The nature of peer-feedback in a MOOC: A Case Study

Denisse López Benavides
Doctoral Programme in E-Research and Technology Enhanced Learning

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Department of Educational Research,

Lancaster University, UK

The nature of peer-feedback in a MOOC: A Case Study

Denisse López Benavides

This thesis results entirely from my own work and has not been offered previously for any other degree or diploma.

The word count for this thesis excluding references is of 46,645 words.

Signature DENISSE LÓPEZ B.

Denisse López Benavides

The nature of peer-feedback in a MOOC: A Case Study

Doctor of Philosophy, March 2018

Abstract

The present research study aimed to describe the nature of peer-feedback in a Massive Open Online Course (MOOC) offered by a German University. The MOOC was team-based, it offered the support of different actors, and allowed learners to participate in two different roles: students or supporters. Considering participants' diversity and voluntary participation in the MOOC, three aspects of feedback were explored: 1) the quantity and quality of feedback provided throughout the course; 2) the feedback provided by students and supporters; and 3) the feedback provided by females and males in their roles as students and supporters. For this purpose, the method of content analysis and a regression model with a Poisson distribution were employed.

Findings from the research study indicate a positive trend in the quantity of peer-feedback provided throughout the course, a constant use of those types of feedback expected to support the formative function of assessment, and an alternate use in other types of feedback identified as quality feedback. A statistically significant difference could be established between both roles. Supporters used a set of types of feedback more often than students. Finally, although no statistically significant differences could be established between both genders, differences could be established in relation to the types of feedback used in the role they had performed.

The study concludes that the purpose and value of assessment were not threatened by the peer-feedback offered by its participants, despite their diversity and voluntary participation. Additionally, it recommends to present assessment criteria principally as a suggestion. Lastly, it encourages researchers on the field to inquire into a) mechanisms that motivate learners to engage in the voluntary activity of peer-assessment, thereby contributing to a sustainable participation throughout a complete course; and b) understanding participants' commonalities, needs and expectations by analysing the feedback they provide.

Acknowledgements

The completion of this piece of work was possible thanks to the support and encouragement of different individuals. Firstly, I would like to acknowledge the professional guidance and always valuable feedback provided by my supervisors, Dr. Kyungmee Lee and Dr. Don Passey (TEL Programme Director). Also, I would like to thank Christian Friedrich and his team at the Leuphana Digital School for supporting me with all necessary information and data that made this study possible.

I am particularly grateful to my husband Lars and my son Finn, who have supported me unconditionally throughout my years of study. I am very fortunate to have them in my life.

I would like to express gratitude to my parents, who have always been there for me and who have followed with interest this journey, to my brothers and my parents-in-law. I am especially grateful to my brother Mario, who supported and encouraged my research and thanks to whom I feel I have grown as a research scientist. Special thanks goes to Dr. Ibis Alvarez, who has been an example and inspiration for me and for my research. She has motivated me to initiate my academic journey and to keep on striving towards my goals. Finally, I would like to thank my fellow doctoral students from Cohort 7 and friends for supporting me along the way with their best wishes.

Thanks for all your encouragement! It is true that words cannot express how grateful I am for having had you around during my studies.

Contents

List of Figures

List of Appendix Figures

List of Tables

List of Appendix Tables

Chapter 1 Introduction

Traditionally the educational process consists of a cycle of three steps: teach, learn, and assess (Frederiksen & Collins, 1989), the third step being a crucial one for further guiding instruction and verifying learning (Costello & Crane, 2013; Hattie & Timperley, 2007; Suen, 2014; Topping, 1998). However, whilst the first two steps appear to be easily transferable across learning settings and to remain less affected, the third one can be seriously affected. An example of this may be seen in Massive Open Online Courses (MOOCs). The scale of these courses makes difficult the possibility of providing one-on-one feedback and support. In other words, assessment has become a concern that instructional designers have addressed by introducing manifold assessment strategies of qualitative and quantitative nature. Peer-assessment/feedback is an example of the former, where learners from a same course are asked to assess each other's work. In fact, peer-feedback has become one of the essential features of many MOOCs and some authors argue that this trend will continue in order to compensate the massive aspect of the courses (O'Toole, 2013; Pilli & Admiraal, 2017).

However, concerns about using peer-assessment/feedback continues to increase with the realisation that learners are diverse, and that their presence in the courses is voluntary. These two characteristics can put at risk the purpose and value of this very last step of the educational process cycle. Accordingly, Suen (2014) argues that 'flexibility' is a keyword when incorporating peer-assessment strategies in these kinds of learning settings. He suggests that peer-assessment needs to be:

1) simple and easy to understand for students;

2) efficient in execution, without occupying much time;

3) limited in that each student rater [assessor] is asked to rate [assess] no more than a handful of other students' assignments. (p. 4)

Nevertheless, considering participants' diverse characteristics in a MOOC, the very first requirement appears difficult to meet. According to Popov et al. (2012), and Popov, Biemans, Kuznetsov and Mulder (2014), this diversity presumably implies different attitudes, behaviours and expectations towards the activities that are proposed in any course.

Up to now, the main focus of studies regarding assessment on MOOCs has been on quantitative approaches that aim at scaling peer-assessment (e.g. Balfour, 2013; Kulkarni, Bernstein, & Klemmer, 2015; Kulkarni et al., 2013; Piech, Chen, Koller, & Ng, 2013; Suen, 2014). Notwithstanding, studies with qualitative approaches are also found. These have focused on participants' experiences and perceptions as participants in their respective MOOCs but there has been a lack of studies that examined peer-feedback or assessment (e.g. Aharony & Bar-Ilan, 2016; Cho & Byun, 2017; Veletsianos, Collier, & Schneider, 2015; Zheng, Rosson, Shih, & Carroll, 2015).

As pointed out by various authors, many variables underpinning the intricacies of peer-assessment (e.g. gender, culture, learning styles, reception and use of feedback) have not been rigorously and independently researched (Evans & Waring, 2011; Shute, 2008; Topping, 2010; van Zundert, Sluijsmans, & van Merriënboer, 2010). This is a pressing issue that applies fully to the MOOC context and to which this study is aiming to contribute.

1.1 Problem statement

Despite the aforementioned concerns when employing peer-assessment/feedback strategies in MOOCs, little is known about what peer-feedback actually looks like in such courses. As Luo, Robinson and Park (2014) note, empirical approaches exploring peer-feedback in a MOOC in its different dimensions have yet to be undertaken. Thus, a research study investigating those characteristics of concern together with other variables, and a method for exploring peer-feedback's actual content has never been more significant and necessary.

Currently, research studies on MOOCs appear to be limited to specific data collection and analysis methods, naming the increased use of quantitative, clickstream and observational data as an example of this (Veletsianos et al., 2015; Veletsianos & Shepherdson, 2016). However, as these authors point out qualitative approaches and a combination of methods are still required to understand learning and participation in open online learning environments. By doing so, findings can be triangulated, methods exploited, and phenomena analysed from different perspectives taking into account the context surrounding it. An over-reliance on computational methodologies may be misleading for those in charge of the techno-pedagogical design of the online learning offers. Findings may not provide a comprehensive understanding on MOOC participants who create and assign meaning to their experiences and participation in activities (Veletsianos, 2013). Practices around MOOCs are still developing, and interpretive approaches –as the of the present study– are useful for researchers seeking to expand their understanding of the meaning of MOOCs, its possibilities and limitations.

Due to the quick developments of MOOCs, researchers have searched for alternative and faster publication outlets than traditional academic journals. These outlets include conference proceedings and journals that issue special editions on MOOCs. It was there, in one of those outlets, that the study by Krogstie, Horgen and Hjeltnes (2015) was found. That study inquired into peer-feedback in a MOOC with the purpose of measuring peer-feedback quality in a qualitative way. Yet, the characteristics of the environment and its participants were of no actual concern for using peer-feedback as an assessment strategy, as previously discussed.

1.2 Purpose of the study

The purpose of this study is to describe the nature of peer-feedback in a MOOC. To this aim, empirical evidence was gathered to explore: 1) the quantity and quality of peer-feedback provided throughout the course; 2) the feedback provided by participants in their respective roles (as students or supporters); and 3) the feedback provided by female and male participants in their respective roles.

1.3 Research questions

Three research questions were formulated to address the purpose of the study. These resulted after having recognised important gaps in the MOOC literature and from having access to field data of a specific course.

1. How does feedback evolve over time in terms of quantity and quality?

2. How does the nature of feedback differ between students and supporters?

3. How does the nature of feedback differ between female and male participants in their respective roles as students or supporters?

This research study employs a couple of terms that are presented below to better understand the context of the research questions.

1. **Assessment**: "the process of evidencing and evaluating the extent to which a candidate has met or made progress towards the assessment criteria" (Busuttil-Reynaud & Winkley, 2006, p. 12), whereby the emphasis is placed on the product and outcome rather than the process (Topping & Ehly, 1998).

2. **Feedback**: "qualitative information about their [learners'] performance given to students after an assessment" (Busuttil-Reynaud & Winkley, 2006, p. 51).

3. **Participants' roles**: The MOOC enabled two ways of participation.

- Students[1]**:** aimed at obtaining a university certificate that represents 5 ECTS (European Credit Transfer System). Within their role, a set of requirements needed to be fulfilled. These are explained in sub-section 1.5.

- Supporters: were interested in the course and its progress, and could freely determine their degree of participation in it. Their

[1] The original term was 'participant', but it was changed to 'student' for a better distinction.

principal role was to support students and their groups to become better throughout the course.

1.4 Contributions of this study

This study presents a type of MOOC that emphasized on peer-to-peer and mentored project-based learning by using a case-study method as the over-arching type of task. MOOCs of this type are scarcely found in the literature, thus of interest for practitioners who seek alternative perspectives of course design that can help enrich the current MOOC didactical landscape. Besides a rich variety of didactical elements that are included in the design of the analysed course to support teaching and learning, a great value has been placed on assessment. Scaffolding elements that make up for the scale of the courses have been included and peer-feedback is an example of this. Peer-feedback as an assessment and learning strategy –although not new– has become commonly employed in manifold learning contexts. However, its use has been thought to fulfil logistical functions over formative ones, and this can be seen in the pedagogical design of the courses. The increasing number of MOOCs and the greater demand for inclusion of peer-feedback in their pedagogical design of MOOCs, justify the need of exploring its content by taking into account participants' characteristics. Different groups of persons can benefit from this study:

Educational entities considering the creation or implementation of a MOOC can become aware of the intricacies of peer-feedback as reported in this study, and so be able to better define the characteristics of a prospective course, and to be considerate towards the audience to which the course will be offered. On the other hand, for educational entities that are already running MOOCs and using peer-

feedback in their pedagogical concepts, the results of this study will be valuable for comparing, reflecting on and evaluating the pedagogical design of their courses.

Instructional designers who are conceiving a MOOC for the first time can greatly benefit from the present study. By acknowledging the many variables involved in the provision of peer-feedback, they can carefully focus on the elements they consider that can better support the provision of feedback and plan specific moments in the learning phases when these can be introduced. Instructional designers with experience in the conception, design and implementation of MOOCs can benefit from this study in mainly two ways. First, by employing the derived instrument of this study. They can analyse, evaluate and reflect on the nature of feedback in their courses, and so be able to review and adapt aspects of the course's pedagogical design (e.g. assessment criteria, learning tasks). Second, by considering the practical recommendations herein presented they can create learning experiences that can be considered more meaningful to the course participants and that can motivate them to engage in the proposed activities.

Researchers of higher education and online education will benefit from the empirical data provided in this study, the derived instrument, the practical recommendations, and the ideas for further research. Each of the aforementioned areas have great potential that can help researchers to expand their knowledge and understanding of peer-feedback and its implications when employed in a community where diversity is its main characteristic.

New researchers can benefit from the different chapters of this study by using them as a guide for their own research topics.

MOOC participants who show an interest in the dynamics of such a course, can gain an understanding of the development of peer-feedback in a MOOC. Moreover, findings can help them self-evaluate their readiness for or commitment to undertaking such an activity in the future. Additionally, they can identify the types of feedback that they employ or that can be employed to support the formative purpose of feedback.

Participants who are new to MOOCs and who are confronted with a course in which peer-feedback is proposed or demanded can understand the purpose of this learning strategy as well as the requirements that it poses. This way they can better prepare for their role and input (if they can choose).

Finally, besides contributing to the previous groups, the group of developers of the tool for qualitative analysis that was employed for this study benefited as well. Based on the particular way of coding undertaken by the researcher during content analysis, developers had a case from which they could further enhance the product.

The present investigation studies a specific single case, and the above uses may be limited by this fact. Beyond the methodological choices and practical limitations, the selection for the exploration of this particular course was based on the research topic and the attributes of the course, whereby peer-feedback is an essential element in its pedagogical design. Despite the limitations that the contextual characteristics of MOOCs may bring (i.e. participants' diversity and voluntary

participation), in this course an effort has been made to integrate all three steps of the educational process' cycle: teach, learn and assess, and make sure each step is addressed properly.

The following section provides detailed information about the studied case that is crucial for making sense of the different chapters in this document.

1.5 Contextual background – "Managing the Arts: Marketing for Cultural Organizations"

Leuphana Digital School (LDS) is a sub-division of the Leuphana University of Lüneburg in Germany that has been in charge of developing online education formats for global learners since its establishment in 2012. LDS' offerings include Massive Open Online Courses (MOOCs) as a way of reaching out to diverse cultures; creating a name; experimenting with innovative concepts in online teaching and lastly, building a profile and collecting experience for future offerings, i.e. programmes. With the idea of further developing MOOCs, LDS created the concept of Mentored Open Online Courses. Just like most of the MOOCs, this type aims at contributing to "open up education and provide free access to university level education for as many students as possible" (Yuan & Powell, 2013, p.6). The design of LDS MOOCs draw on sociocultural approaches to learning (Eun, 2010). They are built on the concepts of peer-learning and problem-based learning, where cooperation in small groups up to a maximum of five persons is required, so as to allow for scaffolding by different supporting actors. One of these actors is the mentor, from which the courses receive the name. This actor provides guidance, support and encouragement to participants in all content-related matters. See Appendix One for more details on all supportive actors.

The courses run in English and consist of consecutive learning phases and assignments, which grow in complexity over time. Each learning phase is two-weeks long and concludes with the submission of an assignment. In order to ensure a more personalised support to the learners, the courses are designed for a smaller number

of participants. However, as a means to meet the expectations of the 'massive' aspect from a MOOC, the courses allow two ways of participation: as a student and as a supporter.

