "progression" (see table 11). If a participant mentioned multiple keywords, all of them were scored. For example, only four participants mentioned discussion as part of the content in the pre-test over both conditions. However, in the post-test, 15 participants mentioned discussion in both conditions. A repeated measures ANOVA revealed a significant effect for time, $F\ (1, 38) = 8.36$, $p < .05$, as well as for type of intervention, $F\ (1, 38) = 4.33$, $p < .05$. The interaction was not statistically significant, $F\ (1, 38) = 0.62$, $p = .435$ (see figure 3).

Table 11

*Summary of Structural Aspects of a MOOC*

| | | Article condition | | Discussion condition | |
|---|---|---|---|---|---|
| Category | Keywords | Pre-test | Post-test | Pre-test | Post-test |
| Content | Information | 7 | 6 | 6 | 8 |
| | Exercises | 7 | 3 | 1 | 0 |
| | Reflection | 2 | 3 | 0 | 0 |
| | Discussion | 3 | 10 | 1 | 5 |
| | Evaluation | 3 | 2 | 5 | 2 |
| | Variation | 0 | 5 | 0 | 1 |
| | Narrative | 0 | 3 | 0 | 0 |
| Extent | Strict | 2 | 1 | 0 | 0 |
| | Concise | 0 | 2 | 0 | 0 |
| Progression | Visible progress | 2 | 1 | 1 | 0 |
| | Linear progress | 6 | 3 | 10 | 8 |
| | Flexible progress | 0 | 1 | 0 | 8 |
| | Easy to difficult | 4 | 1 | 2 | 1 |
| | Broad to specific | 1 | 3 | 2 | 1 |

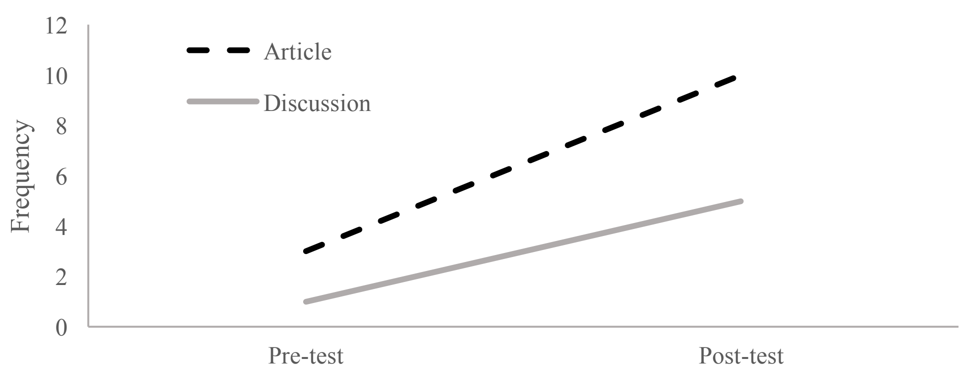


*Figure 3.* Frequency of how often discussion was mentioned as part of the structure of a MOOC before and after the intervention.

The number of participants advocating a linear progress decreased from 16 to 11 over both conditions. A repeated measure ANOVA revealed no significant effect for time, $F$ (1, 38) = 1.45, $p$ = .236, but a significant effect for type of intervention, $F$ (1, 38) = 4.59, $p$ < .05. The interaction was not statistically significant, $F$ (1, 38) = 0.06, $p$ = .811 (see figure 4).

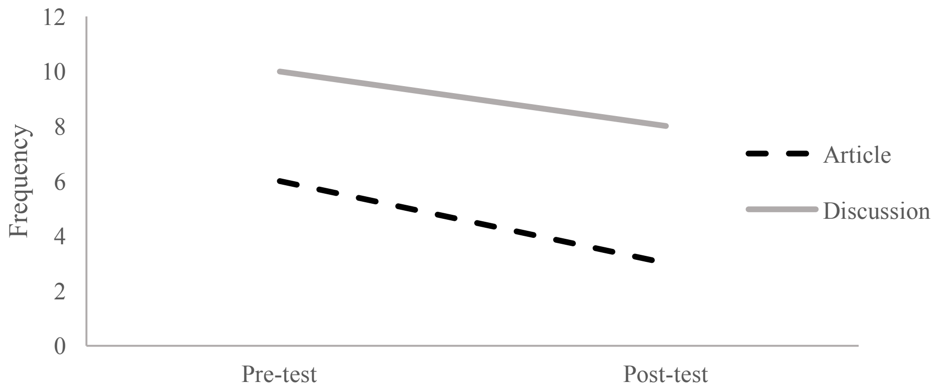


*Figure 4.* Frequency of how often a linear progress was mentioned as part of the structure of a MOOC before and after the intervention.

However, the number of participants mentioning a flexible progress increased from zero in the pre-test to nine participants in the post-test. A repeated measure ANOVA revealed a significant effect for time, $F\,(1{,}38) = 13.38$, $p < .001$, as well as a significant effect for type of intervention, $F\,(1, 38) = 8.09$, $p < .001$. The interaction was statistically significant, $F\,(1, 38) = 8.09$, $p < .001$ (see figure 5).

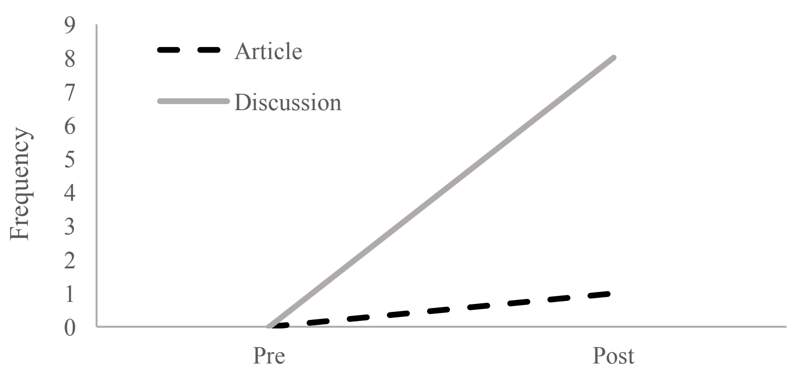


*Figure 5.* Frequency of how often a flexible progress was mentioned as part of the structure of a MOOC before and after the intervention.

The third item in the MOOC questionnaire asked what a good video in a MOOC looks like. Participants' answers were analysed and assigned to three categories, namely *length*, *content* and *presentation*. If a time range was given, i.e. 5-10 minutes, the lower length was counted for the category *length*. Two keywords were assigned to the category *content* and three keywords were assigned to the category *presentation* (see table 12). If a participant mentioned multiple keywords, all of them were scored.

Table 12

*Summary of Video Conception*

| | | Article condition | | Discussion condition | |
|---|---|---|---|---|---|
| Category | Keywords | Pre-test | Post-test | Pre-test | Post-test |
| Length | 5 minutes | 4 | 4 | 2 | 7 |
| | 10 minutes | 11 | 6 | 6 | 7 |
| | 15 minutes | 3 | 4 | 3 | 3 |
| | >20 minutes | 0 | 0 | 3 | 1 |
| Content | Limited | 5 | 6 | 3 | 2 |
| | Differentiated | 1 | 2 | 0 | 1 |
| Presentation | Engaging | 5 | 9 | 3 | 8 |
| | Understandable | 6 | 6 | 4 | 2 |
| | Powerpoint | 1 | 0 | 3 | 0 |

For the category *Length* especially, participants changed their answers after the intervention. For example, one participant in the discussion group stated before the intervention: *"The video should be no longer than 15-20 minutes, because after that time you lost interest while looking at the computer."* This answer changed to: *"(The video) should be between 5 and 15 minutes long, because if it is longer, you can't concentrate anymore."* A Mann-Whitney test indicated that the change in video length was significantly greater in the discussion condition compared to the article condition, $U = 49.5$, $p < .05$ (Mann & Whitney, 1947). Differences in both condition were not sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010), so the Mann-Whitney test was used. See table 13. Only answers which included a value in minutes in both pre-test and post-test were included in this analysis.

Table 13

*Descriptive Statistics Associated With the Change of Video Length from Pre-test to Post-test*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 14 | 1.43 | 4.57 | -2.16 | 6.26 |
| Discussion condition | 14 | 5 | 8.55 | 2.26 | 5.47 |

With regard to content of videos, there were only minor differences between pre-test and post-test. However, regarding the presentation of the video, there was an increase in the number of participants from both conditions who stated that the video should be engaging. A repeated measure ANOVA revealed a significant effect for time, $F\ (1, 38) = 5.94$, $p < .05$, but not for type of intervention, $F\ (1, 38) = 0.45$, $p = .508$. The interaction was not statistically significant, $F\ (1, 38) = 0.07$, $p = .788$ (see figure 6). Additionally, the number of participants who mentioned a powerpoint presentation as part of the video decreased from four to zero over both conditions.

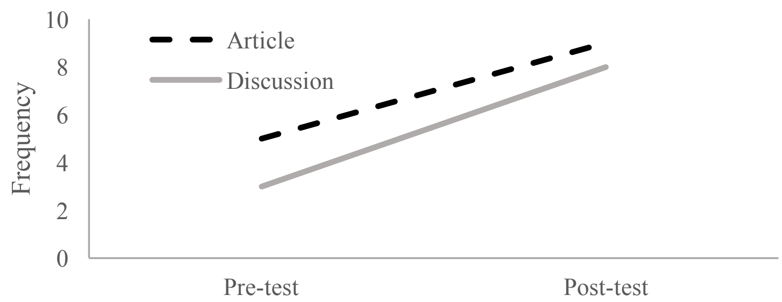


*Figure 6.* Frequency of how a good video was expected to be engaging before and after the intervention.