Whilst students are required to participate actively in the course by completing all assignments successfully in a group, supporters do not need to fulfil any such requirement.

However, both types of participants enrol in the course for learning's sake (Veletsianos, 2013), and so are willing to help create and sustain a learning community where support is offered to each other and where collaboration is an essential element (Lewis & Allan, 2004). This commitment is based on the learning benefits that they find in the different possibilities of interaction present in the environment (Wenger, McDermott, & Snyder, 2002). Both students and supporters have access to all the learning materials, assignment descriptions, discussion forums, and the assignments submitted by the groups in each of the learning phases. They also may establish contact with anyone in the learning community at any time.

During the learning phases students in their groups are encouraged to share drafts of their assignments with the learning community so as to benefit from their comments before handing in their final submissions.

1.5.1 The assessment process in LDS' MOOCs

Once a learning phase has been completed, students in their groups are required to submit the final version of their solution for the assignment by publishing it on the learning platform. These products can be accessed by all members of the learning

community and are ready to undergo the process of assessment. The products are assessed in two forms: formally and informally.

Whilst mentors are responsible for the first form of assessment and only their evaluation counts towards the final grade of the students, the second form of assessment is voluntary and can be undertaken by anyone participating in the course. It is the latter form of assessment, namely the one of peers, which is the focus of the present study and that will be examined.

In a learning community participants progressively create a common sense of responsibility, a safe environment, and they reach a certain degree of dependence (Charalambos, Michalinos, & Chamberlain, 2004). However, this does not guarantee that participants will assume more responsibilities than those to which they signed up for. In other words, participants must not assess the work of their peers and provide them with feedback if not wished. For this reason, mechanisms that motivate participants to get involved in the process of voluntary assessment, as well as benefiting from being involved in it, have been initiated. LDS' courses have introduced two types of rewards: badges of two types and a Statement of Accomplishment. These are explained at the end of this section.

Unlike other MOOCs that rely only on peer-feedback mechanisms as a way of assessment, LDS' courses have mentors that provide a more personalised support to learners. Mentors are assigned to specific groups, for which they are responsible throughout the course. All other participants receive suggestions, from the learning platform, about three specific groups to which they could provide feedback. The

algorithm created for the latter ensures that there is an equal number of reviewers for each group. Apart from the suggestions received, everyone is free to assess any group of their preference and as many as they wish.

A single template generated by the learning platform is used for both forms of assessment. (see Figure 1.1).

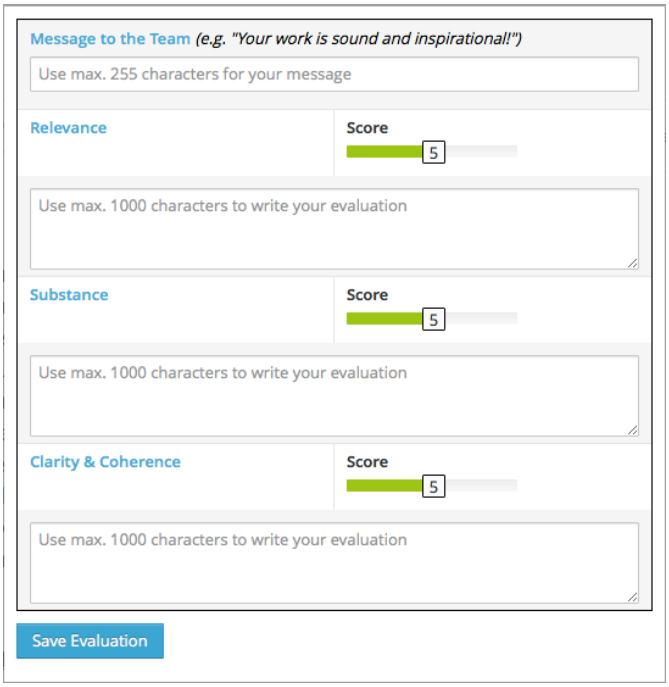


Figure 1.1 Assessment template for "Managing the Arts: Marketing for Cultural Organisations".

Three evaluation criteria were used in the course: *Relevance*, *Substance*, and *Clarity and coherence*. Whilst mentors used further criteria and specifically designed rubrics for each learning phase, all other voluntary assessors were guided through questions under each criterion (see Table 1.1). Criteria were presented to the participants through different communication channels: within the learning platform (e.g. as a formal announcement with a link to a static page) and outside of it (i.e. newsletters). However, these were not discussed with the participants. Both types of assessors

(mentors and participants) were allowed a maximum of 1000 characters for writing their feedback for each criterion. Moreover, they could add a title to their feedback using up to 255 characters. Finally, they could score the work of the groups using a ruler with values between 1 and 10.

Table 1.1 Guiding criteria for evaluating peer-work: MOOC *"*Managing the Arts".

Relevance	• Did the team properly address the tasks at hand? • Are all claims backed with sound argument? Does the submission present valid, well-grounded insights? • Are the conclusions drawn grounded in data and/or theory? Are subjective opinions clearly discernible as such? • Are positions/statements that conflict with the team's position/thesis considered and properly addressed?
Substance	• Did the team allow for frameworks that foster the co-existence of multiple sources, materials and claims? • How deeply did the team penetrate the material to answer questions and solve problems? • Are arguments built upon valid deductions, salient examples and/or theoretical frameworks? • Is reference made to academic or other appropriate sources or positions in the broader discourse (also: references to discussions within the online platform)?

Clarity & coherence	• Does the submission show internal coherence? • Is the argument structured in a clear and concise manner? • Can an outside reader understand the submission, being unfamiliar with the team and its communication habits? • Are there any elements in the problem's context left out without adequate justification? • Is the process of creating a coherent answer made visible?

Formal and informal assessments –namely the feedback provided as well as the authors' name (as used in their profile)– are made visible to everyone in the learning community and remain so until the completion of the course.

All participants involved in the assessment process may be rewarded with badges and a Statement of Accomplishment, representing symbolic recognitions for the effort and time invested during the course. When assessing the works, two types of badges acknowledging the quantity and quality of the feedback provided are awarded. The former type of badge is called Active Evaluator and is awarded after assessing three works. The latter type of badge is called Popular Evaluator and is awarded if the feedback provided has been perceived as useful by those reading it. For this, a star system is used. Feedback messages which are rated five times with

four or five stars are awarded with this type of badge. The Statement of Accomplishment may be earned only after having obtained six Active Evaluator badges and three Popular Evaluator badges.

Having introduced the contextual situation of this study, a review of the literature is presented next.

Chapter 2 Literature Review

The essential expectation underpinning this study is that any MOOC participant who has worked on solving a similar task under the same conditions as other participants is thought to be capable of offering feedback to her or his peers, regardless of their gender. However, sociocultural factors and personal skilfulness may play a key role regarding the nature of feedback provided and more importantly, the way feedback is perceived by the students receiving it. Based on this, this chapter is divided into seven sections. The first section aims at presenting MOOC features and the implications these bring to its participants. The second section discusses MOOCs' cultural diversity from a sociocultural lens, highlighting the learning possibilities and limitations that are entailed by this diversity. In the following section peer-feedback as an assessment strategy is presented underlining its essential elements and functions. Subsequently, the conditions for fulfilling the formative function of feedback are described. The fifth section focuses on feedback content and how it has been reported in the literature. The sixth section sets the stage for the present investigation that focuses on the exploration of the content of peer-feedback in a MOOC. Finally, the last section presents and justifies the hypotheses used for this study.

2.1 An Introduction to MOOCs

MOOCs are a relatively new phenomenon that arose from the ideals of openness in education, in which the desire to learn should not be constrained by demographic, geographic or economic factors. The development of the Internet and mobile Internet have addressed and facilitated this desire to many (Yuan & Powell, 2013).

MOOCs have attracted the interest of hundreds of thousands of people around the world. However, they have been apparently benefiting mostly those who already enjoy or have enjoyed higher education (Christensen et al., 2013; Dillahunt, Wang, & Teasley, 2014; Emanuel, 2013; Liyanagunawardena, 2015; Liyanagunawardena, Williams, & Adams, 2013).

With the aim of creating a definition for MOOCs the European Commission –in one of their reports– listed a set of features underlying these courses. These features provide an indication of the implications of MOOCs (European Commission, 2014, p. 2):

- 'Massive' in the sense of no limit to attendance
- 'Open' in the sense that it can be accessed by anyone anywhere as long as they have an internet connection
- Free of charge (not including the certificate of completion)
- Online
- Structured around a set of learning goals in a defined area of study
- Executed according to a specific time frame and completion point
- Offers possibilities for interaction, such as social media channels, forums, blogs or RSS feeds that contribute to building a learning community
- Provides course materials such as videos, readings, and others, are provided free of charge by the course designers, teachers, facilitators or even students themselves (cMOOCs)
- Includes assessment or evaluation, self-assessment or peer-assessment

These features indirectly set requirements for potential participants, which may range from technological to meta-cognitive aspects. Some of those features are discussed below.

In a current review on MOOCs Liyanagunawardena (2015) argues that these courses have indeed increased access to higher education, but have not broadened access. She suggests that those currently benefiting from these courses have better 'access' to MOOCs, have acquired adequate self-directed learning skills, wish to update their skills and continue their professional development, or are leisure learners. Warschauer (2003) suggests that 'access' does not only imply having a technological device or a connection to the Internet, but that it comprises manifold factors such as "physical, digital, human and social resources, and relationships" (p. 6). Studies carried out by Fini (2009) and Kop, Fournier and Mak (2011) on MOOCs go in line with Warschauer's claim and bring up the following additional factors to be considered under 'access': time available, command of the language in which the courses are delivered, ICT skills, confidence level at the course start, trust and comfort levels, power relations in the course, the use of tools and language, among others. It is evident that when potential participants are affected by any or a combination of these factors, their participation in the courses may most likely be a more passive one (Fini, 2009; Kop et al., 2011). However, assuming a passive role may also be a result of deficiencies in self-directed learning skills. Without entering into details, as learner autonomy will be explored in sub-section 2.4.1, it is worth noting that various factors have an impact on the way in which learner autonomy is developed. According to Bouchard (2009), among these factors the ones related to

environmental issues need to be considered. This author claims that the value that learners place on learning, on the investment they feel to be making for their own future –be it for refreshing skills; extending knowledge on areas that are already familiar or for entering new areas– directly reflects learners' degree of autonomy.

This said, although the courses are open, free, online and most of them may not require previous knowledge on the topics treated, they indeed require an additional set of skills and knowledge that cannot be assumed from a wide population.

According to some studies (Daniel, 2012; Rodriguez, 2012; Yuan & Powell, 2013), the concept of MOOCs has been driven by two diverging pedagogical directions: xMOOCs and cMOOCs. Whilst the former follows behaviourist approaches to learning with more individualistic learning, the latter place an emphasis on connected, collaborative learning and has connectivism as the underlying learning theory. However, as courses rapidly developed, it became obvious that a binary classification was not useful nor characteristic, and that MOOC pedagogy needed to be seen at the "micro level of individual course design" (Bayne & Ross, 2014, p. 8). This observation had been already noted by Clark (2013) and later by Sanchez Gordon and Luján Mora (2014). The former acknowledged the differences in the existing MOOCs and attempted to offer a categorisation of them into eight –not-mutually exclusive– types. The MOOC serving as the case study for this investigation has characteristics of four of the following types:

- synchMOOCs, have a fixed start and end dates, as well as assignment deadlines;

- madeMOOCs, are innovative in the use of video, are quality driven in the design of material and the design of challenging learning assignments, harness peer-work and uses peer-assessment techniques;
- groupMOOCs, focus on collaboration in small groups and have mentors;
- connectivistMOOCs, rely on the connections across a network of peers

This study only focuses on one aspect of those features underlying MOOCs, namely: peer-assessment. The following sections help to develop an understanding of peer-assessment in the context of MOOCs from different perspectives.

2.2 Learning with and from MOOCs' diversity

Despite the various requirements that MOOCs posit on potential participants, these courses still take place and their main characteristic is participants' cultural diversity (Suen, 2014). Besides differences in terms of age, gender, level of education and experience, the underpinning differences lie on language and culture. Popov et al. (2012) note that diversity in its different degrees is a reality that creates benefits but that also poses challenges for all actors involved in the learning process.

Whilst courses in other learning settings have a certain degree of homogeneity (e.g. assume learners to have a similar level of education and knowledge on specific topics), MOOCs welcome anyone that is interested in the topic offered (Liu et al., 2014), thus increasing this way its heterogeneity. Some authors referring particularly to online environments, including MOOCs, claim that the probability of miscommunication increases as the cultural gap between learners in a learning community becomes greater (Bartholet, 2013; Leber, 2013; Liyanagunawardena et

al., 2013; Mak, Williams, & Mackness, 2010; Reeder, Macfadyen, Roche, & Chase, 2004). Miscommunication is problematic as it can modify the function, meaning and interpretation of information that is exchanged among participants in a course, such as feedback. As Popov et al. (2014) note, miscommunication may emerge from different attitudes, styles of communication and behaviour patterns for acting and interacting, which are distinct in each culture. In fact, individuals are a result of the multiple realities of their cultures and according to Alfred (2002) this is what is brought into the classroom.

One of these realities regards the cultural expectations that are connected to sex categories (e.g. female, male) (Risman, 2004; West & Zimmermann, 1987) and hence gender. Gender is here understood as a "social system that restricts and encourages patterned behavior" (Risman & Davis, 2012, p.8). Thus, gender – assumed as a structure – discriminates possibilities and limitations in relation to a certain sex category (Risman, 2004).

In the context of education and specially in online learning, gender-based differences have attracted the interest of various researchers (e.g. Yukselturk & Bulut, 2009). Differences have been reported to be found in participation, contribution and interaction styles of female and male participants (Gunn & McSporran, 2003; Price, 2006; Yukselturk & Bulut, 2009), their language styles (Herring, 1993), and voices (Blum, 1999; Rovai, 2001), to name a few. However, the studies had characteristics in common that differed from the of the present study and that may give an explanation to the differences found. Beyond the facts that students were familiar with each other, the format of the courses was blended and combined face-to-face

and online phases, the sizes of the classes were rather small, and participation was compulsory, the principal characteristic was that there were minimal cultural differences between the participants. If cultural expectations were said to be connected to both sex categories –as commented on previously– then it appears understandable that differences would emerge for each sex category. Some of the studies noted that access to and confidence with technology may have played an important role in the findings. Nevertheless, the contextual characteristics of a learning setting such as a MOOC can challenge findings from other traditional learning settings, and gender differences may be one of them.