***Design Experience Questionnaire.*** Answers to this questionnaire were analysed and coded. The first item asked for the participants' satisfaction with their outlines. For example, one participant stated: "*I think in the end I created an outline for a MOOC which is acceptable, but not very good. Since beforehand, I had a different expectation about what the goal of a MOOC is, as I had no experience." (Article group).* Ten participants form the article condition stated they were satisfied with the outline they produced, while twelve were not satisfied. In the discussion condition, nine participants were satisfied, while eleven were not satisfied with their outlines (see table 14). A chi-square test of maximum likelihood was performed to examine the relation between condition and satisfaction. The relation between these variables was not significant; $\chi^2$ (1, $N$ = 42) = 0.0009, $p$ = .976. Thus, the level of satisfaction is not associated with the type of intervention.

Table 14

*Satisfaction With MOOC Outline*

| Satisfaction | Article condition | Discussion condition |
|---|---|---|
| Yes | 10 | 9 |
| No | 12 | 11 |

The second item assessed whether participants found it difficult to transfer their face-to-face outline into a MOOC outline. For example, one participant said: *"After the discussion we had and after I knew more about MOOCs, I realised that it is really difficult to transfer it, since there was actually a practical part included." (Discussion group).* In the article condition, seven participants found this difficult, while 15 did not report difficulties. In the discussion condition, eight participants experienced difficulties, while twelve did not report any difficulties (see table 15). A chi-square test of maximum likelihood was performed to examine the relation between condition and perceived difficulty. The relation between these variables was not significant; $\chi^2$ (1, $N$ = 42) = 0.31, $p$ = .580. Thus, the perceived difficulty of transferring the face-to-face outline into a MOOC outline is not associated with the type of intervention.

Table 15

*Perceived Difficulty of Transfer*

| Difficulty | Article condition | Discussion condition |
|---|---|---|
| Yes | 7 | 8 |
| No | 15 | 12 |

The third item asked participants to name the difficulties they experienced in transferring the face-to-face outline into the MOOC outline. One participant from the article group stated: *"The most difficult part in transferring the face-to-face to an online was making sure there was space for communication between the lecturer and participant."*. Another participant from the discussion group found *"picking the information and how you're going to present it as short and interesting as possible"* the most difficult part. A total of five different difficulties were mentioned, while six participants did not experience any difficulties. Adapting the exercises and structure were mentioned most often in the article condition, while adapting the exercises and the content of the course was the most common difficulty in the discussion condition (see table 16). A chi-square test of maximum likelihood was performed to examine the relation between condition and experienced difficulties. The relation between these variables was not significant; $\chi^2$ (5, $N$ = 42) = 5.55, $p$ = .353. Thus, the experienced difficulties in transferring the face-to-face outline into a MOOC outline are not associated with the type of intervention.

Table 16

*Experienced Difficulties*

| Difficulties | Article condition | Discussion condition |
|---|---|---|
| Exercises | 7 | 7 |
| Structure | 4 | 1 |
| Discussion | 2 | 1 |
| Time | 2 | 1 |
| Content | 3 | 8 |
| Nothing | 4 | 2 |

The forth item investigated if and at what point participants felt like they realized the fundamental difference between face-to-face courses and MOOCs. Answers were analysed and coded. For example, one participant in the article group stated: *"I had a small opinion about the difference before, but I really saw the fundamental difference after I read the article"*, while a participant from the discussion group said: *"I fully realised and understood the differences at the point of our discussion with the group."*. In the article condition, eight participants did not see a fundamental difference between face-to-face courses and MOOCs. Of the remaining 14 participants, seven realized a fundamental difference while reading the article (intervention), whereas five realized a difference while listening to the presentation. Two participants realised the difference during the design of their own MOOCs. In the discussion condition, all but two participants realized a fundamental difference. Of the remaining 18 participants, 15 realised the difference during the group discussion (intervention), while two realised it during the presentation and another participant during designing his own MOOC (see table 17).

Table 17

*Perceived Point of Conceptual Change*

| Activity | Article condition | Discussion condition |
|---|---|---|
| Intervention | 7 | 15 |
| Presentation | 5 | 2 |
| Design activity | 2 | 1 |
| No conceptual change | 8 | 2 |

A Mann-Whitney test was conducted to compare the two conditions regarding their contribution to a conceptual change (Mann & Whitney, 1947). For this analysis, answers were either scored as 1 point *(participant experienced conceptual change*), or 0 points (*participant did not experience conceptual change*). Differences in both condition were not sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 18. The number of participants to experience a conceptual change was significantly greater in the discussion condition than in the article condition; $U = 162$, $p < .05$. Thus, the discussion condition was more effective in producing a conceptual change in participants.

Table 18

*Descriptive Statistics Associated With Conceptual Change vs. no Conceptual Change*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 0.64 | 0.49 | -0.61 | -1.80 |
| Discussion condition | 20 | 0.9 | 0.31 | -2.89 | 7.03 |

In order to analyse the effect of the intervention on conceptual change, an independent samples *t*-test was conducted, comparing the experience of conceptual change during and outside the intervention over both conditions. Only participants who had experienced a conceptual change were included in this analysis. See table 17: *Intervention* is scored as conceptual change during the intervention, while *Presentation* and *Design Activity* are scored as conceptual change outside the intervention. If participants had experienced conceptual change during the intervention, they scored 1 point, if they had experienced it outside the intervention, they scored 0 points. Scores in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 19. The criteria of homogeneity of equal variances was not satisfied using Levene's *F*-test, $F(1, 30)= 4.38$, $p < .005$, therefore, unequal variances were assumed. Scores in the discussion group were significantly larger than scores in the article group; $t(23) = 2.01$, $p < .05$. Cohen's *d* is estimated at 0.75, which resembles a medium effect (Cohen, 1992). Thus, the discussion intervention was significantly more effective in producing a conceptual change as part of the intervention than the article condition.

Table 19

*Descriptive Statistics Associated With Conceptual Change During vs. Outside Intervention*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 14 | 0.5 | 0.52 | <0.01 | -2.36 |
| Discussion condition | 18 | 0.83 | 0.38 | -1.96 | 2.04 |

The last item in the questionnaire asked participants to name the most helpful aspects of the workshop which made them realize the fundamental difference between face-to-face and online courses. If multiple aspects were mentioned, the first one was counted. One participant from the article condition mentioned: *"Reading the article and summarizing it for me was the best way to keep information in mind and think about it more effectively."*, while one participant in the discussion group stated: *"The discussion was particularly helpful for me, since I only understood MOOCs completely because of that. I had a good idea before, but important aspects were missing."*. In the article condition, the intervention was mentioned by nine participants. Additionally, the activities of designing an outline and revising an outline were mentioned five and four times respectively. Four participants did not find any activity particularly helpful. In the discussion condition, the discussion intervention was mentioned by 16 participants. Designing an outline was mentioned by two participants. Another two participants did not mention any activity as particularly helpful (see table 20). A chi-square test of maximum likelihood was performed to examine the relation between condition and most helpful aspect. The relation between these variables was significant; $\chi^2$ (3, $N$ = 42) = 7.83, $p$ = < .005. Cramer's $V$ is estimated at 0.43, which resembles a large effect (Cohen, 1988). Thus, the aspect that participants found most helpful is associated with the type of intervention.

Table 20

*Most Helpful Aspect*

| Activity | Article condition | Discussion condition |
|---|---|---|
| Intervention | 9 | 16 |
| Design outline | 5 | 2 |
| Revise outline | 4 | 0 |
| Nothing | 4 | 2 |

***Change Reflection Questionnaire.*** The first item in this questionnaire asked participants to describe the changes they have made from their first MOOC outline to the final version. Answers were analysed and coded. If multiple changes were mentioned, each of them was counted. In the article condition changes to the content, structure and to include social learning were mentioned most frequently. Changes to the addition of a narrative was also mentioned often. For example, one participant in the article condition said: *"I changed mostly the way of how I want to provide the information. First it was the typical teacher explains in front of the class way, and after it I thought about a more creative way to teach lessons."* In the discussion condition changes to the structure, content, amount of social learning and the purpose of the MOOC were most common (see table 21). A participant in the discussion condition stated: *"I now think that MOOCs can also be interactive but on a different level than traditional courses, since MOOCs provide space for interaction between different cultures and parts of the world whereas traditional courses provide space for interaction between students and teacher."*

Table 21

*Changes to MOOC Outline*

| Changes | Article condition | Discussion condition |
|---|---|---|
| Structure | 8 | 7 |
| Social learning | 8 | 7 |
| Content | 10 | 5 |
| Narrative | 4 | 1 |
| Big Question | 2 | 2 |
| Learning goal | 1 | 0 |
| Teaching method | 0 | 1 |
| Purpose | 0 | 3 |
| No change | 0 | 1 |

The second item in this questionnaire asked what inspired participants to make the changes they had mentioned. If multiple sources of inspiration were mentioned, all of them were included. In the article condition, almost all participants mentioned the article intervention. One participant explained: *"In the text I read that for example conversation is an important point so I added activities and possibilities in the MOOC which are important for a successful MOOC."* In the discussion condition, all but one participants mentioned the discussion intervention. One participant reflected: *"The group discussion and different aspects and insights that came to the surface during the discussion inspired me. I heard what others considered to be important and tried to implement it in my outline. Further that everybody had a different opinion of what a MOOC should look like helped me identify aspects I did not think about before. It was also interesting to see a short example of what a MOOC looks like to develop a structure of the course."* Additionally, the presentation, prior MOOC experience and the adaption of the face-to-face to the MOOC outline were mentioned (see table 22).

Table 22

*Inspiration for Changes*

| Inspiration | Article condition | Discussion condition |
|---|---|---|
| Intervention | 18 | 19 |
| Presentation | 1 | 2 |
| MOOC experience | 1 | 0 |
| Adapt F2F outline to online outline | 0 | 1 |

**Quality.** In order to assess the quality of the outlines, they were scored with the help of a rating rubric. Scores for the categories are presented in the following order: Introduction, content, interaction, transfer of knowledge, structure and indicators of conceptual change.

***Rating rubric.*** The three outlines (face-to-face, MOOC first draft, MOOC final version) that participants produced were scored using a rating rubric. Points in each of the five categories (introduction, content, interaction, transfer of knowledge and structure) were added separately, as well as summed up for a total score. An overview of the possible points, as well as highest and lowest scores by individual participants is given in table 23.