Apparently, among the diversity of topics that have been studied regarding feedback in online learning environments, studies combining feedback and gender seem scarce. Notwithstanding, as an example, it is worth mentioning a study carried out by Read, Francis and Robson (2005), where gender in written feedback was explored in different disciplines, which reported similarities between genders in the way they presented feedback.

Vygotsky's sociocultural theory assumes that beliefs and attitudes are instilled from a sociocultural context, and that these shape the way in which learning and development may happen (Vygotsky, 1978). This study builds on the assumption that all individuals have the same capabilities and recognises that the sociocultural contexts in which they are and have been involved are directly associated with their beliefs, attitudes, expectations and actions towards gendered behaviours.

Vygotsky stresses that learning and development are processes that require intentional involvement of a person in a larger sociocultural context, and that these are supported by dialogue and social mediation, with language as its backbone (Alfred, 2002; Vygotsky, 1978; Wertsch, 1991). At the same time, the theory identifies the different experiences and perspectives of individual persons as opportunities to enrich and benefit the learning experience of a community (Alfred, 2002; Leber, 2013; López-Benavides, 2014; Nasir, Rosebery, Warren, & Lee, 2005; Niewolny & Wilson, 2006).

Some authors (Gutierrez & Rogoff, 2003; Kulkarni, 2014; Kulkarni et al., 2016; Nasir et al., 2005; Nkuyubwatsi, 2014) suggest that a way of addressing cultural diversity is by designing meaningful and flexible learning experiences where learners can relate to, adapt and transfer to their own cultural settings and realities while still being exposed to diversity. Nevertheless, language –as indicated previously– still continues to be a latent challenge, especially in the MOOC context.

LDS' MOOCs acknowledge participants' cultural diversity and attempt to enable significant learning experiences in which opportunities for expression, reflection, interaction and collaboration are made available through various moments, spaces and tools. For instance, peer-assessment is one of those moments and spaces. Additionally, a context and a structure are provided in order to support a learner's development, the emergence of new ideas, and enhancement of creativity. Nonetheless, language –as the principal mediating tool for initiating any of the aforementioned processes– poses limitations. The information that is exchanged during the processes of feedback and assessment, and the way in which this

information is interpreted and understood by each course participant is decisive in the learning process. The present study focuses on two specific characteristics of the participants: the role they performed in the course (student or supporter) and their gender.

In the following section, the importance of the information delivered in a message containing feedback is discussed.

2.3 Peer-feedback as an assessment strategy

Acknowledging that social interaction can shape learners' experiences and that learners' perspectives and experiences are enriching and beneficial opportunities for learning in community, then feedback as an assessment strategy appears to be an ideal option to support this notion. This strategy can –under different conditions and settings– address what Vygotsky (1978) introduced as a Zone of Proximal Development (ZPD), which is:

> the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers. (p. 86)

The diversity of learners in a MOOC community presumably implies the presence of more knowledgeable others. However, according to some authors, the exclusive use of peer-assessment techniques in courses of any kind may jeopardise the formative function of assessment (van der Pol, van den Berg, Admiraal, & Simons, 2008), which is that of enhancing learners' knowledge, skills and comprehension of a topic studied

(Shute, 2008). Its sole use may consequently hinder assessed-learners from reaching ZPD.

Because of the importance of assessment, LDS' MOOCs have introduced different supportive actors in its courses. The role of the mentor, for instance, has been especially conceived to support students reach ZPD. This, however, does not undervalue peers or their feedback. Having mentors reassures students in that they can rely on continuous and rich feedback despite the varied types of learning tasks they are expected to accomplish; this reassurance would not otherwise be possible when peer-assessment is voluntary. Rich feedback is interpreted here as useful and valuable information that learners understand and can use to enhance their products.

The desired learning evidence in LDS' MOOCs is reflected in participants' capacity to find solutions to problems, come up with ideas or create products, for which high order thinking skills (e.g. Krathwohl, 2012) are required. Primarily, the type of feedback that is expected in LDS' MOOCs is *task feedback* (Hattie & Timperley, 2007), which focuses on different aspects of the learning task, and which –according to DeNisi and Kluger (2000)– proved to show more effectiveness than in other types of feedback. The overarching type of task that is the focus of assessment is the case study (Cohen, Manion, & Morrisson, 2000; Falchikov, 2005). Herein, problems are raised from real cases or situations and solutions are expected to be provided by means of different products such as: analyses, reports, design of processes, plans and finally proposals. The case study is addressed through different sub-tasks that include theory and practice. The former allows students to develop conceptual

understandings that will enable them to offer well-grounded arguments for the latter. The types of tasks can play a significant role when considering undertaking peer-assessment. In fact, Neubaum, Wichmann, Eimler and Krämer (2014) have reported that the willingness to engage in the process of assessment and the degree of elaboration of the feedback in MOOCs is influenced by the task, the length of the product to be assessed and its content. Whilst those tasks involving theoretical aspects may be perceived as abstract to many assessors, the practical tasks in which real or fictional situations are presented may facilitate the transmission and understanding of ideas, making it easier to digest and more appealing to others.

2.3.1 Functions and foundational elements of feedback

Whether the tasks to be assessed are appealing or not, feedback must be provided and it should be rich. Black and Wiliam (1998) highlight the importance of delivering rich feedback in order to support learning so that feedback is formative.

According to those authors, feedback has two functions: 1) to identify the weaknesses and to point out what needs to be revised (*directive*), and 2) to offer comments and suggestions that learners can use as a guidance for their own revision and understanding (*facilitative*). The first one evidently reflects the view of assessment as summative and corrective, whereas the second one views it as formative. Regarding formative feedback (or assessment-for-learning), Shute (2008) defines it as "information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of learning" (p. 154).

As claimed by Kulhavy and Stock (1989) and Narciss (2006, 2008) the types of information that learners are provided with formative feedback include *verification* and *elaboration*. Whilst the former provides a quick indication on whether the created product or parts of it have met specified requirements, the latter provides relevant information that guide the learner in the process of meeting those requirements. *Elaboration* is a type of information that can be combined in different ways, and can include verification. This combination can meet the *directive* and *facilitative* purposes of feedback highlighted by Black and Wiliam (1998). Nevertheless, it is important to consider that the topic to be learned and the learning goals, determine the types of information that are required in the feedback.

Besides acknowledging the functions of feedback, assessors undertaking assessment, or here specifically peer-assessment, are also required to understand the criteria that guides the assessment and to engage with the product in terms of analysis and comprehension. The former appears to be the foundation for being able to undertake assessment and thus support learners' further development.

Having in mind the characteristics of MOOCs and its prticipants, these two conditions may become problematic. As previously mentioned, peer-assessment is a voluntary and non-monitored activity, and as such it may not be taken seriously (Krause, 2013; Suen, 2014) or may even be ignored. Moreover, the quantity and quality of feedback provided may be affected if in those environments the assessment criteria are difficult to understand, are insufficiently described, or are not discussed prior to using (Krause, 2013). Here, once again, the principal tool

enabling the process of assessment is language, which is used for reading, understanding, interpreting and finally writing.

2.3.2 Perceptions regarding the value of peer-feedback

Feedback is said to fulfil logistic, pedagogic, cognitive, metacognitive and motivational functions (Admiraal, Huisman, & van De Ven, 2014; Butler & Winne, 1995; Narciss & Huth, 2004; Nicol & Macfarlane-Dick, 2006; Topping, 1998; Tseng & Tsai, 2007), yet the value of implementing peer-feedback as an assessment method is debatable. Whilst some authors doubt the predictability of peer-feedback (Chen, Wei, Wu, & Uden, 2009), its efficiency (Boud, 2000) or its validity (Strijbos & Sluijsmans, 2010), there are advocates of this method that have presented positive evidence. Different methods of qualitative and quantitative nature have been employed to measure positive evidence, which in most of the cases report on participants' perceptions. Nevertheless, Evans (2013) who carried out a literature review on assessment in higher education, reported that those studies addressing peer-feedback were rather small scale, opportunistic and regarding a single topic, and that these conditions put into question the validity and reliability of the findings of the studies. This observation is shared by Hounsell, McCune, Hounsell and Litjens (2008) and Poulos and Mahony (2008).

Studies carried out by Sluijsmans, Brand-Gruwel and van Merriënboer (2002) and Falchikov (2005), as well as studies reviewed by the latter author in which peer-feedback was undertaken, report positive acceptance towards this method. The studies had similarities in the sense that the learning settings were face-to-face, the groups were relatively homogeneous (e.g. same level of education, similar ages,

minimal cultural differences), the sizes of the classes were manageable, and training in peer-feedback was provided. However, there are also studies which - under similar conditions as the previously mentioned ones - report a rather negative acceptance, rating feedback as ambiguous and irrelevant (e.g. Tsai, Lin, & Yuan, 2002). Van Zundert et al. (2010) notify that the amount of published studies regarding positive perceptions or attitudes towards peer-feedback are considerably higher than those with contrary findings. However, in the context of online learning, this appears to be the opposite.

Discrepancies regarding students' perceptions exist. Some examples of studies report negative findings and categorise the peer-feedback received as questionable, unattractive, unfair and inaccurate (Cheng & Warren, 1997; Davies, 2006; Kaufman & Schunn, 2011; Liu & Carless, 2006; Rushton, Ramsey, & Rada, 1993) also in MOOCs (e.g. Suen, 2014), while some others affirm the contrary (e.g. Prins, Sluijsmans, Kirschner, & Strijbos, 2005) also in MOOCs (Neubaum et al., 2014). In summary, peer-feedback may prove positive for many, but not for all (Fund, 2010), and the learning setting and other variables may play a role in this perception.

2.4 Providing rich feedback

It is evident that providing rich qualitative information to peers is challenging and that academic skilfulness seems to play an important role (van Zundert et al., 2010). In the MOOC context a disparity of abilities, skills, commitment and engagement with the course within community members is expected. Thus, sharing interests and knowledge from a common area do not seem sufficient to be able to provide rich feedback.

As in many other learning activities, peer-assessment requires training and practice (Black & Wiliam, 1998; Boud, 2000; Evans & Waring, 2011; Falchikov, 2005; Nicol, 2008; Sadler, 1998; Sluijsmans et al., 2002; Topping, 2010; van Zundert et al., 2010; Vickerman, 2009). Topping (2010) claims that the preparation received by the learners will most likely be reflected in the results of peer-feedback use and in their perception towards the method. Most MOOCs are courses structured around a topic with a specific timeframe for completion, and training on peer-assessment is not stipulated. Based on the information provided by the participants when enrolling in LDS' courses, it can be said that the great majority have enjoyed higher education to different degrees. This information may be misleading for those creating and organising MOOCs, as in many cases it is implied that participants bring experience to some extent as receptors, or even better as feedback providers. However, as noted by Scott et al. (2011), it may not be taken into consideration that those participants coming from workplace environments may find it difficult to cope with higher education practices, such as engaging with and undertaking peer-assessment.

In any case, having enjoyed higher education does not imply being familiar with the method of peer-assessment, having been trained on it, or having collected experiences of any kind with it. The complexity of peer-assessment in MOOCs increases even more as learners are not required to bring any knowledge on the course's topic.

Besides time, commitment and engagement, various abilities and skills are required to create and elaborate feedback that fulfils the formative function of assessment

(Costello & Crane, 2013; Nicol, Thomson, & Breslin, 2014; Strijbos, Pat-El, & Narciss, 2010; Tuzi, 2004; van den Berg, Admiraal, & Pilot, 2006).

According to some authors, rich feedback is defined by its *validity, reliability and usefulness* (Falchikov & Goldfinch, 2000; Hattie & Timperley, 2007), and for others by its *content* and /or *style characteristics* (Cho & MacArthur, 2010; Gielen et al., 2010; Kim, 2005; Prins, Sluijsmans, & Kirschner, 2006; Sluijsmans et al., 2002).

In either perspective, it is necessary to understand what it takes to produce and deliver this kind of feedback. Thus, the requirements that need to be met by the persons creating it, and the qualitative information that is delivered, need to be carefully considered. Once the requirements have been outlined, the focus will be placed on the content that feedback requires to be considered rich.

2.4.1 Learners' attributes

Nowadays, education aims to support students to develop into critical, thoughtful, and autonomous learners, and peer-assessment provides a great landscape for addressing this (Falchikov, 2005; Tuzi, 2001). Notwithstanding, these characteristics are acquired with hard work, and the journey to become critical, thoughtful and autonomous will most likely be long and require persistency, perseverance, continuity, guidance and support. Understanding what is meant by being critical, thoughtful and autonomous provides a more concrete idea of the requirements and also the conditions for becoming so.

Although an agreed definition on critical thinking does not appear to exist in the literature, there are different approaches to it. Considering these approaches, Moon (2005) defines critical thinking as:

> a capacity to work with complex ideas whereby a person can make effective provision of evidence to justify a reasonable judgment. The evidence, and therefore the judgment, will pay appropriate attention to context. (p. 7)

This author argues that the capacity of thinking in a critical way essentially depends on learners' understanding of knowledge as relative and not as absolute. It also appears that that understanding is directly influenced by the nature of the discipline learners study. Developing this understanding requires undergoing four different stages, which Baxter Magolda (2007) refers to as: absolute, transitional, independent, and contextual. Progressing through the different stages evidently indicates a shift towards becoming thoughtful and autonomous. This becomes clearer when understanding the characteristics of each stage.

The first phase sees knowledge as absolute or dualist, where only correct or incorrect answers exist. Students in this stage rely completely on the knowledge or information that is transmitted by the teacher. In the second stage students start to discover incongruences and start questioning the possibility of knowledge not being as absolute as they first thought. Teachers are seen here as those helping to understand this possibility. When reaching the third phase students are aware that knowledge is contextual. They start to form own opinions and also give value to those of their peers. Students expect teachers to provide them with a context for

them to explore and try out. In the final stage students recognise that knowledge is relative and that it is constructed. This empowers them to develop own beliefs and opinions, stand up for them and distinguish them from others'. Teachers are no longer seen as the only experts holding knowledge, but as partners with whom knowledge can be co-constructed.

The journey to reaching thoughtfulness and autonomy, which allows students to become critical thinkers, requires students to trust themselves and build contextual knowing.