Table 23

*Possible Scores, Lowest Scores and Highest Scores for Each Category Over Both Groups*

| Category | Score possible | Face-to-face | | MOOC first draft | | MOOC final draft | |
|---|---|---|---|---|---|---|---|
| | | Low | High | Low | High | Low | High |
| Introduction | 4 | 0 | 2.5 | 0 | 2.5 | 0 | 3 |
| Content | 6 | 0 | 5 | 0.5 | 5 | 0.5 | 6 |
| Interaction | 3 | 0 | 2 | 0 | 2 | 0 | 2.5 |
| Transfer of knowledge | 4 | 0 | 2 | 0 | 2.5 | 0 | 3 |
| Structure | 4 | 0 | 2 | 0.5 | 2.5 | 0.5 | 3 |
| Total | 21 | 1 | 8 | 2.5 | 11 | 3.5 | 14 |
| Indicators of conceptual change | 8 | 0 | 4.5 | 0 | 3.5 | 0 | 5.5 |

Average scores for participants in each condition were calculated for each outline separately. See table 24 for an overview of average scores. In the following paragraphs scores for each category are presented. Specifically, using a paired-sample *t*-test, scores for the first MOOC outline are compared with scores for the final MOOC outline for each condition separately, as these outlines have been created before (MOOC first) and after (MOOC final) the intervention. Additionally, scores for the MOOC final outlines are compared between both conditions using an independent sample *t*-test. Further, the difference in scores between first and final MOOC outline is compared between conditions to analyse the effect of the intervention on the difference in scores.

Table 24

*Means and Standard Deviations for Outline Scores in Both Conditions*

| | Article condition | | | | | | Discussion condition | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Face-to-face | | MOOC first | | MOOC final | | Face-to-face | | MOOC first | | MOOC final | |
| Category | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Introduction | 1.07 | 0.64 | 1.09 | 0.73 | 1.68 | 0.84 | 0.98 | 0.62 | 1.15 | 0.69 | 1.38 | 0.65 |
| Content | 1.09 | 0.93 | 2.20 | 1.08 | 3.66 | 1.36 | 1.58 | 1.25 | 2.78 | 1.20 | 3.85 | 1.24 |
| Interaction | 0.27 | 0.57 | 0.25 | 0.51 | 1.09 | 0.75 | 0.25 | 0.38 | 0.18 | 0.34 | 0.73 | 0.62 |
| Transfer of knowledge | 0.77 | 0.55 | 1.09 | 0.67 | 1.57 | 0.68 | 1.13 | 0.63 | 1.30 | 0.52 | 1.45 | 0.54 |
| Structure | 0.86 | 0.64 | 1.45 | 0.60 | 1.95 | 0.77 | 1.28 | 0.85 | 1.65 | 0.61 | 1.90 | 0.60 |
| Total | 4.07 | 2.21 | 6.09 | 2.14 | 10.00 | 2.87 | 5.15 | 2.46 | 7.03 | 1.98 | 9.25 | 2.48 |
| Indicators of conceptual change | 0.91 | 0.84 | 1.20 | 0.88 | 3.41 | 1.49 | 1.45 | 1.11 | 1.55 | 1.01 | 3 | 1.44 |

*Introduction.* A maximum of four points was possible in this category. An overview of scores for both conditions on each outline is presented in table 24. To compare the first MOOC outline with the final MOOC outline, paired sample t-tests were conducted for each condition separately. In the article condition, the scores of the first MOOC outline were significantly lower than the scores of the final outline; $t(21) = -4.05$, $p < .001$. Cohen's $d$ is estimated at 0.86, which resembles a large effect (Cohen, 1992). In the discussion condition, scores of the first MOOC outline and the final outline differed significantly; $t(19) = 2.09$, $p < .05$. Cohen's $d$ is estimated at 0.59, which resembles a medium effect (Cohen, 1992). Thus, scores of the introduction significantly improved from the first to the final MOOC outline in both the article condition and the discussion condition.

To compare scores of the final outline between conditions, an independent sample *t*-test was conducted. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). See table 25. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40)= 1.74$, $p = .231$. Scores for the final outline in the article group did not differ significantly from scores for the final outline in the discussion group; $t(40) = 1.32$, $p = .195$. Thus, introductions of the article group were not of significantly higher quality than introductions of the discussion group.

Table 25

*Descriptive Statistics Associated With Introduction Scores for Final MOOC Outlines*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 1.68 | 0.84 | -0.37 | -0.28 |
| Discussion condition | 20 | 1.38 | 0.65 | 0.03 | -1.17 |

Differences in scores of the first and final MOOC outlines between both condition were compared. For this, an independent samples *t*-test was conducted. Differences in both conditions were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 26. The criteria of homogeneity of equal variances was not satisfied using Levene's *F*-test, $F(1, 40)= 4.47$, $p < .001$, therefore, unequal variances were assumed. Differences between outlines in the article group were significantly larger than differences between outlines in the discussion group; $t(33) = 2.17$, $p < .05$. Cohen's $d$ is estimated at 0.65, which resembles a medium effect (Cohen, 1992). Thus, the quality of introductions in the article group improved significantly more than the quality of introductions in the discussion group.

Table 26

*Descriptive Statistics Associated With Differences in Introduction Scores*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 0.59 | 0.68 | .62 | -1.12 |
| Discussion condition | 20 | 0.23 | 0.38 | 1.39 | 0.41 |

***Content.*** A maximum of six points was possible in this category. An overview of scores for both conditions on each outline is presented in table 24. To compare the first MOOC outline with the final MOOC outline, paired sample *t*-tests were conducted for each condition separately. In the article condition, scores of the first MOOC outline were significantly lower than scores of the final outline; $t(21) = -6.60$, $p < .001$. Cohen's $d$ is estimated at 1.41, which resembles a large effect (Cohen, 1992). In the discussion condition, scores of the first MOOC outline were significantly lower than scores for the final outline; $t(19)$

= -4.36, *p* < .001. Cohen's *d* is estimated at 0.97, which resembles a large effect (Cohen, 1992). Thus, the score for content significantly improved from the first to the final MOOC outline in both the article condition and in the discussion condition.

To compare scores between conditions, an independent sample *t*-test was conducted. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 27. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, *F*(1, 40)= 0.23, *p* = .381. Scores for the final outline in the article group did not differ significantly from scores for the final outline in the discussion group; *t*(40) = .47, *p* = .638. Thus, content of the article group was not of significantly higher quality than content of the discussion group.

Table 27

*Descriptive Statistics Associated With Content Scores for Final MOOC Outlines*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 3.66 | 1.36 | -0.42 | -0.58 |
| Discussion condition | 20 | 3.85 | 1.24 | -1.12 | 2.17 |

Differences in scores of the first and final MOOC outlines between both condition were compared. For this, an independent samples *t*-test was conducted. Differences in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 28. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, *F*(1, 40)= 1.32, *p* = .930. Differences between outlines in the article group were not significantly larger than differences between outlines in the discussion group; *t*(40) = 1.15, *p* = .257. Thus, the quality of content in the article group did not improve significantly more than the quality of content in the discussion group.

Table 28

*Descriptive Statistics Associated With Differences in Content Scores*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 1.45 | 1.03 | 0.35 | -0.78 |
| Discussion condition | 20 | 1.08 | 1.10 | 0.02 | -0.34 |

*Interaction.* A maximum of three points was possible in this category. An overview of scores for both conditions on each outline is presented in table 24. To compare the first MOOC outline with the final MOOC outline, paired sample *t*-tests were conducted for each condition separately. In the article condition, scores of the first MOOC outline were significantly lower than scores of the final outline; *t*(21) = -5.07, *p* < .001. Cohen's *d* is estimated at 1.08, which resembles a large effect (Cohen, 1992). In the discussion condition, scores of the first MOOC outline were significantly lower than scores for the final outline; *t*(20) = -4.40, *p* < .001. Cohen's *d* is estimated at 0.98, which resembles a large effect (Cohen, 1992). Thus, the score for interaction improved significantly from the first to the final MOOC outline in both the article condition and in the discussion condition.

To compare scores between conditions, an independent sample *t*-test was conducted. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). See table 29. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, *F*(1, 40)= 2.94, *p* = .346. Scores for the final outline in the article group did not differ significantly from scores for the final outline in the discussion group; $t(40) = 1.72$, $p = .09$. Thus, the interactive parts of outlines in the article group were not of significantly higher quality than the interactive parts of outlines in the discussion group.

Table 29

*Descriptive Statistics Associated With Interaction Scores for Final MOOC Outlines*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 1.09 | 0.75 | 0.31 | -0.63 |
| Discussion condition | 20 | 0.73 | 0.62 | 0.87 | 0.02 |

Differences in scores of the first and final MOOC outlines between both condition were compared. For this, an independent samples *t*-test was conducted. Differences in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 30. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40)= 1.90$, $p = .06$. Differences between outlines in the article group were not significantly larger than differences between outlines in the discussion group; $t(40) = 1.38$, $p = .175$. Thus, the quality of interaction in the article group did not improve significantly more than the quality of interaction in the discussion group.

Table 30

*Descriptive Statistics Associated With Differences in Interaction Scores*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 0.84 | 0.78 | 0.42 | -0.48 |
| Discussion condition | 20 | 0.55 | 0.56 | 1.04 | 0.91 |

***Transfer of knowledge.*** A maximum of four points was possible in this category. An overview of scores for both conditions on each outline is presented in table 24. To compare the first MOOC outline with the final MOOC outline, paired sample *t*-tests were conducted for each condition separately. In the article condition, scores of the first MOOC outline were significantly lower than scores of the final outline; $t(21) = -3.95$, $p < .001$. Cohen's *d* is estimated at .84, which resembles a large effect (Cohen, 1992). In the discussion condition, scores of the first MOOC outline were significantly lower than scores for the final outline; $t(19) = -2.85$, $p < .05$. Cohen's *d* is estimated at .64, which resembles a medium effect (Cohen, 1992). Thus, the score for transfer of knowledge improved significantly from the first to the final MOOC outline in both the article condition and the discussion condition.