In a series of studies in different face-to-face learning settings reviewed by Falchikov (2005) it was reported that *confidence* was the main limiting factor for students to undertake peer-assessment. The lack of sufficient knowledge on the topic made students hesitant about providing fair, objective and responsible feedback to their peers. However, *confidence* was reported to be gained when students were supported and trained (Cheng & Warren, 1997; Dippold, 2009; Dochy, Segers, & Sluijsmans, 1999; Read et al., 2005; Sluijsmans et al., 2002). Whilst peer-assessment is a monitored and required activity in different learning contexts, it is not in MOOCs. In the latter, learners need to overcome limitations of a different nature, starting with those associated to language and culture. Therefore, in this specific learning context, *confidence* appears to be essential for peer-assessment to actually take place. Within this learning setting it has been reported that *confidence* is gained when participants start discussing or sharing knowledge with others in the community (Urrutia, Fielding, & White, 2016). Although their study referred to participants who had performed a mentoring role in a MOOC and therefore had a

responsibility to perform such role, the possibility that this would apply to all participants engaging in peer-assessment activities is open.

Confidence, understood here as the acknowledgement of being able to contribute and to see own contributions as valid and valuable, appears to be the foundation for learners to start the development of critical thinking and the process of becoming autonomous.

In the context of mediated learning settings, learning autonomy has been identified as an essential aspect (Bouchard, 2009). An important part of this autonomy is already being reflected in participants attending MOOCs, where they decide whether to learn, what to learn and when to do it. However, these are only a few of the different competencies that are associated with learning autonomy. The 'how' to learn is more an aspect of learner autonomy that is directly connected to self-directed learning (SDL). According to Knowles (1975), SDL is a process initiated by the learners themselves in which they identify own learning interests and needs, set own learning goals and also measures for learning. These learners are able to identify the resources required to meet their goals as well as establish whether and when they require assistance. These characteristics are included in what Bouchard (2009) identifies as dimensions influencing learners to apply autonomous learning strategies. According to Bouchard, two of the dimensions include psychological and pedagogical issues.

The former he calls '*conative'* and considers four aspects: drive, motivation, initiative and confidence. The latter he calls '*algorithmic*' and considers the previously

mentioned characteristics of SDL. Boyer, Edmondson, Artis and Fleming (2014) claim that the four elements in the first dimension provide the basis for entering the second dimension.

In a way, one could argue that MOOC participants meet at least many of these characteristics already. This, however, does not imply that the 'contextual' knowing stage has been reached by all. In fact, participants' diversity in all its aspects will differ in the stages of knowing reached. Nevertheless, even if the conditions are met –namely, that participants are critical thinkers, thoughtful and autonomous and are thus able to create rich feedback– the conditions for feedback delivery need to be assured.

2.4.2 Information attributes

It is evident that formative assessment requires communication (Sadler, 1998) either written or spoken. In the MOOC context, this communication is essentially written, and participants are confronted with different obstacles. In this particular environment, the command of the language in which the course is offered represents the main obstacle. However, beyond the linguistic limitations, other issues related to the attributes of the information require closer attention.

According to some authors, feedback primarily needs to be sufficiently clear and precise (Ferguson, 2011; Lizzio & Wilson, 2008; Moon, 2005; Prins et al., 2005) in order for its receptor to understand it and make use of it. Clarity and precision are reflected in information being simple, structured, concise and stimulating (Strijbos, Narciss, & Dünnebier, 2010). However, the capacity to write clearly is directly connected to critical thinking and to a writer's proficiency with language (Moon, 2005). In fact, writing represents the thinking process. The way feedback is structured and how evidence is presented reflect the quality of information delivered.

Clarity and precision can be present in four different types of feedback: authoritative, interpretive, probing and collaborative (Lockhart & Ng, 1995). Interestingly, these types appear to reflect the stages of critical thinking that were discussed in the previous section. The information contained in the *authoritative* type focuses on problems or errors without providing any explanations or suggestions for improvement. The *interpretive* type evaluates the product and provides suggestions for improvement based on personal opinions. In the third type

probing, the context in which the product has been developed is taken into account to offer evaluative comments and suggestions for improvement. Finally, the *collaborative* type contains eliciting information that invites dialogue in order to negotiate meaning. The first type (*authoritative*) resembles the first stage of critical thinking (*absolute*), in which an own opinion is lacking, thus no explanations or recommendations are offered. The second type (*interpretive*) mirrors the second phase of critical thinking (*transitional*). Here, information reflects an own opinion, yet it does not offer an evaluation based on contextual knowing. The third type (*probing*) corresponds to the third phase of critical thinking (*independent*) in which comments and suggestions are provided based on the context of the assessed product. This and the fourth type (*collaborative*) are compatible with the fourth stage of critical thinking (*contextual*), which understands knowledge as contextual and is interested in the co-construction of meaning. Based on this reflection and having in mind the type of task of the analysed MOOC, the desired types of feedback would be *probing* and *collaborative*. However, a writer's language ability will be decisive in elaborating feedback that is simple, structured, concise and stimulating.

As a conclusion, it can be argued that learners' stage of critical thinking (e.g. *absolute, transitional, independent*, or *contextual*) and abilities with language determine the way in which information is delivered (*clear* and *precise - simple, structured, concise* and *stimulating*). Although learners' stage of critical thinking can be reflected in the type of feedback that is delivered (*authoritative, interpretive, probing* or *collaborative*), the type of task determines the type of feedback that is required. This said, learners' stage of critical thinking will be reflected in the content

and style characteristics of the feedback provided, as will be the degree of validity and reliability. However, there are no implications regarding feedback's usefulness.

According to Narciss (2006) the components of the information delivered within feedback may concern five areas: task (requirements), concepts (for processing the task), inaccuracy of information (misconceptions), procedure (how to continue), and metacognition (strategies). The information that can be provided within the different areas can refer to hints, explanations, examples or guiding questions.

Each of the components may prove valid and meaningful depending on the current state of learning of the person who is receiving the qualitative information. This reflection goes in line with a conclusion made by Prins et al. (2006) and Read et al. (2005) that the value of feedback lies in the perception of the person receiving it. Nevertheless, it appears that specificity and elaboration in the information (Goodman, Wood, & Hendrickx, 2004; Strijbos et al., 2010) can have an effect on readers' perceptions and uses. The following section presents the different types of feedback that have been reported as useful in the literature.

2.5 Useful types of feedback

As stated in the previous section, rich feedback can be seen from two perspectives, and usefulness of feedback seems to be a debatable topic. It appears that the literature has not reached any agreement on the types of feedback that are useful and why they are useful (Nelson & Schunn, 2009), as feedback quality or usefulness may depend on the receptor's perception as well as on the moment in learning in which feedback was received. Nevertheless, it seems that an agreement has been

reached concerning the important role that the social dimension of feedback plays along the content and organisation of the feedback's message (Evans, 2013).

In fact, the different focus within the literature addressing feedback in terms of validity, perceptions, acceptance, implementation or usefulness (i.e. its quality), recognise that feedback plays cognitive, metacognitive and motivational functions (Narciss & Huth, 2004; Nicol & Macfarlane-Dick, 2006).

Narciss and Huth (2004) argue that the value of a feedback message can be influenced by three factors: the nature and quality of the message itself, the characteristics of the learning setting and the characteristics of learners individually. Whilst awareness towards the second and the third factors has been raised in previous sections, in which the diversity of the participants in the MOOC setting has been highlighted, for the purpose and interest of this investigation the focus will now be placed on the first factor.

The 'nature and quality' of the message contained in the feedback is determined by the following three dimensions: functional, structural and semantic. The first dimension refers to the learning objectives, hereby acknowledging the cognitive, metacognitive and motivational functions of feedback. The second one considers the formal and technical aspects of the delivery of the feedback message (e.g. who, when, how long). Finally, the third dimension regards the content of the feedback message and its meaning. It is the third dimension that is explored in the present investigation.

Studies considering the semantic dimension of feedback have principally focused on finding the types of feedback that are considered useful and effective from two perspectives: that of the authors, and that of the participants. Studies from the former perspective have tried to find out correlations between the feedback provided (either by the instructor or peers) and participants' qualities of performance in different learning settings. Studies from the latter perspective have taken into account participants' perceptions on the topic. The context of the reviewed studies ranges from secondary education to graduate education level, and the education format include face-to-face and online environments. The research methods employed in both perspectives have been quantitative and qualitative in nature.

In the MOOC setting, the semantic dimension of feedback appears still unexplored, being the study by Krogstie et al. (2015) the only example found in the literature. Based on the works by Hattie and Timperley (2007) and Gielen et al. (2010), their study created a set of criteria to evaluate the quality of peer feedback in a Norwegian MOOC for vocational training in which participants were teaching practitioners from that country. Addressing this gap in the literature, López-Benavides (2015) –the researcher of the present investigation– carried out an exploratory study to analyse the semantic dimension of peer-feedback in one of LDS' MOOCs. The study followed a case study approach and employed the method of content analysis to explore peer-feedback data from two out of five learning phases. Details on this study are presented in Section 2.6.

In the following sub-section, the types of feedback that have been considered useful based on the author's perspectives, are presented.

2.5.1 Usefulness of feedback from an author's perspective

Different studies were reviewed in order to understand the manifold types of feedback that have been considered useful by the authors undertaking the studies, as they have led to improvements of student's products. The most salient studies are presented below.

To begin with, the study by Tseng and Tsai (2007) is presented. This study analysed peer-feedback in a high-school computer course (n=184) and concluded that two types of feedback –namely: *reinforcing* and *suggestive*– proved to be useful as an improvement in students' work could be established. The first type, *reinforcing,* recognised what was done correctly by highlighting the good parts of the work, and the second type, *suggestive,* advised of a problem encountered in the work, without further specifications. This second type only proved useful at an early stage of the course.

A second study undertaken by van der Pol et al. (2008) investigated peer-feedback provided by two different groups in an online environment. In the first group feedback from 'Health Care Education' students from a college (n=27) was studied, and in the second group feedback of students (n=38) from an 'Educational Science' course in a university were analysed. For the first group, it was reported that the type of feedback *revision* was the one leading to improvements in student's works, as it contained direct and concrete suggestions. For the second group an

improvement in student's work was reported when feedback was aimed at understanding the work and its content (*analysis*), and when it compared the content of the work with the task requirements (*evaluation*).

Another study carried out by Cho and MacArthur (2010) in an undergraduate course (n=30) in 'Research Methods of Psychology' found that the writing quality of the participants increased when *non-directive* feedback was provided. The writing quality was measured if complex repairs and extended content revisions had been done. *Non-directive* feedback refers to comments on general matters without making any suggestion. A further study carried out by Gielen et al. (2010) that analysed written assignments in secondary education (n=43) reported that feedback providing *justification* improved students' performance. In providing a justification, the reviewer supported the judgements they made.

In the context of collaborative writing, a study carried out by Guasch, Espasa and Álvarez (2010) reported that students' performance had increased when students received feedback that combined *suggestions* with *questions*, or *suggestions* with *corrections*. The study took place in an online university and involved a collaborative writing task in a post-graduate course (n=83). The same findings were reported in a further study carried out by Alvarez, Espasa and Guasch (2012), claiming that through those types of feedback the quality of argumentation within students' texts greatly improved.

Building on the previous two studies, Guasch, Espasa, Alvarez and Kirschner (2013) analysed the improvement of the quality of a collaborative written task in a course in

a ‘Psychology’ bachelor’s degree (n=201). They concluded that an improvement was shown when students were questioned about certain parts of their work (*epistemic*) or when students were questioned about any part of their work and at the same time received a hint on how to improve (*epistemic + suggestive*). Guasch and Espasa (2015) concluded the series of the previous three studies in the same context, and suggested that in order for feedback to contribute to higher quality learning, feedback needs to offer questions and proposals for improvement, emphasising that feedback needs to contain both *epistemic and suggestive* feedback. These authors point out that the individual use of these types of feedback did not prove effective and that if used that way they may even yield opposite effects.

The authors argue that the combination of *epistemic and suggestive* feedback include the components of *verification* and *elaboration* proposed by Kulhavy and Stock (1989) and that were introduced in sub-section 2.3.1. When feedback contains questions, it enables others to understand that something in the work is either unclear or incorrect, implying the *verification* component. When feedback includes hints for improvement, it invites the learner to explore and expand and so it accomplishes the *elaboration* component.

Findings of the previous studies confirm a variation in the perception of useful feedback. All studies compared the types of feedback used and the improvements that were made in students’ products based on the types of feedback received. Findings seem to imply that different types of feedback may prove useful in different contexts of study, types of participants, types of required tasks, and possibly the cultural context in which peer-feedback was analysed.

Following, the perception of useful feedback as regarded by participants will be exemplified by important works from the literature.

2.5.2 Usefulness of feedback from a participant's perspective

The first study reviewed was the one carried out at university level by Cho, Schunn and Charney (2006) with undergraduates (n=88) in the course 'Cognitive Psychology' for Non-Majors. The authors reported that the *directive* and *praise* types of feedback were perceived positively by students, since they preferred detailed comments with specific suggestions (*directive*) as well as comments expressing admiration for the work or parts of it (*praise*).

The study undertaken by Prins et al. (2006) involved participants training to be general practitioners (n=46) from a post-master level. *Descriptive*, *reflective* and *personal* types of feedback were perceived as useful, in which reflective questions were posed along with examples and suggestions.

Lizzio and Wilson (2008) studied the perceptions on feedback received by psychology, law and arts students (n=57) in their second or third year of university studies, and by students (n=277) in the programmes of psychology, criminology, science and engineering from various levels of university studies. The authors concluded that the perceptions of both groups could be understood in three areas: developmental, encouraging and fair feedback. The first group perceived feedback as useful when it was *transferable,* when it showed evidence of *deeper engagement* from the reviewer's side, and when it used *socioemotional components,* for example by being considerate in their tone while offering a criticism, and by letting the

student know that their work was valuable. The second group perceived feedback as useful when it posed questions and reflections that enabled the reader to think (*engaging*), and feedback that praised the work and acknowledged the effort (*encouraging*).

A study undertaken by Poulos and Mahony (2008) inquired into students' perceptions on feedback. The number of participants was not provided. Participants were undergraduate students from the Faculty of Health Sciences at a university. Based on students' comments, the author concluded that a joint understanding of the meaning of effective feedback, or the way it could be used, did not exist. However, those that were able to identify types of feedback, said to prefer *timely* and *specific* feedback over late and general one. The study concludes with an interpretation of useful feedback provided by one of the participants "… feedback needs to be provided to you so you can actually make a change … if you can't make a change from what's provided then it's useless" (p.153).