To compare scores between conditions, an independent sample *t*-test was conducted. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 31. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40)= 0.39$, $p = .535$. Scores for the final outline in the article group did not differ significantly from scores for the final outline in the discussion group; $t(40) = 0.62$, $p = .537$. Thus, the transfer of knowledge in the article group was not of significantly higher quality than the transfer of knowledge in the discussion group.

Table 31

*Descriptive Statistics Associated With Transfer of Knowledge Scores for Final MOOC Outlines*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 1.57 | 0.68 | -0.14 | 0.62 |
| Discussion condition | 20 | 1.45 | 0.54 | -0.64 | -0.72 |

Differences in scores of the first and final MOOC outlines between both condition were compared. For this, an independent samples *t*-test was conducted. Differences in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 32. The criteria of homogeneity of equal variances was not satisfied using Levene's *F*-test, $F(1, 40)= 0.76$, $p < .001$. Unequal variances are therefore assumed. Differences between outlines in the article group were significantly larger than differences between outlines in the discussion group; $t(29) = 2.48$, $p < .05$. Cohen's *d* is estimated at .74, which resembles a medium effect (Cohen, 1992). Thus, the quality of transfer of knowledge in the article group improved significantly more than the quality of transfer of knowledge in the discussion group.

Table 32

*Descriptive Statistics Associated With Differences in Transfer of Knowledge Scores*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 0.48 | 0.57 | 0.74 | -0.93 |
| Discussion condition | 20 | 0.15 | 0.24 | 0.95 | -1.24 |

*Structure.* A maximum of four points was possible in this category. An overview of scores for both conditions on each outline is presented in table 24. To compare the first MOOC outline with the final MOOC outline, paired sample *t*-tests were conducted for each condition separately. In the article condition, scores of the first MOOC outline were significantly lower than scores of the final outline; $t(21) = -3.69$, $p < .005$. Cohen's *d* is estimated at .79, which resembles a medium effect (Cohen, 1992). In the discussion condition, scores of the first MOOC outline were also significantly lower than scores for the final outline; $t(21) = -2.24$, $p < .05$. Cohen's *d* is estimated at .5, which resembles a medium effect (Cohen, 1992). Thus, the score for structure improved significantly from the first to the final MOOC outline in the article condition, as well as in the discussion condition.

To compare scores between conditions, an independent sample *t*-test was conducted. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 33. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40)= 0.06$, $p = .321$. Scores for the final outline in the article group did not differ significantly from scores for the final outline in the discussion group; $t(40) = 0.25$, $p = .8$. Thus, the structure of outlines in the article group was not of significantly higher quality than the structure of outlines in the discussion group.

Table 33

*Descriptive Statistics Associated With Structure Scores for Final MOOC Outlines*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 1.95 | 0.77 | -0.09 | -0.13 |
| Discussion condition | 20 | 1.90 | 0.60 | 0.22 | -0.49 |

Differences in scores of the first and final MOOC outlines between both condition were compared. For this, an independent samples *t*-test was conducted. Differences in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 34. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40)= 1.98$, $p = .256$. Differences between outlines in the article group were not significantly larger than differences between outlines in the discussion group; $t(40) = 1.41$, $p = .167$. Thus, the quality of structure in the article group did not improve significantly more than the quality of structure in the discussion group.

Table 34

*Descriptive Statistics Associated With Differences in Structure Scores*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 0.5 | 0.64 | 1.07 | -0.05 |
| Discussion condition | 20 | 0.25 | 0.5 | 0.18 | 2.55 |

*Overall score.* A maximum of 21 points was possible across all categories. An overview of scores for both conditions on each outline is presented in table 24. To compare the first MOOC outline with the final MOOC outline, paired sample *t*-tests were conducted for each condition separately. In the article condition, scores of the first MOOC outline were significantly lower than scores of the final outline; $t(21) = -8.74$, $p < .001$. Cohen's *d* is estimated at 1.86, which resembles a large effect (Cohen, 1992). In the discussion condition, scores of the first MOOC outline were significantly lower than scores for the final outline; $t(19) = -6.06$, $p < .001$. Cohen's *d* is estimated at 1.35, which resembles a large effect (Cohen, 1992). Thus, the overall score improved significantly from the first to the final MOOC outline in both the article condition and in the discussion condition.

To compare scores between conditions, an independent sample *t*-test was conducted. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 35. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40)= 0.81$, $p = .359$. Scores for the final outline in the article group did not differ significantly from scores for the final outline in the discussion group; $t(40) = 0.9$, $p = .373$. Thus, the overall quality of outlines in the article group was not significantly higher than the overall quality of outlines in the discussion group.

Table 35

*Descriptive Statistics Associated With Total Scores for Final MOOC Outlines*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 10 | 2.87 | -0.65 | -0.63 |
| Discussion condition | 20 | 9.25 | 2.48 | -0.82 | 0.4 |

WHAT IS THE EFFECT OF A GROUP DISCUSSION COMPARED TO INDIVIDUAL READING ON DESIGNING AN OUTLINE FOR A MOOC?

Differences in scores of the first and final MOOC outlines between both condition were compared. For this, an independent samples *t*-test was conducted. Differences in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 36. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40) = 8.28$, $p = .121$. Differences between outlines in the article group were significantly larger than differences between outlines in the discussion group; $t(40) = 2.88$, $p < .01$. Cohen's *d* is estimated at 0.89, which resembles a large effect (Cohen, 1992). Thus, the overall quality in the article group improved significantly more than the overall quality in the discussion group.

Table 36

*Descriptive Statistics Associated With the Difference in Total Scores*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 3.91 | 2.1 | 0.02 | -0.90 |
| Discussion condition | 20 | 2.23 | 1.64 | -0.45 | 0.27 |

*Conceptual change score.* In addition to the analysis above, a score for conceptual change was computed. For this, eight items across all categories which closely reflect conceptual change were selected. An overview of scores for both conditions on each outline is presented in table 24. To compare the first MOOC outline with the final MOOC outline, paired sample *t*-tests were conducted for each condition separately. In the article condition, scores of the first MOOC outline were significantly lower than scores of the final outline; $t(21) = -7.42$, $p < .001$. Cohen's *d* is estimated at 1.58, which resembles a large effect (Cohen, 1992). In the discussion condition, scores of the first MOOC outline were significantly lower than scores for the final outline; $t(19) = -6.25$, $p < .001$. Cohen's *d* is estimated at 1.40, which resembles a large effect (Cohen, 1992). Thus, the indicators of conceptual change increased significantly from the first to the final MOOC outline in both the article condition and in the discussion condition.

To compare scores between conditions, an independent sample *t*-test was conducted. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 37. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40) = 0.81$, $p = .979$. Scores for the final outline in the article group did not differ significantly from scores for the final outline in the discussion group; $t(40) = 0.9$, $p = .373$. Thus, there were not significantly more indicators of conceptual change in the article group than in the discussion group.

Table 37

*Descriptive Statistics Associated With Conceptual Change Scores for Final MOOC Outlines*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 3.41 | 1.49 | -0.38 | -0.9 |
| Discussion condition | 20 | 3 | 1.44 | -0.57 | -0.55 |

Differences in scores of the first and final MOOC outlines between both condition were compared. For this, an independent samples *t*-test was conducted. Differences in both condition were sufficiently normal distributed for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 38. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 40) = .89$, $p = .083$. Differences between outlines in the article group were not significantly larger than differences between outlines in the discussion group; $t(40) = 1.97$, $p = .055$. Thus, the conceptual change in the article group was not significantly larger than in the discussion group.

Table 38

*Descriptive Statistics Associated With the Difference in Conceptual Change Scores*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 2.2 | 1.39 | 0.16 | -1.04 |
| Discussion condition | 20 | 1.45 | 1.04 | -0.32 | -0.8 |

**Self-efficacy.** Participants' scores in the OSS-SF before the intervention (pre-test) are compared with their scores after the intervention (post-test) using a paired-samples *t*-test. This is done separately for both groups. The article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). An overview of the scores is presented in table 39. Participants in the article condition show significantly higher levels of self-efficacy in the post-test than in the pre-test; $t(21) = 6.95$, $p < .001$. Cohen's *d* is 1.48, which constitutes a large effect (Cohen, 1992). In the discussion condition, participants differ significantly in their self-efficacy scores in the pre-test compared to the post-test; $t(19) = 2.98$, $p < .05$. Cohen's *d* is 0.67, which constitutes a medium effect (Cohen, 1992).

Table 39

*Self-efficacy Scores in Both Conditions*

| | Article condition | | Discussion condition | |
|---|---|---|---|---|
| Self-efficacy | *M* | *SD* | *M* | *SD* |
| Pre-test | 2.83 | 0.64 | 3.07 | 0.71 |
| Post-test | 3.51 | 0.49 | 3.51 | 0.70 |

To compare changes in scores between groups, an independent sample *t*-test was conducted. Change in scores in the article group and discussion group distributions were sufficiently normal for using a *t*-test (i.e., skew <|2.0| and kurtosis < |9.0|; Schmider et al., 2010). See table 40. Additionally, the criteria of homogeneity of equal variances was satisfied using Levene's *F*-test, $F(1, 41)= 1.85$, $p = .212$. The independent sample *t*-test revealed no significant effect, $t(40) = 1.36$, $p = .181$. Thus, the increase in self-efficacy in the article group was not significantly larger than in the discussion group.

Table 40

*Descriptive Statistics Associated With Change in Self-efficacy*

| | *N* | *M* | *SD* | Skew | Kurtosis |
|---|---|---|---|---|---|
| Article condition | 22 | 0.68 | 0.46 | 0.16 | -0.86 |
| Discussion condition | 20 | 0.44 | 0.66 | 0.31 | 0.003 |

In summary, self-efficacy scores increased in both the article and the discussion condition from the pre-test to the post-test. However, improvement in one condition was not significantly different from the other condition.