Finally, a study conducted by Ferguson (2011) provides a detailed insight into the perception of useful and effective feedback by graduate students (n=465) and undergraduate students (n=101) who studied teacher education at a university. The main characteristics of useful feedback were *timely* and *personalised* feedback that was specific to their piece of work. The content of the personalised feedback should contain *positive* and *constructive* comments, recognising the invested time in the work and offering suggestions for improvement. Many students noted that *positive* feedback was useful and necessary for the purpose of gaining confidence and keeping motivated. Most students agreed that a clear link between the assessed task

and the assessment criteria needed to be provided. Students in this study had clear expectations regarding useful feedback, especially on how feedback should be structured. For a normal assignment, the majority of students expected to receive a paragraph with a brief summary of the assessment and a longer paragraph that highlighted both their work's strengths and weaknesses. The latter should be carefully formulated and also detailed in order to guide students to improve their work.

This literature review also shows that the perception of useful feedback is different among students of different levels and disciplines of study. The last study provided a more focused view on feedback and the reason was the type of learners and the discipline studied. Revisiting the purpose of formative assessment, it is said that it aims at enhancing learners' knowledge, skills and comprehension of a topic studied (Shute, 2008). However, as pointed out by Lizzio and Wilson (2008), whilst feedback that helps students to further develop is academically greatly valuable, its value may decrease if it does not contain an encouraging element. This clearly supports the functions that feedback play in the cognitive, metacognitive and motivational dimensions (Narciss & Huth, 2004; Nicol & Macfarlane-Dick, 2006).

To conclude this section, it can be argued that feedback usefulness is reflected by its quality (Black & Wiliam, 1998), and that quality is:

> not just the technical structure of feedback (such as its accuracy, comprehensiveness and appropriateness) but also its accessibility to the learner

(as a communication), its catalytic and coaching value, and its ability to inspire confidence and hope. (Sadler, 1998, p. 84)

2.6 The semantic dimension of peer-feedback in a MOOC

As mentioned earlier, the content of peer-feedback in MOOCs has not yet been explored to my knowledge. Only after having analysed the content of peer-feedback in a course is it possible to describe it in terms of quality, as well as to determine and describe possible differences or similarities based on the characteristics of the participants providing it.

In response to this gap in the literature, an exploratory study was carried out by López-Benavides (2015). Through a case study methodology and by means of content analysis, findings of the study confirmed the presence of types of feedback fulfilling the cognitive, meta-cognitive and motivational functions of feedback as suggested in the literature (Narciss & Huth, 2004; Nicol & Macfarlane-Dick, 2006). Also, types of feedback were found that addressed *Style*, *Content* and *Structure*, which have been referred to as *aspects* of feedback by van den Berg et al. (2006) and van der Pol et al. (2008).

In the sample analysed, which considered feedback in two learning phases, 11 types of feedback were identified. The types of feedback that stood out the most were those relating to the motivational functions of feedback, followed by the aspects of feedback and finally the cognitive functions of feedback. In order of appearance, these were: *Praise, Structure, Mitigating language, Content, Style, Evaluation,*

Analysis, Not-applicable, Inflammatory language, Revision, and *Explanation* (see Appendix Two).

These types of feedback had also been identified in the literature in other learning contexts, presumably indicating that types of feedback unique to MOOCs may not exist.

The pilot study presented limitations in regard to the size of the sample and the reliability of the instrument that was created to analyse feedback content. The former could have had implications in the types of feedback found, as the data was only extracted from two learning phases with respectively two different types of learning tasks. The latter could have had implications for the consistent analysis of the content of peer-feedback.

It is that study on which the present investigation builds. The aim of the present study is to further explore the nature of feedback in the MOOC learning context by addressing the before-mentioned limitations. Content of peer-feedback will be analysed from all learning phases of a course (n=6) considering the types of tasks that needed to be submitted in each learning phase. For this, the first step required is enhancing and validating the instrument that aims to analyse feedback content.

Based on the findings of all previous studies reviewed, recognising the differences in the learning contexts and their research limitations, hypotheses for each of the research questions of this study are presented next.

2.7 Hypotheses and justification

The principal aim of this investigation is to explore the nature of peer-feedback present in a MOOC. The specific research questions guiding this study and respective hypotheses are presented below.

Research Question 1

How does feedback evolve over time in terms of quantity and quality?

Hypothesis 1-1

The quantity of peer-feedback provided will fluctuate throughout the learning phases.

Justification

Considering the learning setting, participants' diversity and the fact that the assessment activity is voluntary and not monitored, then a fluctuation in the quantity of peer-feedback is principally expected. However, beyond these considerations, the fluctuation in the quantity of feedback is argued to depend heavily on the type of task of a given learning phase. As mentioned earlier, the types of tasks included in the case study were theoretical and practical in nature, and some of them combined both. Thus, it is possible that participants showed more preference for some types of tasks than for others, and that this preference will be reflected in the quantity of feedback provided. Findings of the study by Neubaum et al. (2014) in this specific context support this hypothesis, but also highlight other

factors that may cause a variation in terms of quantity, such as the length and content of the work to be assessed.

Hypothesis 1-2

Independent from the quantity of feedback provided, its quality will tend to increase over time.

Justification

It is evident that each type of task has different requirements in terms of feedback content, form or even presentation as discussed by Narciss and Huth (2002), thus the quality of feedback can only be compared to tasks of a similar nature. Although training on peer-assessment was not offered in the course, it is assumed that the different learning phases provided all participants with a space for training in the provision of feedback, for becoming more familiar with the process of assessment, its criteria, the topic learned, their peers' work (e.g. their writing styles and working topics) and that the feedback provided by other peers have served as inspiring or deterrent examples from which to learn. As a result, it is expected that participants undertaking peer-assessment have developed the confidence required to provide feedback of quality which –as considered in this study– include questions about the work or parts of it, and ideas on what can be improved and how it could be improved.

Research Question 2

How does the nature of feedback differ between students and supporters?

Hypothesis 2-1

Students' peer-feedback will more likely focus more on the cognitive functions of feedback than on the motivational functions. Conversely, supporters' feedback will focus more on the motivational functions of feedback than the cognitive ones.

Justification

As explained earlier, the roles of each type of reviewer were different. In short, students had obligations to attend, and supporters had only expectations to meet. In order to motivate both types of participants to engage in the process of peer-assessment, badges were awarded.

The course's learning material (e.g. literature, video keynotes, assignments) was available to both types of participants. However, only students were required to have engaged with the learning materials in order to complete the tasks. This requirement allowed students to undertake a cognitive analysis for each of the tasks (Narciss & Huth, 2002), which is an essential step for being able to identify the elements that need to be contained in the feedback. From this perspective, it is believed that students were in an advantageous position compared to supporters. Because of the difference in roles, the presence of types of feedback specific to each role is expected.

Research Question 3

How does the nature of feedback differ between female and male participants in their respective roles as students or supporters?

Hypothesis 3-1

There are likely no differences to be found between the nature of feedback provided by female or male participants.

Justification

Despite the debatable findings in the literature regarding gender and communication behaviour (Yukselturk & Bulut, 2009) as discussed in Section 2.2, this study builds on the understanding of gender as a social category, that is socially constructed. Because of the diversity that has been previously highlighted in this kind of course, stereotypical behaviours such as those reported in the literature (e.g. Blum, 1999; Gunn & McSporran, 2003; Herring, 1993; Price, 2006; Rovai, 2001; Yukselturk & Bulut, 2009) are not expected to be found. However, this is not to say that differences will not emerge, but it acknowledges other potential factors (e.g. role assumed in the course, age, educational background, first language, type of learning task) as responsible for the differences.

Chapter 3 Methodology

This chapter presents the design of the investigation. It starts by introducing the methodology and principal method employed, as well as the different phases that were planned for the investigation. Details on the case of study are provided, and the analytical strategy that was undertaken for each of the planned phases of study is explained and justified. Next, the analyses of the study are presented and the procedures undertaken for data analysis are described in detail. The chapter concludes with a critical reflection on the methodological decisions made for this investigation.

3.1 Research Design

This study follows the principles of the constructivist inquiry paradigm and also fulfils its requirements (Lincoln & Guba, 1985). Compliance with the paradigm and the trustworthiness of the study in its different stages are discussed in Section 3.6.

Little investigation has been done on the focus of the present study, thus exploration through the case study approach was deemed suitable. This type of approach can be viewed as a research process (Yin, 1994), as the unit of study (i.e. the case) (Stake, 1995), or as an end product (Merriam, 1988). However, as Merriam (1998) points out “the single most defining characteristic of case study research lies in delimiting the object of study, the case” (p.27). The case, Merriam defines as “a thing, a single entity, a unit around which there are boundaries” (p.27). The case under investigation is a single MOOC from the online learning offers provided by the Leuphana University and the focus of study laid on peer-feedback. The boundaries mentioned by Merriam refer to a specific phenomenon that happens in a bounded

context and from which its variables cannot be separated. Thus, the present study is interested to explore how peer-feedback occurred in a specific online learning event. It takes into account the context in which it took place and acknowledges the variables directly relating to the phenomenon under study. This case study is synchronic, particularistic, descriptive and heuristic (Merriam, 1998; Mills, Durepos, & Wiebe, 2010). The first feature indicates that the purpose of the study is not affected by the time that has passed since the event. The second suggests that the focus of study is placed on a specific situation, event, program or phenomenon. The Third that the study is concerned with stating how things have been or have happened in the case at hand. Finally, the last feature suggests that the case study sheds light on the understanding of a specific phenomenon.

This approach is suited to researchers who want to gain deep understanding of an instructional context guided by descriptive and explicative questions (Flick, 2009; Yin, 2003). Moreover, the relationships and processes that are of interest for the research can be examined by employing diverse methods for data collection, types of data and research methods as part of the investigation (Denscombe, 2003, 2010; Merriam, 1998).

Given that the principal interest of the study lies in analysing participants' written feedback, it was considered that the method of content analysis would be appropriate. This method starts with the analysis of qualitative data and is completed with a quantitative output of the analysed data, facilitating its summarisation and further exploration. This quantitative output was used descriptively and for carrying out statistical analyses (Denscombe, 2010; Stan, 2010).

The investigation was designed to be undertaken in five phases.

The first phase served to become familiar with all relevant information around the case at stake and course participants. In relation to the former, information about the conception and pedagogical design of the course, information published in the learning platform, and reports in numbers about participants' activity during each learning phase, were studied. Regarding course participants, survey reports helped to understand their views on peer-feedback. Moreover, in this phase, primary data were gathered for the second and subsequent phases of analysis.

In the second phase the nature of feedback in the MOOC was explored and identified. To this end, a coding system that allowed for categories addressing the aspects and functions of feedback was used. In the subsequent phases, the quantitative output of the previously analysed data was used to report on the quantity and quality of the feedback provided during the course and on the differences on the frequency of use of types of feedback distinguishing the role of the participant in the course and their gender in their respective role.

Figure 3.1 displays the systematic order that was followed in this investigation.

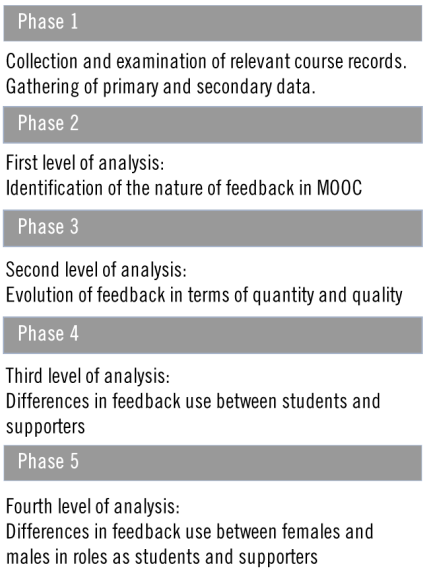


Figure 3.1 Phases of the present investigation.

3.2 Context

3.2.1 The course

This study focuses on the informal assessment, specifically the written feedback provided by students and supporters in all six learning phases of the fourteen-week course “Managing the Arts: Marketing for Cultural Organizations”, which was organised by the Leuphana University in cooperation with the Goethe Institute. The course ran from February 19 until May 28, 2015; it was conducted in English, and had 17,000 enrolled learners from around 170 countries. From the enrolled learners, only 800 were allowed to register as students. A raffle was run for the process of selection. Details on the topics treated in each learning phase and suggested assignments are presented in the findings section.

The target group were practitioners in the field of arts management and cultural marketing (e.g. arts and cultural managers, artists, cultural marketing experts), as

well as graduate and undergraduate students of associated fields. However, anyone interested in the MOOC was welcome. The aim of the course was to offer additional skills and academic insights into the management and marketing of cultural organisations, for which four different case studies were explored.

It seems important to note that the researcher of this study had worked at the Leuphana Digital School and in the running of different MOOCs. However, in the case at hand, the researcher participated merely as an observer. Details on the assumed role as well as the implications for the study are discussed in section 3.6.

3.2.2 Participants and sampling

A total of 824 participants undertook assessment at some point during the course. This population was studied to inquire into the nature of feedback in the course.

For all subsequent phases of analysis, purposive sampling was used (Denscombe, 2010). From the population of 824 participants, 235 were awarded with the Popular Evaluator type of badge. It was interesting to analyse and describe the types of feedback that this group of participants used throughout the phases of the course and that were regarded as useful. Likewise, possible differences in terms of participants' characteristics were worthy of inquiry.

Table 3.1 summarises the demographic information of the population. Information was retrieved from participants' profile and records from the learning platform. Note that a third gender option named 'hidden' could be chosen.

Table 3.1 MOOC participants' demographic information (N = 824).

	Gender		
Career level	female	hidden	male
Academic	50	4	16
Professional	409	27	143
Retired	0	0	5
Student	103	6	32
Unemployed	25	1	5
Educational level			
Bachelor	200	8	97
Doctorate	37	6	3
High school	16	1	12
Master	297	18	79
Other	37	5	9
Average age	32	35	34

Note: Information is based on that retrieved from the learning platform.

3.2.3 Data and ethical considerations

Data comprise the content in participants' peer-feedback, which was provided throughout the entire MOOC. Data were formally collected in 2015 after the completion of the course and were treated with confidentiality. Participants' original identities were protected by the use of identification (ID) numbers that were automatically generated by the learning platform as participants enrolled in the course. This project examined records relating to humans, yet it was considered of low risk. The focus of analysis was placed on the content and not on the author of the data. Furthermore, although no contact with the participants was required, the study made use of participants' demographic information as a means to provide deeper contextual information on the case. Participants agreed to the Privacy Policy and Terms of Service of the course, which stated that data that had been entered and generated during the course and within the learning platform may be used for

research purposes. Ethical approval for this study was approved by the Lancaster University Research Ethics Committee on April 21, 2015.

3.3 Analytical strategy – data gathering and analysis

The first phase of the study considered the collection and examination of relevant course records that would help gain deeper contextual information on the course, its progress throughout the learning phases and its participants. To this purpose, secondary data of different types were retrieved from the learning platform in which the course took place. Special interest was taken in the information concerning four main topics: the course itself, roles, badges and evaluation. The first one considered the presentation of the topic of the course, the learning objectives, the assignments and its descriptions. The second one included descriptions of the types of roles assumed in the course and its responsibilities. The third comprised the explanation of the types of awards that could be earned and the conditions for earning them. Finally, the fourth regarded the presentation and description of the three assessment criteria that were required when evaluating the work of the participants.

The progress of the course throughout the learning phases could be followed by course updates that were published and learning phase reports that were issued. The former were important announcements that were published anytime by *the MOOC Facilitator* depending on the development of the course. The latter consisted of statistical reports issued by the learning platform provider after each learning phase.

The course updates helped to recall and understand the rhythm, the activities, the 'ups and downs' during the course, and they contributed to provide possible explanations for the development of feedback at certain moments of the course. The reports provided an overview in numbers of the course activity during each learning phase. For instance, the number and type of badges that had been awarded in each phase were reported there. This information could be compared and further explored in the datasets that were created for the study.

Finally, informal e-mail conversations were held with the creators of the course in order to gain deeper understanding of the design of the different elements considered under peer-evaluation. During those conversations, post-course survey reports were shared with the researcher. A specific section of the survey that considered the topic of feedback was extremely helpful in understanding the perceptions of those participants taking part in the survey. The survey took the form of a structured web questionnaire and included open-ended type of questions, multiple-choice, and scaled items. The latter used a 5-point Likert scale (Applies fully, Applies, Applies slightly, Does not apply at all, No answer).

The questionnaire was slightly different for each of the roles. There were six items relating to peer-evaluation that were addressed to the students, and five addressed to the supporters.

A total of 472 participants took part in the post-course questionnaire; however, not all participants completed it. Out of this total, 190 were students and 282 supporters.

Students' answers helped to understand how feedback provided by their peers was perceived. To the statement: "The evaluations helped me and my team to improve our work in the following learning phases", students answered as displayed in Table 3.2.

Table 3.2 Students' answer to statement: The evaluations helped me and my team to improve our work in the following learning phases. (n= 171)

	Applies fully	Applies	Applies slightly	Does notapply atall	No answer	No response
	17.5%	33.3%	26.3%	8.7%	9.3%	4.6%

Furthermore, to the question: *During which learning phase(s) did you receive the most useful evaluations?* (multiple answers were possible) students responded as shown (see Table 3.3).

Table 3.3 Students' answer to question: During which learning phase(s) did you receive the most useful evaluations? (n= 171)

	Learning phases							
	1	2	3	4	5	6	All	none
	14%	13%	11%	19%	17%	17%	22%	0.5%

Out of all participants taking part in the questionnaire, 129 students and 101 supporters said to have evaluated the work of their peers at least once. On average, the estimated amount of time invested per peer-evaluation was 20 minutes for students and 30 minutes for supporters.

Some of the background questions that were asked of participants in both roles and that are of interest for this study are summarised in Table 3.4.

Table 3.4 Questions and answers about feedback addressed to students and supporters.

	Likert-scale					
	Applies fully	Applies	Applies slightly	Does not apply at all	No answer	No response
Evaluating my peers' submissions was very enriching for me.						
Supporter	35%	30%	24%	2%	4%	5%
Student	19%	41%	32%	4%	-	4%
I evaluated more submissions in order to earn the "Active Evaluator badge".						
Supporter	16%	21%	3%	26%	30%	4%
Student	12%	10%	5%	30%	43%	-
I put a great deal of effort into providing meaningful evaluations in order to receive good ratings and earn the "Popular Evaluator badge".						
Supporter	22%	17%	2%	24%	30%	5%
Student	15%	15%	10%	22%	36%	2%
I wanted to earn the required badges to be eligible for the Statement of Accomplishment.						
Supporter	23%	19%	6%	21%	25%	6%
Student	14%	17%	9%	22%	37%	1%
Before evaluating the work of my peers, I reviewed each evaluation criterion (relevance, substance, quality & clarity), its explanations and guiding questions.						
Supporter	35%	37%	17%	2%	2%	7%
Student	25%	46%	15%	8%	1%	5%
The explanation for each evaluation criterion and its guiding questions were easy to understand.						
Supporter	22%	38%	24%	6%	4%	7%
Student	19%	36%	30%	5%	-	5%
I am confident that the feedback I provided with my evaluation was very useful for my peers.						
Supporter	20%	39%	24%	1%	10%	6%
Student	15%	35%	41%	3%	-	6%

Note: n= 129 students, *n*=101 supporters.

Once deeper contextual information on the case at hand was gained, data for the second and subsequent phases of analysis were gathered. Primary data were retrieved directly from the learning platform's database as .csv files. Two datasets were created from the retrieved data with those records of interest for the study. The first dataset was used for the first level of analysis that aimed at exploring the nature of peer-feedback within the MOOC, and the second dataset was employed

for all subsequent analyses that examined the quantity and quality of the feedback provided during the course and the frequency of use of types of feedback distinguishing the role of the participants in the course and the gender in their respective role.

The first dataset covered the feedback provided to all submitted assignments in all six learning phases, the ID of the participant providing it, the role and gender of the participant.

The second dataset referred to the records of awarded badges. The records that could be retrieved for this set included the ID of the participant earning them, together with the participant's role and gender and the type and quantity of badges that were earned.

Once data for both datasets were retrieved, the process of data cleaning started. To support this purpose, a guideline was prepared for each dataset in order to follow a consistent and also systematic procedure. The final aim was to create clearly arranged databases that would help the researcher subsequently during the processes of analysing and coding. For both processes, the software Microsoft (MS) Excel 2016 was used.

The first dataset was organised in columns as displayed in Table 3.5.

Table 3.5 Data required for the first dataset.

				Feedback for:		
Phase	ID	Gender	Role	Relevance	Substance	Clarity and coherence

Once the database was organised, documents were created and saved per learning phase for each of the three assessment criteria. A total of 18 documents resulted from this process. These were exported and saved as a rich text format (.rtf) and imported to the software for qualitative analysis Atlas.ti (version 7).

The total number of messages analysed in the study, as well as the number of messages written by the participants in their respective role and gender option, are displayed in Table 3.6. From the total messages numbering 8,479, the first learning phase showed the highest number of messages (923), followed by learning phases two and five with 821 and 708 messages respectively. Learning phases three and six reported a similar number of messages, 669 and 692 respectively, and learning phase four showed the lowest number of messages with 605. Students wrote 4,418 messages and supporters 4,061. Out of the students, female participants wrote a total of 3,127 messages, followed by male participants with 1,014, and by hidden participants with 277. Comparably, in the group of supporters, the number of messages written by female participants was 3,092, while male participants wrote 769 and hidden participants wrote 200.

Table 3.6 Number of messages written during all learning phases (N=824).

Phase	Role Student			Total	Role Supporter			Total	Grand total
Gender	female	hidden	male		female	hidden	male		
1	623	64	236	923	796	40	239	1075	1998
2	585	53	183	821	664	22	203	889	1710
3	450	43	176	669	504	14	106	624	1293
4	410	42	153	605	331	12	64	407	1012
5	512	54	142	708	381	81	81	543	1251
6	547	21	124	692	416	31	76	523	1215
Total	3127	277	1014	4418	3092	200	769	4061	8479

Note: 1 message contains feedback for each of the three assessment criteria: *Relevance*, *Substance*, and *Clarity and coherence.*

Table 3.7 summarises information from the second dataset that considered the awarded badges. The amount of Active and Popular Evaluator badges awarded per learning phase during the course are displayed there. A total of 3,395 badges were awarded during the course, 2,713 corresponding to the type Active Evaluator (referring to quantity), and 682 to the type Popular Evaluator (referring to quality). The latter type of badge, which is of interest for this study, demonstrates interesting behaviours throughout the phases. Whilst in the first phase only 45 badges were awarded, in the next three subsequent phases this figure was doubled. Intriguingly enough, the number of badges awarded in phases five and six tripled the amount of the badges awarded in the first phase.

Table 3.7 Number of badges awarded during the course per learning phase.

	Badge type		
Phase	Active Evaluator	Popular Evaluator	Total
1	576	45	621
2	542	107	649
3	428	93	521
4	335	110	445
5	429	158	587
6	403	169	572
Total	2713	682	3395

Note: Numbers in this table were obtained from the learning platform provider.

Whilst a total of 824 participants provided feedback at some point during the course, only 235 of them were awarded with the Popular Evaluator badge. Feedback provided by these participants were analysed in detail in phases three, four and five of the study in which specific analyses were planned. Data available from these participants were organised in a MS Excel table and included: IDs, role and gender, number of badges received in total[2], and the quantity of messages (feedback) written per learning phase. Table 3.8 provides an overview of the number of messages written by these participants during the learning phases of the course.

[2] Data retrieved did not display the learning phase in which individual participants earned their badges, nor the specific messages that were awarded with the badge.

Table 3.8 Number of messages written during all learning phases (n=235).

Phase		Role Student			Total	Role Supporter			Total	Grand total
	Gender	female	hidden	male		female	hidden	male		
1		427	44	113	584	236	24	41	301	885
2		414	51	129	594	427	10	65	502	1096
3		369	35	129	533	370	8	64	442	975
4		320	41	121	482	244	8	23	275	757
5		418	57	119	594	282	77	43	402	996
6		509	31	88	628	302	20	48	370	998
Total		2457	259	699	3415	1861	147	284	2292	5707

Note: 1 message contains feedback for each of the three assessment criteria *Relevance, Substance,* and *Clarity and coherence*

In summary, students wrote 49% more messages than supporters. In both roles, messages written by female participants were considerably higher than those written by male and hidden participants. Specifically, male students wrote 71% less messages than female students and hidden students 89% less. Male supporters wrote 87% less messages than female supporters and hidden supporters 92% less.

3.4 Levels of analysis

The base of communication and interaction within the case of study was written language, and this was analysed by employing the method of content analysis. The levels of analysis that were undertaken in the study are described next.

3.4.1 First level of analysis (Phase 2) - Identifying the nature of feedback in the MOOC

This first analysis focused on investigating the nature of peer-feedback within the MOOC by identifying content that referred to the aspects and functions of feedback. In order to achieve this, the definition of a basic unit of categorisation used for

computer conferencing transcripts was required. From the different units of analysis that are available for this purpose such as physical units, syntactical units, referential units, propositional units and thematic units (De Wever, Schellens, Valcke, & Van Keer, 2006; Donnelly & Gardner, 2011; Rourke, Anderson, Garrison, & Archer, 2001), the one selected was the thematic unit (TU henceforth), as defined by Budd, Thorp and Donohue (1967) "a single thought unit or idea unit that conveys a single item of information extracted from a segment of content" (p.34). Unlike other units, TUs do not follow a predetermined logic, making the identification less objective. However, these units capture the construct that is of interest for the researcher. Besides the suitability of this unit for the type of analyses proposed in the study, its context as well as data available, the unit was chosen as it proved useful in previous studies carried out by the researcher, in which the method of qualitative content analysis was employed.

Having decided on the unit for categorisation, the next step was to use a coding system that would help to identify feedback in relation to its aspects and functions. This study considered using a coding system created by López-Benavides (2015) for a pilot study in which the types of peer-feedback in a MOOC were explored. However, the instrument presented some limitations that were addressed in this study. Sub-section 3.5.1 explains in detail the process for the enhancement of the coding system.

After the enhancement of the coding system, categories within it were revised and discussed by three evaluators. Each evaluator applied the codes to a number of

randomly selected messages from one of the six learning phases of the course. The inter-rater reliability process is described in sub-section 3.5.1.1.

The analysis of the TUs was carried out using the software for qualitative analysis Atlas.ti (version 7).

Once the process of content analysis was completed, the subsequent analyses were performed.

3.4.2 Second level of analysis (Phase 3) - Describing the evolution of feedback in terms of quantity and of quality

In this second step, the evolution of peer-feedback in regard to quantity and quality was observed and described. The quantity aspect reports on the frequency of appearance of the categories identified in the coding system, which corresponds to feedback aspects and functions.

The quality aspect zooms in on specific types of feedback considered from two perspectives, and reports on their frequency of appearance. Perspective one considers the types of feedback that have been identified in the literature as fulfilling the objectives of formative assessment, and that were discussed previously in sub-section 2.4.1. Perspective two examines those types of feedback that were perceived as useful by the course participants. The analysis designed to address the quality factor from this second perspective will be further explained in sub-section 3.5.2.

3.4.3 Third level of analysis (Phase 4) - Differences in the use of feedback between students and supporters

The quantitative output of the qualitatively analysed data was used to report on the differences in the use of types of peer-feedback in relation to reviewer's role. The mean frequency of use of the different types of feedback were studied in conjunction with statistical analyses to explore possible relations between dependent and independent variables. The dependent variable refers to the types of feedback (the categories identified from the coding system), and the independent variable explored in this analysis was participant's role. Furthermore, an additional analysis was performed to establish the type of association between the variables.

The statistical analyses were carried out using the software Statistical Analysis System (SAS) (version 9.4).

3.4.4 Fourth level of analysis (Phase 5) - Differences in the use of feedback between female and male participants in their respective role

This analysis zooms in on the previously described analysis. It makes use of the same quantified data to explore possible differences in the use of types of peer-feedback in relation to reviewers' gender in their roles as students and supporters. The gender variable was included in this analysis to run the same statistical analyses as the ones presented in the previous level of analysis. The description of the statistical procedures for this and the previous analysis is covered in sub-section 3.5.3.

3.5 Procedure for data analysis

This section describes in detail the procedures for data analysis that were undertaken for all four levels of analysis.

3.5.1 First level of analysis: Identifying the nature of feedback in the MOOC

The analysis of the nature of peer-feedback in the MOOC required the use of an instrument: a coding system. The system would allow categorising the types of peer-feedback in terms of its aspects as well as its functions, as suggested by the reviewed literature and by a previous study carried out by the researcher. To this aim, the coding system created by the latter was used (López-Benavides, 2015). It was developed following the directed approach to content analysis (Hsieh & Shannon, 2005), where codes emerge from the theory and from data itself.