**Intervention Check.** Results revealed no difference between conditions with regard to how satisfied learners were with the outline they produced. Additionally, participants in both conditions found transferring the face-to-face outline into a MOOC outline equally difficult. Further, they mentioned similar difficulties that they experienced during the design phase. The different workshops appear to have been similar enough, apart from the intervention, to allow for comparing the two conditions. This notion is further support by the fact that all but five participants over both conditions stated the intervention as their inspiration for making changes to their outlines. Also, significantly more participants found the intervention, rather than any other activity, to be the most helpful aspect of the workshop. Thus, the group discussion and the individual reading respectively appear to have been the critical aspect of the workshops with regard to having an influence on participants.

## Discussion and Conclusion

This study compares the effects of a group discussion to those of individual reading on conceptual change, quality of MOOC outlines and self-efficacy. Group discussions were expected to be more efficient in increasing the quality of MOOC outlines and participants' self-efficacy. Further, it was hypothesized that group discussions are more beneficial to conceptual change than individual reading. Results are discussed and related to each other in the following paragraphs.

### Conceptual Change

Several measures in this study attempted to evaluate whether participants experienced a conceptual change. On the one hand, a variety of questionnaires which included items to measure conceptual change were used: MOOC Concept, MOOC Questionnaire, Design Experience Questionnaires and Change Reflection Questionnaires. On the other hand, the rating rubric included a category which counted indicators of conceptual change.

Results indicate a significant increase in conceptual change for both the article reading intervention as well as the group discussion intervention. Based on the MOOC outlines, results showed no evidence for one intervention to be more effective than the other with regard to conceptual change. However, effects for conditions have been found when specific thresholds are considered. The evidence elaborated below provides support for the passing of three out of four particularly relevant thresholds by a significant number of participants. These four thresholds are:

1. Educators often do not realize the distinctive nature of the online learning environment, which does not imitate face-to-face education.
2. Different material is used for online courses to encourage interaction with and among learners.
3. Learning does not happen through passive absorption of knowledge, but through interaction and active knowledge construction
4. The content should be humanized by packaging it in a story and making the educator visible (Northcote et al., 2011).

First, educators should understand that MOOCs follow a distinct structure and are substantially different from face-to-face courses (Northcote et al., 2011). Results make this evident when comparing the perceived differences between MOOCs and face-to-face courses. In both the individual reading condition as well as the group discussion condition, participants had a similar focus before the intervention, which was flexibility. After the intervention, participants had more diverse ideas about the differences. Possibly, this increase in diversity was due to the time participants spend on designing their own MOOC outlines. Thus, the differences they listed after the intervention, at the end of the workshop, might be influenced by the design they had in mind for their own MOOC. Thus, it remains unclear if the intervention had an effect on the diversity of differences, or if the mere exercise of designing an own MOOC outline would suffice as an explanation. Even though both conditions were equal with regard to the differences mentioned by participants, there was a significant increase in the number of categories mentioned for the main purpose of a MOOC in the group discussion condition, but not in the individual reading condition. Thus, the cooperative exchange of ideas might have inspired participants to see a greater variety of main purposes and allowed them to reconsider what they believed to be the main purpose of a MOOC. Additional evidence for the beneficial influence of the group discussion condition is provided by how participants proposed to structure and progress through a MOOC. In both the individual reading condition and the group discussion condition, the number of participants who proposed a linear progression decreased significantly after the intervention. However, in the group discussion condition significantly more participants mentioned a flexible structure as an alternative than in the individual reading condition. The cooperative learning activity thus stimulated conceptual change with regard to the structure of a MOOC better than the individual reading condition. The support for a flexible structure in the group discussion condition could be interpreted as a focus on finding a solution (flexible structure) rather than avoiding a problem (linear structure). Learners in a discussion are exposed to more ideas and might thus think more solution oriented than problem oriented. This is in line with research on conceptual change, that has found cooperative learning activities to encourage conceptual restructuring (Bilgin, 2006; Liu & Hmelo-Silver, 2009).

Second, educators need to realize that different material should be used for MOOCs compared to face-to-face courses (Northcote et al., 2011). For this reason, participants were asked what a video in a MOOC should look like. There is a significant improvement in both the individual reading condition as well as the group discussion condition for video content. Participants in both conditions stressed the importance of videos to be engaging significantly more often after their interventions. Additionally, the number of participants suggesting PowerPoint slides as a foundation for a video in a MOOC decreased to zero. Neither intervention appeared more effective with regard to video content. However, the group discussion condition was significantly more effective with regard to the appropriate length of videos in a MOOC, as participants decreased the suggested length of an ideal video to a greater degree than in the individual reading condition and consequently approximated the ideal length of 5 minutes. Thus, both individual reading and group discussion appear to positively effect conceptual change with regard to video content, but only group discussion also contributes to conceptual change with regard to video length. This is in line with research that attributes the detection of misconceptions to sharing ideas and experiences (Bilgin, 2006; Liu & Hmelo-Silver, 2009).

Third, educators should understand that knowledge is not passively absorbed, but actively constructed through interaction (Northcote et al., 2011). The results show that participants in both conditions have passed this threshold after the intervention, because their focus moved from surface characteristics of MOOCs to social and interactive characteristics when asked for the main differences between face-to-face courses and MOOCs. Additionally, participants in both conditions stated social learning as the main purpose of MOOCs significantly more often after the intervention. When asked how a MOOC should be structured, participants in both conditions mentioned the importance of discussions significantly more often after the intervention than before the intervention. This effect was more pronounced for the individual reading condition. Thus, both individual reading as well as cooperative learning seem to benefit conceptual change with regard to the integration of social learning in MOOCs. These results are in line with the researcher's expectations. They are not surprising, as both the article as well as the group discussion had a strong focus on social and interactive aspects of MOOCs. Thus, it is very likely that these aspects were salient in participants' thoughts when considering the main purpose and structure of a MOOC. It seems likely that the interventions were crucial to passing this threshold, because their content is closely related to this threshold. However, to eliminate the possibility that the mere practice of designing MOOC outlines was responsible for this effect, a control group with no intervention would need to be included in future research.

Fourth, the educator should be visible and a MOOC should tell a story (Northcote et al., 2011). There was no significant evidence for participants to have passed or even come close to this threshold. Perhaps the emphasis in the article and group discussion on this threshold was not strong enough to manifest in participants' thoughts. This could be tested for by redesigning the interventions. It is further possible that the limited amount of time for the intervention did not allow participants to internalise more than three threshold concepts. Possibly, more time should be planned for the intervention in future research.

When considering the perception of conceptual change in participants, evidence suggests a beneficial effect of the group discussion intervention. Not only did significantly more participants in the discussion condition experience a conceptual change, the group discussion intervention was also significantly more effective in producing a conceptual change during the intervention, rather than at any other point during the workshop. These results are in line with the researcher's expectations, as she assumed the group discussion to facilitate conceptual change. Apparently, the group discussion activity stimulated participants to reconsider their beliefs and helped them to let go of their misconceptions. Literature agrees that cooperative learning activities promote conceptual restructuring and thus are beneficial for conceptual change (Bilgin, 2006; Liu & Hmelo-Silver, 2009). The researcher therefore assumes the group discussion to constitute the pivotal building block of the discussion condition in facilitating conceptual change. Perhaps the MOOC design exercises also play an important role in combination with the group discussion. In order to test for this possibility, future research could include one condition with design exercises before and after the intervention, one condition with design exercises only before the intervention, one condition with design exercises only after the intervention and one condition without any design exercises at all. This would allow more clarification if and at what point in the workshop the MOOC design exercises are beneficial for conceptual change.

In summary, there was no significant difference in strength of conceptual change between conditions in the MOOC outlines. Both cooperative learning and individual reading have been shown to stimulate conceptual change on various aspects. However, when closely examining the results, differences between conditions become apparent. Individual reading appears to enhance conceptual change on social and interactive aspects in MOOCs, while cooperative learning seems to facilitate conceptual change on the structure of a MOOC and the lengths of videos in a MOOC. Further, learners who participate in cooperative learning experience conceptual change during those activities to a greater degree than learners who read individually do during their reading. Consequently, any intervention aiming for conceptual change should include cooperative learning activities. However, reserving time for an individual reading activity on social and interactive aspects in MOOCs is advisable. Possibly, other topics could be covered in the individual reading part, only to then be discussed with the group. This combination might appeal to learners who prefer individual learning and yet still reap the benefits for conceptual change, as there is a cooperative activity.

**Quality**

This study examined the quality of the MOOC outlines that participants produced by evaluating them using a rating rubric. Outlines were scored and compared intrapersonally as well as interpersonally between conditions.

The results of this research suggest that the group discussion intervention as well as the article reading intervention were effective in increasing the quality of MOOC outlines. As expected, there was a significant increase in quality of MOOC outlines from the pre-test to the post-test for both groups not only in the overall score, but also in all individual categories that were evaluated (introduction, content, interaction, transfer of knowledge and structure). Consequently, these results support the claim that both cooperative learning and individual learning can lead to improved quality, as both are a form of scaffolding. These results are in line with research that has found scaffolding to increase academic achievement and therefor quality of deliverables (Lin et al., 2012; Lin & Liu, 2014). However, it is possible that the mere repetition of the task could have increased the quality of MOOC outlines even without any intervention. To eliminate this doubt, it is advisable to include a control group which does not receive any intervention in future research on this matter.

Surprisingly, the quality of MOOC outlines which were generated after the article reading intervention seemed equal to the quality of outlines produced after the group discussion intervention. Although there was no evidence of a significant difference in quality of outlines between conditions, a closer look at the scores revealed that outlines in the article condition received higher scores for introduction, interaction, transfer of knowledge, structure and overall. The only category in which the discussion condition received higher scores was content. Yet, these differences are not of statistical significance, so further research is necessary to evaluate if individual reading interventions could possibly lead to higher quality outlines.