The instrument was enhanced and validated by a further inductive and deductive analysis. The inductive analysis reconsidered the reviewed literature (specified later on in this sub-section) from both face-to-face and online learning environments that had contributed to the identification and definition of types of feedback when assessing products or outputs. The deductive analysis resulted from the data themselves. The categorisation process required codes to be re-defined, created, specified and combined until reaching a consistent coding system.

To understand the development of the coding system that was employed in this study, a condensed summary of the process of creation of its original version is presented, and the adjustment for the current one is described.

In its original version, the coding system consisted of 11 categories. On the basis of the three assessment criteria provided in that course (*Presentation*, *Substance* and *Consistency*), a correspondence between the description of the criteria and the literature was identified. Criteria matched to what van den Berg et al. (2006) had referred to as aspects of feedback, and of which there are three: *Style*, *Content* and

Structure. The aspects of feedback focus on specific features of the assessed piece of work: the formal aspects of writing (*Style*), the information within it (*Content*) and the connection of ideas (*Structure*). During the data analysis, feedback provided within these criteria revealed the presence of other types of feedback that appeared to fulfil cognitive and motivational functions, as suggested in the literature (Narciss, 2008; Narciss & Huth, 2004; Nicol & Macfarlane-Dick, 2006).

The assessment criteria were guided by questions which allowed the assessor to explore different dimensions from which the work could be evaluated. According to Flower, Hayes, Carey, Schriver and Stratman (1986), when assessing a work of this type, the assessor engages in four steps: *Analysis*, *Evaluation*, *Explanation*, and *Revision*. These four steps were incorporated as categories in the coding system representing the cognitive function of feedback. The descriptions as defined by van den Berg et al. (2006) were used.

Regarding the motivational function of feedback, data displayed similar characteristics as those reported in the works of Nelson and Schunn (2009). Three categories referring to *Praise*, *Mitigating language* and *Inflammatory language* were identified, and therefore adopted and redefined according to what the data had revealed. Lastly, for feedback that was considered off-topic, the category *Not applicable* was created.

The original coding system can be appreciated in Figure 3.2.

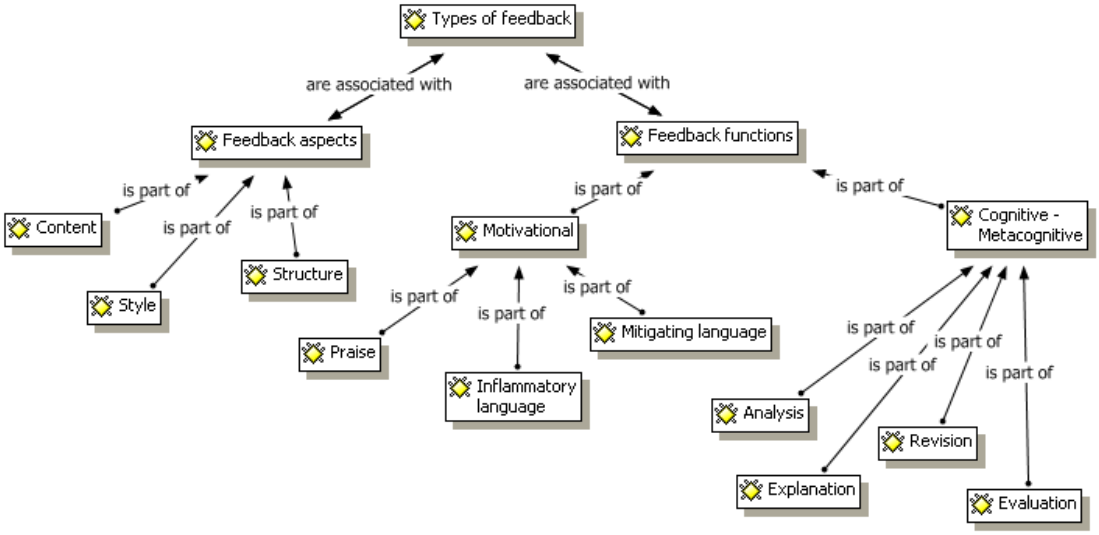


Figure 3.2 Coding system for the identification of types of feedback in a MOOC developed by López-Benavides (2015).

The case at hand differed in the definition of criteria for assessment. Whilst criteria in the previous case addressed all aspects of feedback, the ones in this course solely focused on the content aspect. The three assessment criteria defined for this course were: *Relevance*, *Substance and Clarity and coherence.* Thus, categories specifically referring to these criteria were incorporated in the coding system, and its descriptions were guided by the same information provided to the participants. Despite the redefinition of criteria in this course, the data kept displaying the presence of the other two aspects of feedback: *Style* and *Structure*.

In order to pilot and continue enhancing the coding system, a further familiarisation with data was required. Thus, the process of codification was planned to be undertaken in three rounds.

First round

In this round, TUs were identified for the first dataset (18 documents) and codified into 3 categories: learning phase (Ph1, Ph2, Ph3, Ph4, Ph5 or Ph6), reviewer's role (supporter, student) and reviewer's gender (female, hidden or male). Besides becoming more familiar with the data, this round served to identify sub-categories within the categories defined under the motivational function of feedback.

It was during this first round that the inter-rater reliability process for the coding system was undertaken, which will be described separately in the following sub-section.

This labour-intensive process concluded with a coding system that was employed for the analysis of the nature of feedback in the case at hand (see Figure 4.1). The coding system representing the types of feedback identified in the MOOC is presented in Chapter 4, Figure 4.1.

Second round

After having undergone an inter-rater reliability check, the second round of codification started. Here, TUs were reviewed and coded into the categories identified for the coding system. Thereby, it was possible to still correct the range of the TUs, where required. It is to note that the focus here was placed on the TUs preventing possible bias when analysing content from female or male participants.

In order to maintain consistency when coding the TUs, and given the large sample, a set of guidelines was prepared with extended descriptions of the categories, various

types of codified examples, and a summary of key topics resulting from the discussions during the inter-rater reliability test. Some examples of this set of guidelines are presented below. Details include names of the codes that were used to categorise TUs, therefore, it is recommended to the reader of this document to become familiarised with the coding system.

To start with, it was considered that content that did not contribute to the formative function of feedback, or did not appear to refer to any of the defined aspects or functions, was to be categorised as *Not applicable*. Most examples that fit in this category contained short phrases or single words. Examples of this were: “nothing to say”, “cero comments”, “you can do it”, “now you can sit and relax”, “I am speechless”, “thank you for the nice work”.

As soon as some short phrases mentioned keywords that were contained within the assessment criteria or the description of the suggested assignment, these were categorised differently. In cases where features of the work were positively highlighted, then the TU was firstly categorised as *Praise*. However, this category (and all others from the motivational dimension of feedback) presented further sub-divisions: contextualised or decontextualised. Whilst the former referred to the content or presentation of the work or a part of it, the latter was kept general or lacked specification. Examples of contextualised TUs were: “outstanding analysis, great keyword comparison”, “helpful table”, “clear and logical MorphoBox”, “good structured proposal”. Examples of decontextualised TUs were: “easy to read”, “very well written”, “great text”, “fantastic work”, “nicely done overall”, “very thorough work”.

TUs referring to the work itself may be categorised within the aspects or the cognitive function of feedback. When coding TUs that possibly suited the aspect Content, then a differentiation between the three criteria (*Relevance*, *Substance, and Clarity and coherence*) needed to be made.

Coding TUs into *Clarity and coherence* required specific attention. As agreed in the discussion with the evaluators, feedback needed to be accepted as it was provided, without questioning it. However, the definition provided for the course implied a meaning relating to content, and from the many examples found in the data set, the criterion seemed to have been interpreted as relating to the structure of the work (e.g. clear and concise/coherent, well structured). And in that case, those TUs were coded as *Praise – decontextualised*. Therefore, in order for TUs to classify for *Clarity and coherence*, they required more elaboration that helped evaluators to identify them (e.g. "The ideas presented in the analysis are well connected and are in tune with the case").

Some TUs appeared to suit the cognitive function of feedback; however, they did not. For example, the TU "the analysis of the parameters was very detailed, but also confusing" was categorised under *Mitigating language – content* (motivational function). Although the content of this TU does not directly suggest revising the work or a part of it, it may direct students who received the feedback to revise it. Nevertheless, in order for a TU to fall under the category *Revision*, then it needed to offer a direct suggestion that could enrich the work content-wise.

The set of guidelines also comprised fictive examples based on real topics and examples. Variations from the same topic were applied throughout the categories. The examples were especially helpful at a later stage as a quick reference, once the definitions of each of the categories had been interiorised. These examples can be found in Appendix Three.

Third round

This last round was planned in order to address the second, third and fourth levels of analysis. Using the second dataset that regarded the badges, participants who had earned the Popular Evaluator badge were identified. Their IDs were coded in the TUs that had been identified for them from the first dataset.

3.5.1.1 Inter-rater reliability process

As previously mentioned, the process of inter-rater reliability started during the first round of codification and included various steps.

The first step of the process was to search for and elect two external evaluators with experience in qualitative research and online education, who were familiar with the method of content analysis. The first evaluator was the supervisor of this study, who is a lecturer at Lancaster University (England) and her research interests focus on online higher education theories and practices. The second evaluator was a senior lecturer at the Autonomous University of Barcelona (Spain) and a consultant instructor at the Open University of Catalonia (Spain). Her research interests focus on teaching and learning strategies.

Based on the discussion with the first evaluator, the sample used for piloting the coding system was randomly selected, and included 30 messages from the fourth learning phase. Considering the three assessment criteria (*Relevance*, *Substance, and Clarity and coherence*), 10 messages were selected for each criterion.

In the second step, a guideline for the evaluators was created and data were prepared for its codification.

The guideline was created to brief evaluators on three topics: contextual information of the case, presentation of the coding system, and training on the use of the coding system. The contextual information included the name of the learning phase, its learning objective, the description of the suggested assignment, the keywords that were likely to be encountered in the messages when coding, and finally the type of output expected from the assignment. This information would help evaluators make sense of the messages they were going to codify.

The coding system was presented with descriptions and examples for each category. Finally, the last section of the guideline explained how codification was expected to be carried out. By means of varied types of examples and explanations justifying its codification, all categories were covered.

In a parallel way, data were prepared with the intention of providing all evaluators with the same starting condition. All 30 messages comprised single thought units or ideas, and segmentation in TUs was not necessary. Both guideline and data were shared with the evaluators.

The third step of the process was carried out after each evaluator had followed the guideline and codified the suggested sample. Synchronous discussions were planned with each evaluator individually to compare the codification and understand the rationale behind it.

In the meeting with the *first evaluator*, the codification agreement was first compared. The agreement rate was of 8 TUs out of 30. It became apparent that descriptions differentiating between the categories within the content aspect and those in the cognitive function were not clear. Apparently, the sample consisted mostly of categories within those unclear categories. Also, it became evident that a new category needed to be included. During the meeting, a flow diagram was created in order to better support evaluators during the codification process. The diagram was tested directly during the meeting with the same sample of 30 TUs, and a common understanding on the meaning of each of the categories was reached for each of the TUs.

As a next step, it was agreed to randomly select a new sample of messages from the same learning phase, in order to pilot the coding system. A new set of data and the flow diagram were sent to both evaluators (refer to Appendix Four for the flow diagram).

The first meeting with the *second evaluator* focused on discussing the information in the flow diagram and its mode of use. The results for the codification for the first dataset were compared, and the sample was used to test the flow diagram. Although the agreement rate after comparing our initial coding was of 9 TUs out of 30, going

through the sample with the flow diagram and discussing the categories was productive, and a shared understanding of the intention of each category was constructed.

The diversity in terms of elaboration of feedback was discussed and its effect when coding. Similarly, and building on the understanding that feedback has a formative function, agreement was reached regarding the characteristics of the TUs to be coded into any of the categories defined under the cognitive function.

The results from the discussions from the separate meetings were shared with both evaluators, to assure a common ground when coding. A second meeting was planned to be held after all evaluators had coded the second sample.

A week after the first meeting, the second online meeting with each of the evaluators took place. The focus of the meeting was to compare the agreement on the codified data. The percentage agreement with both evaluators was higher than the first time around. With the *first evaluator* the agreement was of 70% (21 TUs out of 30), with the *second evaluator* it was 73% (22 TUs out of 30), and between both evaluators the agreement rate was 76% (23 TUs out of 30). TUs in which a consensus was not clear were discussed until reaching a common agreement. The percentage agreement that was reached after discussion was 80%. This value was considered to be high enough to progress with the study.

3.5.2 Second level of analysis: Describing the evolution of feedback in terms of quantity and of quality

Whilst reporting on the evolution of peer-feedback quantitatively could be undertaken by providing the overview of the number of TUs coded per learning phase for both areas, aspects and functions, reporting on the evolution of quality required careful consideration. The latter reported on the development of peer-feedback as perceived by the participants.

The available data for the analysed sample of 235 participants was not extensive and therefore presented an important limitation. It only provided information on participants earning the badges (IDs), the number of badges they had received and the number of messages they had written. However, it was not possible to trace information referring to the exact messages that were awarded with badges or even the learning phase in which these were awarded. In other words, although all participants of this sample had earned a Popular Evaluator badge during the course, not all messages they had written were awarded with this type of badge. Thus, a criterion defining quality needed to be established. It was evident that quality was not measured by the highest number of badges earned by a participant but by the correlation between the number of badges earned and the quantity of messages that a participant had written during the course. With this information in mind, a two-step procedure was planned and undertaken to define the required quality criterion. The result of the procedure provides a set of participants who consistently engaged in giving peer-feedback and whose messages were perceived by their peers as having average quality and above. The procedure is described below.

The first step was to take the number of messages written by each participant per learning phase and to calculate the total number of participants with the same number of messages and make those values comparable.

For this:

1. The number of messages written per participant were arranged in ascending order.
2. The number of participants who wrote the same amount of messages was counted.
3. The values were plotted in a graphic.

The plot provided a clear overview of the distribution of the number of messages written by the participants. At the same time, it was possible to observe the area in which most of the population lay. (see Figure 3.3)[3]

[3] Axis X has been shortened for presentation purposes. Seven more participants have not been displayed. These wrote between 62 and 192 messages.