Even though there was no detectable difference for final scores between conditions, there were significant differences in improvement of scores from the first MOOC outlines and the second MOOC outline between conditions. Outlines of the article reading condition improved significantly more than outlines of the group discussion condition in the categories introduction and transfer of knowledge, as well as in the overall score. In all other categories, participants in the article condition improved their scores to a greater degree, but the difference was not statistically significant. Again, this effect to the benefit of the article condition was surprising. Perhaps participants in the article condition benefitted from taking notes while reading and thus rehearsing the information better. Consequently, it might have been more readily available when they worked on their final outlines. Especially when compared with participants in the discussion condition, who have had contact with a broader variety of input and opinions on the different topics. Possibly, participants in the discussion condition would have benefitted from additional time and stimulation to sort through what they had learned during the discussion, rather than being thrown right into the final designing exercise. Future research could provide learners with an exercise to reflect on the discussion and structure their thoughts after the discussion has finished.

Another potential reason for this finding is that participants were university students. This population group naturally spends a great amount of time learning information from literature and can thus be considered experts in gaining knowledge by reading. The experience with group discussions, on the other hand, varies between students. Although each learner participated in the group discussion, there was a wide variety in learners' engagement. Possibly, some participants did not learn much during the group discussion simply because their attention was elsewhere or because they are not used to gaining information from a source like that. Perhaps results would be different for a group of adult learners, who learn more by interacting with colleagues than by reading individually in their daily lives.

In summary, individual and cooperative learning interventions both seemed effective for increasing the quality of MOOC outlines. However, evidence suggests individual reading to be more effective in increasing the overall quality as well as the quality of introductions and transfer of knowledge. Consequently, individual reading is recommended as part of an intervention that aims to increase quality of MOOC outlines.

**Self-efficacy**

This study measured self-efficacy using a questionnaire. Scores were compared intrapersonally as well as interpersonally.

The results of this study suggest that both an individual learning intervention as well as a group learning intervention are effective in increasing self-efficacy in course designers developing a MOOC, as there was a significant increase in self-efficacy from the pre-test before the intervention to the post-test after the intervention in both conditions. The increase in self-efficacy after the group discussion intervention is in line with the hypothesis of this study and can be explained with using Bandura's (1997) work on self-efficacy. According to Bandura, the most powerful way to increase self-efficacy is by letting learners form mastery experiences. As the majority of learners were complete novices in the field of MOOC design, it is plausible that their self-efficacy increased after they have had the opportunity to design their first outline. However, this raises the possibility that no intervention is needed to increase learners' self-efficacy, as more self-efficacy is gained with every task that novices complete. Yet, feedback, termed verbal persuasion by Bandura, is a crucial part to increasing self-efficacy. The information from the group discussion or the individual reading have both served as a source of feedback on what was already in participant's outlines and what might still need to be included. Thus, it seems likely that participants gained self-efficacy after mastering the exercises and receiving feedback on them as part of the intervention. Nevertheless, a control group which receives no intervention at all could be included in future studies to ensure that self-efficacy increases as a result of the intervention. Yet, the question remained if the cooperative intervention was more effective in increasing self-efficacy than the individual intervention.

There was no significant evidence for the group discussion intervention to be more effective than the individual reading intervention in increasing participants' self-efficacy. This finding is in line with research by Ryan, Bordoloi and Harrison (2000), who found that learners who participated in group learning in a data-modelling task did not outperform participants in the individual learning condition with regard to their self-efficacy. The authors reasoned that their novice learners might not have experienced sufficient modelling of the skills to be learned in their group learning activity, so their vicarious experience was limited, thereby hindering the growth of self-efficacy. As the present study's participants are novices in the field of MOOC design as well, a similar explanation for the lack of difference between individual and cooperative learning with regard to self-efficacy is possible. Although the group discussion was moderated by an expert, it might be that the discussion was not concrete enough for learners to gain vicarious experiences. In a recent study, Fernandez-Rio, Cecchini, Méndez-Gimenez, Mendez-Alonso and Prieto (2017) found the influence of self-regulated learning on self-efficacy to be stronger than the influence of cooperative learning on self-efficacy. Thus, participants in the individual reading intervention might have benefitted from being able to self-regulate their learning, while participants in the group discussion intervention did not have this opportunity. Possibly, denying learners in a cooperative learning intervention the opportunity to self-regulate their learning counteracts any benefits of this method in increasing self-efficacy. To further investigate this question, a cooperative learning intervention that allows for a certain degree of self-regulated learning could be included in a future study.

In summary, an intervention which provides feedback to the learner is likely to increase self-efficacy. However, further research is needed to clarify if cooperative learning interventions are more effective in increasing self-efficacy than individual learning interventions. Current evidence does not allow for a recommendation of one type of intervention over the other with regard to self-efficacy.

Considering all aspects of this research, cooperative learning has been proven to facilitate conceptual change in educators, improve quality of MOOCs and increase self-efficacy in educators. However, results indicate the same effects to be true for individual reading. Nonetheless, differences in the effects of the two types of learning have been discovered. If the main purpose of an intervention is to stimulate conceptual change, cooperative learning is the method of choice. Cooperative learning did not only show positive effects on conceptual change in more areas than individual reading, learners also reported to experience a conceptual change more often after participating in a cooperative learning activity than they did after individual reading. However, if the goal of an intervention is to increase the quality of deliverables that learners are producing, individual reading should be preferred, as it promises to be more beneficial. In terms of increasing educators' self-efficacy, neither cooperative learning nor individual reading proved to be more effective than the other in this study. Overall, both types of intervention can be considered effective, as they help learners to realise the difference between face-to-face course design and MOOC design and thus solve the problem commonly experienced by educators in the transfer from traditional education to MOOCs. Both interventions lead to MOOCs of higher quality and increased self-efficacy in educators.

The researcher recommends a combination of cooperative learning elements and individual reading for future interventions. Basic knowledge should be acquired during individual reading time and then be discussed in cooperative learning activities. This will likely reap the benefits of both types of learning. It is recommended to test this type of hybrid intervention in future research. Further, any future research should include a control group which does not include an intervention. Although most effects are very likely the results of the intervention, there can be no certainty due to the lack of a control condition. Another limitation of this research is that the sample consisted solely of university students with very limited teaching experience. Although MOOC designers are usually novices in the online teaching area, they are typically experienced teachers. To evaluate if the findings of this research are generalizable to the population of teachers, future research with that population group is needed.

**List of References**


Adamopoulos, P. (2013). What makes a great MOOC? An interdisciplinary analysis of student retention in online courses. *Proceedings of the 34th International Conference on Information Systems, ICIS, (2013).*

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

Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York: Freeman.

Bilgin, I. (2006). Promoting pre-service elementary students' understanding of chemical equilibrium through discussions in small groups. *International Journal of Science and Mathematics Education, 4*(3), 467-484.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences . Hilsdale. *NJ: Lawrence Earlbaum Associates*, *2*.

Cohen, J. (1992). A power primer. *Psychological bulletin, 112*, 155-159.

Cropper, M. H., Bentley, J. P., & Schroder, K. (2009). How Well Do High-Quality Online Courses Employ Merrill's First Principles of Instruction?. *In Educational media and technology yearbook* (pp. 121-140). Springer US.

da Costa, C., & Galembeck, E. (2016). The evolution of the Krebs cycle: A promising subject for meaningful learning of biochemistry. *Biochemistry and Molecular Biology Education, 44*(3), 288-296.

Doolittle, P. E. (1997). Vygotsky's zone of proximal development as a theoretical foundation for cooperative learning. *Journal on Excellence in College Teaching, 8*(1), 83-103.

Fernandez-Rio, J., Cecchini, J. A., Méndez-Gimenez, A., Mendez-Alonso, D., & Prieto, J. A. (2017). Self-regulation, cooperative learning, and academic self-efficacy: Interactions to prevent school failure*. Frontiers in psychology, 8.*

Gagné, R.M. (1985). *The conditions of learning and theory of instruction* (4th Ed.). New York: Holt, Rinehart and Winston.

Glušac, D., Karuović, D., & Milanov, D. (2015). Massive open online courses-pedagogical overview. In *Carpathian Control Conference (ICCC), 2015 16th International (pp. 142-146*). IEEE.

Gosselin, K. P., & Northcote, M. (2013). Cross-continental research collaborations about online teaching. In D. G. Sampson, J. M. Spector, D. Ifenthaler, & P. Isaias (Eds.), *Proceedings of the IADIS International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA)* (pp. 282-289). Fort Worth, Texas.

Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC design principles: A pedagogical approach from the learner's perspective. *eLearning Papers, 33*(4).

Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *The Journal of the Learning Sciences*, *15*(1), 53-61.

Jasnani, P. (2013). Designing MOOCs. *A White Paper on Instructional Design for MOOCs. Publisher: Tata Interactive Systems.*

Johnson, D. W., & Johnson, R. T. (1994). *Learning together and alone. Cooperative, competitive, and individualistic learning.* Allyn and Bacon, 160 Gould Street, Needham Heights, MA 02194.

Karlsson, N., Godhe, A. L., Bradley, L., & Lindström, B. (2014). Professional Development of Teachers in a MOOC. In *Proceedings of the 22nd International Conference on Computers in Education, Nara, Japan* (pp. 2-4).

Kopp, M., & Lackner, E. (2014). Do MOOCs need a special instructional design. *EDULEARN14 Proceedings*, 7138-7147.

Li, M. P., & Lam, B. H. (2013). Cooperative learning. The Active Classroom, *The Hong Kong Institute of Education.*

Lin, C., & Liu, E. Z. F. (2014, July). The Effect of Self-Efficacy and Scaffolding on Students' Programming Learning in Robotics Learning Context. In *Ubi-Media Computing and Workshops (UMEDIA), 2014 7th International Conference on* (pp. 292-296). IEEE.

Lin, T. C., Hsu, Y. S., Lin, S. S., Changlai, M. L., Yang, K. Y., & Lai, T. L. (2012). A review of empirical evidence on scaffolding for science education. *International Journal of Science and Mathematics Education*, *10*(2), 437-455.

Liu, L., & Hmelo-Silver, C. E. (2009, June). Collaborative scientific conceptual change: a framework for analyzing science learning. In *Proceedings of the 9th international conference on Computer supported collaborative learning-Volume 1* (pp. 194-198). International Society of the Learning Sciences.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of educational research, 71*(3), 449-521.

MacLeod, H., Haywood, J., Woodgate, A. & Sinclair, C. (2014) Designing for the Unknown Learner. Under review, retrieved from European MOOCs Stakeholders Summit papers, 245-248 http://emoocs2014.eu/sites/default/les/Proceedings-Moocs-Sumit-2014.pdf

Margaryan, A., Bianco, M., & Littlejohn, A. (2014). Instructional quality of massive open online courses (MOOCs). *Computers & Education*, *80*, 77-83.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 50-60.

Merrill, M. D. (2001). Star instructional design rating from http://medsci.indiana.edu/c602web/tbl/reading/merrill.pdf

Meyer, J. H., & Land, R. (2003). *Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines* (pp. 412-424). Edinburgh: University of Edinburgh.

Mills, R., Tomas, L., & Lewthwaite, B. (2016). Learning in Earth and space science: a review of conceptual change instructional approaches. *International Journal of Science Education, 38*(5), 767-790.

Moll, L. C. (Ed.). (1992). *Vygotsky and education: Instructional implications and applications of sociohistorical psychology*. Cambridge University Press.

Northcote, M. T., Reynaud, D., Beamish, P., Martin, T., & Gosselin, K. P. (2011). Bumpy moments and joyful breakthroughs: The place of threshold concepts in academic staff development programs about online learning and teaching. *Critical Perspectives on Communication, Cultural & Policy Studies, 30*(2), 75-89.

Northcote, M., Gosselin, K. P., Reynaud, D., Kilgour, P., & Anderson, M. (2015). Navigating learning journeys of online teachers: Threshold concepts and self-efficacy. *Issues in Educational Research, 25*(3), 319-344.

Panitz, T.(1996). *A Definition of Collaborative vs Cooperative Learning.* Deliberations, London Metropolitan University; UK., Retrieved 20 Dec. 2011, from: http://www.londonmet.ac.uk/deliberations/collaborative-learning/panitz-paper.cfm.

Pea, R. D. (1993). Learning scientific concepts through material and social activities: Conversational analysis meets conceptual change. *Educational psychologist, 28*(3), 265-277.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences*, *13*(3), 423-451.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of engineering education, 93*(3), 223-231.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science education, 66*(2), 211-227.

Rigotti, T., Schyns, B., & Mohr, G. (2008). A short version of the occupational self-efficacy scale: Structural and construct validity across five countries. *Journal of Career Assessment*, *16*(2), 238-255.

Ritchhart, R., Church, M., & Morrison, K. (2011). *Making thinking visible: How to promote engagement, understanding, and independence for all learners.* John Wiley & Sons.

Roehl, A., Reddy, S. L., & Shannon, G. J. (2013). The flipped classroom: An opportunity to engage millennial students through active learning. *Journal of Family and Consumer Sciences, 105*(2), 44.

Ryan, S. D., Bordoloi, B., & Harrison, D. A. (2000). Acquiring conceptual data modeling skills: The effect of cooperative learning and self-efficacy on learning outcomes. *ACM SIGMIS Database, 31*(4), 9-24.

Scagnoli, N. I. (2012). Instructional Design of a MOOC.

Schunk, D. H., & Pajares, F. (2005). Competence perceptions and academic functioning. *Handbook of Competence and Motivation*, *85*, 104.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 6*, 147-151.

Sharma, P., & Hannafin, M. J. (2007). Scaffolding in technology-enhanced learning environments. *Interactive Learning Environments*, *15*(1), 27-46.

Shelton, K., & Saltsman, G. (2008). Applying the ADDIE model to online instruction. *Adapting information and communication technologies for effective education, 2,* 40-57.

Shepherd, C., Alpert, M., & Koeller, M. (2007). Increasing the efficacy of educators teaching online. *International Journal of Social Sciences, 2*(3), 172-78.

Slavin, R. E. (1991). Synthesis of research of cooperative learning. *Educational leadership, 48*(5), 71-82.

Terenzini, P. T., Cabrera, A. F., Colbeck, C. L., Parente, J. M., & Bjorklund, S. A. (2001). Collaborative learning vs. lecture/discussion: Students' reported learning gains. *Journal of Engineering Education, 90*(1), 123-130.

Varonis, E. M. (2014). Most courses are not born digital: An overview of the Quality Matters peer review process for online course design. *Campus-Wide Information Systems, 31(4)*, 217-229.

Vygotsky, L. S. (1980). *Mind in Society: The Development of Higher Psychological Processes*. Harvard university press.

Walker, S. L., & Fraser, B. J. (2005). Development and validation of an instrument for assessing distance education learning environments in higher education: The Distance Education Learning Environments Survey (DELES). *Learning Environments Research, 8*(3), 289- 308.

Wu, L., & Looi, C. K. (2013). Incorporation of Agent Prompts as Scaffolding of Reflection in an Intelligent Learning Environment. In *Intelligent and Adaptive Educational-Learning Systems* (pp. 369-391). Springer Berlin Heidelberg.

Zaiontz C. (2017) *Real Statistics Using Excel.* Retrieved from www.real-statistics.com.

Zayapragassarazan, Z., & Kumar, S. (2012). Active learning methods. *Online Submission, 19*(1), 3-5.

## Appendix

Questionnaires used in this study are presented below.

MOOC Concept

In your opinion, what are the main differences between a MOOC/online course and a face-to-face course? Please list as many as you can think of and explain each difference.

| **Difference** | **MOOC/online course** | **Face-to-face/ traditional course** |
|---|---|---|
| | | |
| | | |
| | | |

MOOC Questionnaire

Please answer the following questions. There are no right or wrong answers.

In your opinion, what is the main purpose of a MOOC?

How should a MOOC be structured?

What does a good video in a MOOC look like? (Length, content, etc.)

Design Experience Questionnaire

There are a few last questions you are asked to answer. These are reflective questions about the workshop. Please take a moment to think about each question and write a detailed answer. There are no right or wrong answers.

Do you think you have created a good outline for a MOOC? Please explain.

Was it difficult for you to transfer the face-to-face course outline into a MOOC outline? Please explain.

What was the most difficult part in transferring the face-to-face outline into the online outline?

Did you realize the fundamental difference between face-to-face and online course design at any one point during the workshop? If so, at what point.

Is there any activity that you felt was particularly helpful for you?

Change Reflection Questionnaire

Please take a moment to reflect on the changes you made from your first MOOC outline to the version you have just send in. What did you change?

What inspired you to make those changes? For example something that was discussed or something you read about. Please be specific in your answer.

Self-efficacy

How confident are you that you can…

I can remain calm when facing difficulties in MOOC design because I can rely on my abilities.

| Not at all true | | | | Very true |
|---|---|---|---|---|
| o | o | o | o | o |

When I am confronted with a problem in MOOC design, I can usually find several solutions.

| Not at all true | | | | Very true |
|---|---|---|---|---|
| o | o | o | o | o |

Whatever comes my way in MOOC design, I can usually handle it.

| Not at all true | | | | Very true |
|---|---|---|---|---|
| o | o | o | o | o |

My past experiences in MOOC design have prepared me well for my future.

| Not at all true | | | | Very true |
|---|---|---|---|---|
| o | o | o | o | o |

I meet the goals that I set for myself in MOOC design.

| Not at all true | | | | Very true |
|---|---|---|---|---|
| o | o | o | o | o |

I feel prepared for most of the demands in MOOC design.

| Not at all true | | | | Very true |
|---|---|---|---|---|
| o | o | o | o | o |

**Outline Rating Rubric**

**Participant:**

**Scoring: Yes=1, No=0**

**Introduction**

| | F2F Outline | MOOC first outline | MOOC final outline |
|---|---|---|---|
| The outline mentions an introduction to the topic | | | |
| The outline mentions a learning objective | | | |
| The outline mentions a Big Question | | | |
| The outline mentions the introduction of educators | | | |
| Total | **0** | **0** | **0** |

**Content**

| | | | |
|---|---|---|---|
| The outline includes use of relevant media | | | |
| The outline includes examples | | | |
| The outline includes a variety (at least 3 different forms) of learning activities | | | |
| The outline includes opportunities for discussion | | | |
| The outline includes opportunities for reflection | | | |
| The outline includes a form of assessment | | | |
| Total | **0** | **0** | **0** |

**Interaction**

| | | | |
|---|---|---|---|
| The outline includes opportunity to ask questions | | | |
| The outline includes opportunity for feedback to the learner (from peers or educator) | | | |
| The outline promotes social learning (discussions etc.) | | | |
| Total | **0** | **0** | **0** |

**Transfer of knowledge**

| | | | |
|---|---|---|---|
| The outline includes an opportunity for learners to apply their new knowledge in a different situation | | | |
| The outline includes activation of prior knowledge | | | |
| The outline includes a real world problem | | | |
| The outline provides an opportunity to practice | | | |
| Total | **0** | **0** | **0** |

| | | | |
|---|---|---|---|
| **Structure** | | | |
| The outline provides a clear structure | | | |
| The outline includes a storyline / central theme | | | |
| Duration of each activity is appropriate | | | |
| The outline matches the task (MOOC or traditional) | | | |
| Total | **0** | **0** | **0** |

| | | | |
|---|---|---|---|
| **Total** | **0** | **0** | **0** |

| | | | |
|---|---|---|---|
| Indicators of Conceptual Change | 0 | 0 | 0 |

Material: Article in article condition

# FutureLearn Learning Design Guidelines

## Objective

The objective of this document is to offer guidelines for the learning design of FutureLearn courses, covering the top-level structure of steps and other elements (such as course emails). It gives some basic principles of learning design, followed by guidelines for learning design with some examples from FutureLearn courses, then outline templates for typical course structures.

## What is learning design?

The basic aim of learning design is to take a systematic approach to designing and planning good courses, based on successful learning theory and practice, along with evidence gathered from previous courses.

Online courses are complex systems that have to be managed by focusing on the most important elements, then setting appropriate guidelines, structures and constraints. A central issue to consider is ‘what will success look like?’ – for learners, educators, the institution, and possibly funders and associated bodies. The course should be designed to meet your specific criteria for success.

## Learning design guidelines

Here are some guidelines to consider, based on our experience of developing and improving FutureLearn courses.

### 1. Work with the FutureLearn platform

No educational technology is ‘pedagogy neutral’ - they all have built-in educational structures. Social learning is expressed in the FutureLearn product vision: “Inspire the best learning experiences by telling stories, provoking conversations and celebrating progress”.

FutureLearn courses are structured into steps, activities and weeks (not content, units and modules). We have educators, not instructors; and Learners, not students. The platform supports comment and discussion, not forums. This structure matches the FutureLearn pedagogy. Steps enable learners to hold conversations in the context of educational content. Activities are ways of organising steps to support active learning. Weeks suggest a progression of activities and building of ideas over time. The aim is for Educators to support the building of knowledge, rather than to instruct. We want to distinguish FutureLearn Learners from students taking university degree courses. And Comments and Discussions suggest a more immediate and informal conversation than a separated forum.

We recommend that learners should engage with the course for about 3-6 hours per week, which includes watching and reading course materials, taking part in discussions, writing and reviewing assignments, and engaging in reflective study and additional research. The course should be designed to retain the interest of busy learners and integrate the study into their daily lives. The course should be designed so that learners can pause after any step and then quickly continue with the content and the conversation where they left off.

Each step should be designed with content and structure appropriate to the course, and a sequence of steps is grouped together into an activity, with a title that describes the educational theme or objective of the activity. For example, a 'What is climate change?' activity to introduce the topic might consist of a Video, Article, Discussion, Quiz, Article sequence of steps, where the final article is a synthesis of the material. Thus, there is flexibility in the design of each step, and these can be ordered to create good learning activities using FutureLearn elements.

**2. Start from a learning objective or big question**

Designing a course around one or more learning objectives or big questions gives a focus to the course. It also offers a learner-centred way to assess the effectiveness of the course, by asking learners at the end to give their answers to the question. And it provides an effective way to market the course and entice people to join. Some example questions from FutureLearn courses are:

“Should we geo-engineer our climate?” (Climate Change) “Why is the sun burning so slowly?” (Higgs Boson) “Is your brain just like your desktop computer?” (Good Brain, Bad Brain)

A good approach would be to display the objectives or question in the course description page, and in the first course email. This might be combined with an initial course discussion where learners suggest their own learning objectives, or questions they would like answered. At the end of the course, the questions can be re-visited in a course summary followed by a reflective discussion with the learners saying whether their personal goals have been met.

**3. Every step should contribute to learning**

It is good practice to start a course by introducing the educators and allowing learners to introduce themselves, but this should be a valuable learning experience. Feedback shows that learners respond well to a course led by an enthusiastic and knowledgeable educator, but just as an educational TV programme doesn’t start with an introduction to the narrator and production crew, consider other ways to start the course than a series of introductory steps.

The introduction could be placed after a first step with a snappy title, that poses the big course question, to provoke curiosity or wonder. Or you could use a Discussion step to introduce the educator and facilitators. A good example is to have the lead educator a) ask learners to give a structured introduction to why they are taking the course, and b) introduce himself through a model response. The facilitators can then add their profiles and interests to the discussion.

**4. Design a balanced mix of learning activities**

Varying the types of learning activity will engage learners with differing approaches to learning and provide variety. The exact mix will depend on the course topic and pedagogy. A typical course may have 4 or 5 types of activity in each week. Note that conversational, networked and browsing learning activities are all 'baked into' the FutureLearn platform, but they need to be supported by appropriate educator comments and course emails. A learning activity may embrace more than one type of learning.

| **Learning** | **Learners...** | **FutureLearn example** |
|---|---|---|
| Delivered | are presented with information | Video step |
| Reflective | reflect on activities | Discussion to review previous steps |
| Collaborative | construct shared understanding | Discussion step to explore learners' perspectives and experiences |
| Conversational | converse with others | Comments linked with each step |
| Networked | interact with networks of peers | 'Like' comments and 'follow' other learners |
| Browsing | seek and collate information | Use ToDo to browse the course steps |
| Assessing | learn by receiving constructive feedback | Peer review step |
| Inquiry-driven | investigate authentic situations | Hadrian's Wall course: Exercise step to investigate terrain of Hadrian's Wall with an overlay on Google Maps |
| Problem-solving | try to solve problems or answer big questions | Ebola: Symptoms, History and Origins course: What can be done to stop the spread of the virus? |
| Case-based | investigate individual cases | Introduction to Forensic Science course: murder case presented by series of videos |
| Simulation | interact with a simulated world | Moons course: Exercise step with a 'virtual microscope' |
| Construction | design and make artefacts | Creative Coding course: learners construct computer programs and share results |
| Cross-context | learn across physical or social settings | Ecosystems course: learners use iSpot software to photograph and identify wildlife |
| Game-based | engage with game environments | Moons course: Moon Trumps game |
| Performative | present for an audience | Start Writing Fiction course: learners create short stories for an audience of other learners |

**5. Design for storytelling**

The FutureLearn platform is built to support narrative. The course steps are 'building blocks' that can be put together in different combinations to create flows of activity that drive the learning forwards. The learning content for each step can be designed to assist these flows.

As an example, Week 1 of the Secret Power of Brands course starts with a video of people from around the world talking about the brands they love, to raise interest and show the scope of the course. Then learners are asked in a Discussion step what they want to get from the course. That is followed by a big question - "What is a brand?" - to motivate a sequence of videos to address the question from practitioners and academics. This leads on to a Discussion among the learners about how brands impact their world, and ends with a Test for learners to review their new knowledge, followed by short structured Articles on “Five ways to learn more” and “Our top 20 books to read on brand”. The whole week is a narrative structure that motivates, questions, explains, discusses, reflects, and extends.

On a larger scale, the Forensic Science course is presented week by week through an unfolding story, based on a real murder that took place in 2013. Each week, the learners are challenged to interpret the scene and work out who committed the murder, using the forensic techniques (e.g. fingerprint analysis and blood pattern analysis) introduced that week. The final week encourages learners to offer their views based on the evidence and to cast their vote as members of the ‘jury’. The course is rounded off the the result of the poll of learners and the jury result in the actual case. Thus, a typical ‘soap opera’ story narrative, with tensions, cliffhangers, and reveals was adapted very effectively to the subject of the course.

Not every course can, nor should, fit into an overarching narrative. But FutureLearn supports storytelling as a way of creating coherence and drive for online courses where the learners must motivate themselves and push forward without the benefit of scheduled seminars and lectures.

**6. Design for conversation**

The learning theory that underlies the FutureLearn platform is ‘learning as conversation’. It sees learning as a continuing conversation, with oneself (as we try to interpret and reconcile pieces of knowledge), with teachers, and with peers.

To make the most of learning opportunities, every step should provoke and model conversation. To this end, each Article, Text, Video/audio, and Exercise step is linked with a free-flowing discussion. These are not Moodle forums! They are more like chats around a water-cooler, designed to be informal, easy to enter, and with a simple choice to read, reply, or contribute to the flow. We may hope a learner will, for a step, read the first few comments and replies, scroll down to see some more, click ‘most liked’ or ‘following’ to see the most interesting or relevant contributions, and then perhaps add a reply. Typically, learners will not know how to do this, so the course design can model and prompt discussions, for example by asking learners to summarise the key points from a video, or to offer their own perspective on an issue.

The Discussion step is intended for more focused discussion. David Major from FutureLearn has suggested three types of Discussion:

- *water-coolers* are the free-flowing conversations where the learner replies to an immediate comment or adds to the flow
- *mountains* build knowledge by learners contributing their experiences, their perspectives on a question or issue

- *capstones* round off a topic by encouraging learners to extract key points or synthesise their understanding. A valuable and popular type of capstone discussion is where learners suggest and prioritise (by liking) the questions to be answered by staff in a Google hangout or course email.

Peer Review steps provide more directed assignment-review-reflect conversations. It is a good idea to follow a Peer Review with a Discussion step where learners can reflect on their assignment and review and what they have learned from the process.

A Quiz step is a structured conversation with the educator, where each response to a correct or incorrect answer offers feedback and/or a link back to a previous step to recap knowledge.

## A typical course structure

Course structures will vary to suit the course topic, aims, and resources. The structures below are offered as exemplars to be adapted.

Here is a typical Week 1 course structure (from an invented course on Teaching with Technology').

**Week 1: Teaching with Technology - Introduction**

| | **Topic** | **Description** | **Learning** | **Step** |
|---|---|---|---|---|
| 1.1 | How can we teach better? | The big question or issue that the course is addressing, by the lead educator | Delivered | Video |
| 1.2 | Why are you here? | Introduction to the course team, learners, and ways of learning | Conversational | Discussion |
| 1.3 | Ways of teaching and learning | Introductory course material | Delivered | Video |
| 1.4- 1.7 | Teaching methods | Further core course material | Delivered | Video, Audio, Article |
| 1.8 | How I teach | Learners reflect on the course material and relate it to their personal teaching experience | Reflective | Discussion |
| 1.9 - 1.12 | Technologies for teaching | Further core course material | Delivered | Video |
| 1.13 - 1.14 | Case study: Essa Academy | Ground the content in a case study | Case-based | Video, article |
| 1.15 | What makes a good teaching environment? | Learners address the question by reflecting on the core material in relation to their experience | Reflective | Discussion |

These are just examples, as starting points and templates for designing your course. An Exercise step and practical activity will vary with the course. The Exercise might be to examine a painting, solve a puzzle. The practical activity could involve creating and sharing images, videos, computer programs, or short stories, or it could involve using a piece of simulation software or a mobile app to collect data from the environment.