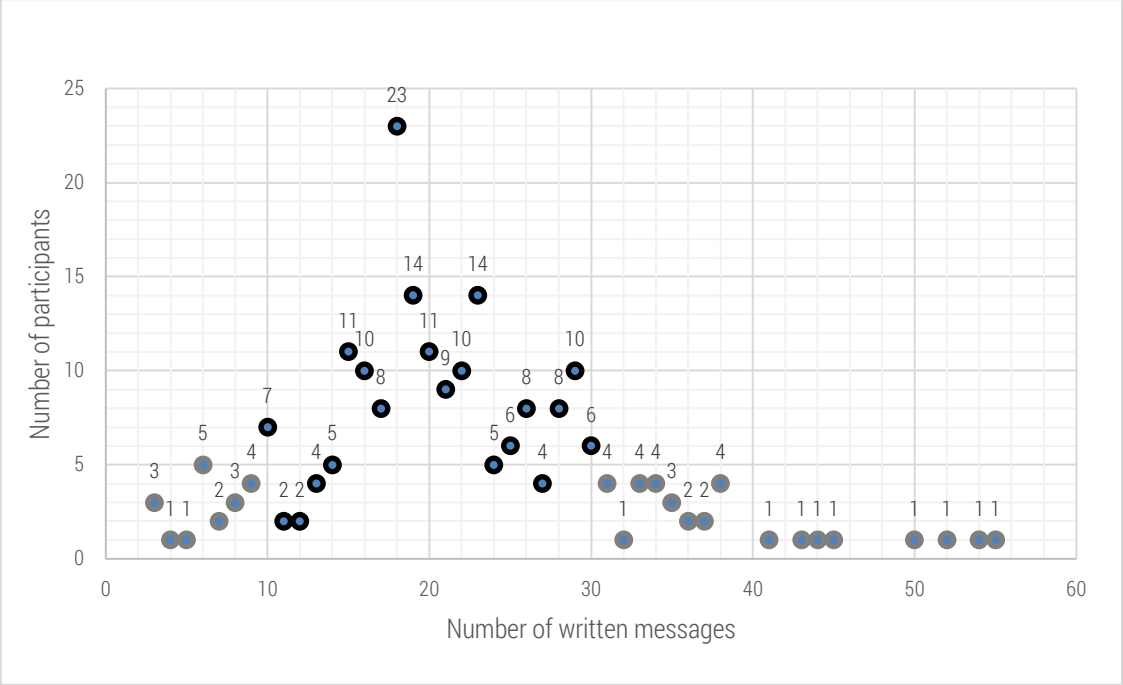


Figure 3.3 Number of messages distributed over the sample (n=235). Axis X displays the number of written messages and axis Y the number of participants writing the messages. The darker area represents the ranges of reference taken for the first part of the procedure described for the definition of the quality criterion.

The ranges of reference taken for the analysis were between 10-7 and 30-6; where 10 were the total number of messages written per person during the course by 7 participants, and 30 the total number of messages written per person during the course by 6 participants. The number of participants included within this range was 177 (75% of the sample). Every value within these two references was defined as suitable for the second step.

The second step was to define the quality criterion that would allow observing the types of feedback employed by those participants fulfilling it. Considering only the values within the aforementioned range, then:

1. The average share of a badge per message was measured. The number of Popular Evaluator badges received by participant was divided by the number of messages written by each participant. The resulting value was converted to a percentage.
2. The average value was calculated.

From this procedure, the average value was of 11.23%, which was the value defining the quality criterion. All numbers included within this percentage and above it correspond to the participants whose feedback was of interest and therefore analysed. A total of 63 participants (27% of the sample) met the quality criterion. On average, each participant wrote 20 messages during the course and earned 3.74 Popular Evaluator badges. The characteristics of these participants can be appreciated in Table 3.9, which shows that 29 participants assumed the role of student and 34 that of supporter. Within the student role, 18 of them were female, 1 hidden and 10 male. Within the supporter role, 29 of them were female and 5 male.

Table 3.9 Details on participants who met the quality criterion.

Role	Gender			Total
	female	hidden	male	
Student	18	1	10	29
Supporter	29	0	5	34
Total	47	1	15	63

The evolution of peer-feedback in terms of quality as perceived by the participants will be reported based on the types of feedback used by these participants that showed the highest average use during the course.

3.5.3 Third and fourth level of analysis: Reporting on differences in the use of feedback in relation to reviewers' role and in relation to reviewers' gender in their respective role

The main aim of these two analyses was to understand and describe the differences in the use of feedback types in relation to the reviewer's role and the gender of the reviewer in the role. Thus, statistical analyses were conducted in order to address this.

To understand the choices made for this analysis, it is worth explaining that: each participant (n = 235) provided feedback during the course but not necessarily in each of the learning phases; and that not all feedback types were used by the participants in each of the phases. In many cases they were used 0 times. This said, the provision of feedback was not binary (they provided feedback or they did not provide feedback), but rather overall ranged from 0 to 357 times. All the data –in this particular case count data– followed a Poisson distribution and therefore a logistic regression model with Poisson distribution was chosen as the best method of analysis. For this, the GLIMMIX procedure from SAS (version 9.4) was used.

The variable inputs included the type of feedback provided. Feedback was also grouped into its two different dimensions: aspects and functions, and their respective sub-groups.

Special consideration was given to ensure that clustering that existed in the feedback provided was accounted for. Specifically, this meant factoring in possible bias in the model response due to the participation of users in one or more of the phases of study. In Chapter 4, the frequency of response least square means ($\pm$ SEM) of the

variable studied is presented. Post-hoc comparisons were calculated based on these means using the procedure PLM of SAS, based on the GLIMMIX model output. The association between these two categorical values was studied using Odds Ratio (OR). Thus, comparison results are presented as Odds Ratio (OR) (95% Confidence Interval) and are accompanied by the respective *p* value. Statistical significance was set at 0.05.

3.6 Critical reflection on the methodology and methods employed

This study claims to adhere to the constructivist inquiry paradigm. The paradigm presents three basic requirements (Lincoln & Guba, 1985), which are considered to be fulfilled in this study. First, this study has been purposefully planned, for which a research proposal was created and accepted. Second, a digital research audit trail was created for the study as a means of maintaining transparent documentation throughout the different stages of the research. Third, the research study was carried out consistently with the ontological, epistemological and the five axiological assumptions, as suggested by Lincoln and Guba (1985), and which are exemplified next. The course was studied in its natural context, and examined from a holistic perspective. On the one hand, the researcher was familiar with the way in which the courses were designed and run. This insider insight allowed the researcher to focus on the observation of the phenomena of interest of the course and immerse herself in the development of the course. On the other hand, as a course participant, the researcher was in the same conditions as all other participants. This implied the need of being involved –even as an observer– in the topics treated in the course and in different spaces in order to be able to contribute to the course in many forms (e.g. providing feedback), if desired. The role performed enabled the researcher to engage in the culture of the analysed course and get insights that would not have been possible otherwise.

Various forms of data were collected to complement and corroborate observations, while at the same time capturing different perspectives. Findings are represented by participants' contributions during the course that reflect the social reality that was

created and experienced there. The study of an online learning event as it developed in real practice was chosen to contribute to the understanding and analysis of a real situation where peer-feedback is used as a method of assessment despite the unique characteristics of such an online course. Besides contributing from a theoretical and methodological point of view, conclusions emergent from the case study may prove useful for the enhancement of future MOOCs, where approaches of assessment of this kind are adopted or planned to be adopted.

This study continues the exploration of a previous study undertaken by the researcher in which the types of peer-feedback in a similar setting were studied. Therein, working hypotheses were created that were used as a basis for this research.

The third and fourth phases of analysis of the study used the quantitative output of the analysed data to report on possible differences in the use of types of peer-feedback in relation to reviewer's role and to reviewer's gender within their role. However, adhering to the understanding that 'causal linkages' cannot be identified (Axiom 4 - Lincoln & Guba, 1985), the results are used to describe possible differences as occurred in the specific studied event, without intending to draw conclusions about the causes and effects of the variables analysed.

Finally, it can be said that the research is value-resonant, as the problem of investigation, the context in which the problem was studied, the theoretical perspective (sociocultural theory) under which the problem was analysed, together with the research paradigm, are all connected.

The content analysis method was used as the principal method for the analysis of data. Besides its clear suitability for the purpose of the study, it is a naturalistic method that supports interpretive approaches. The content of text data is interpreted subjectively; however, it follows a systematic classification process of coding (Hsieh & Shannon, 2005). Qualitative content analysis was employed for the main research question, where the purpose was to identify the nature of feedback in the course. To achieve this, emergent feedback types along with examples that would reflect the social world as constructed in the course, are presented. The quantitative output of the qualitative approach was employed for the second, third and fourth phases of analysis, in which the focus was placed on the description of the frequencies of use of the different types of feedback in relation to the variables of interest for the study.

Although content analysis has been described as a difficult, frustrating and time-consuming method (e.g. Rourke, Anderson, Garrison, & Archer, 2001), it is also a gratifying method in which informed decisions are made based on the results of long and rigorous processes. The amount of data was rich and abundant and the time for its analysis was long and pauses of up to two weeks were made. However, in order to maintain, in the best possible way, a constant quality of analysis, the following procedure was undertaken: after the time of pause and before taking up new data to analyse, a randomly selected segment was re-coded. Both codified versions were compared and when inconsistencies were found, then the descriptions within the created guideline would be revised until reaching the same understanding as when they were coded in the first version.

The quality of the study is discussed according to criteria suggested for the constructivist paradigm, which can be summarised in three items: correspondence between research problem and research design, demonstration of trustworthiness (rigour), and the contribution of the study's results to the community (Lincoln & Guba, 1985).

As stated in the first section of this chapter, the focus of study, the context in which the study was carried out together with the constructivist inquiry paradigm, are congruent.

Rigour (trustworthiness and authenticity) can be achieved when observing the quality criteria that have been defined for the qualitative approach. Creswell (1998) proposed the implementation of at least two of eight techniques. These techniques go in line with the four criteria presented by Lincoln and Guba (1985), namely: credibility, transferability, dependability and conformability. There were six techniques implemented in this study: prolonged engagement, triangulation, peer-debriefing, thick description, the compilation and maintenance of a research audit trail, and a reflexive journal.

Finally, different groups of the community can benefit from the results of this study. The contributions that are claimed to be made with this study have been specified in Chapter 1, sub-section 1.4.

Chapter 4 Findings

This section presents the findings for the four levels of analysis undertaken for the study, and it is respectively divided in four parts. The first part considers the first analysis; it provides an overview of the number of thematic units (TUs) that were studied and categorised and presents the nature of feedback identified within the analysed course. The second, third and fourth parts regard the second, third and fourth analyses in which feedback was studied from different angles. Therein the development of feedback throughout the course in respect to quantity and quality; the nature of feedback in relation to reviewers' role; and finally, the nature of feedback in relation to reviewers' gender in their respective role are reported.

4.1 First level of analysis (Phase 2) - Identifying the nature of feedback in the MOOC

The nature of feedback in the course "Managing the Arts: Marketing for Cultural Organizations" was analysed in the written feedback provided voluntarily by 824 students and supporters during all six learning phases of the course. Specific assessment criteria and guiding questions assisted participants when undertaking assessment. The analysis was carried out employing the previously presented coding system which consisted of twenty categories.

Different types of feedback were identified from the body of content. These were grouped into two main dimensions: aspects and functions of feedback. The aspects of feedback refer to content and presentation, and the functions of feedback to motivation and cognition (see Figure 4.1).

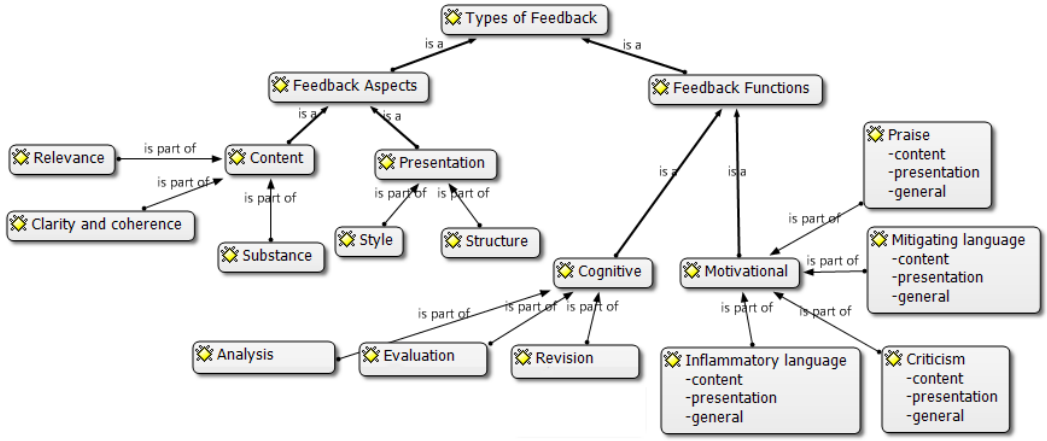


Figure 4.1 Coding system for the identification of types of feedback in a MOOC.

An overview of the number of TUs identified for all four subgroups is presented in Table 4.1 and described next.

Table 4.1 Number of Thematic Units (TUs) per learning phase.

	Aspects		Functions			
Learning phase	Content	Presentation	Motivational	Cognitive	NA	Total
1	522	275	4415	2482	550	8244
2	763	222	4023	2051	410	7469
3	480	155	3258	1214	334	5441
4	605	188	2756	1005	364	4918
5	616	148	2771	1383	439	5357
6	720	142	2775	1212	409	5258
Total	3706	1130	19998	9347	2506	36687

Note: Numbers in columns 2-7 represent TUs.

A total of 36,687 TUs were identified from the data. From this total, 4,836 TUs referred to feedback aspects, 29,345 TUs to feedback functions, and 2,506 TUs to the category *Not applicable* (NA).

The number of TUs identified under feedback functions was six times higher than the number of TUs identified for feedback aspects. Within the feedback functions, the

number of TUs regarding the motivational dimension was double the number referring to the cognitive dimension. Within the feedback aspects, the number of TUs identified as referring to the content of the work was triple the number of TUs identified for the work's presentation. Also, the number of TUs that were classified as NA was double the number of the TUs identified for the presentation aspect of feedback.

In the first learning phase of the course, the number of TUs coded for all areas of feedback aspects and functions, except NA, was higher than in the rest of the phases. In phases three and four, the number of TUs decreased, except for the content aspect and NA, where it increased. The values for the fifth learning phase in all feedback types, except for presentation, increased in comparison to the last two previous ones, and for some of the types the value increased more in the last learning phase. The learning phases that show a noticeable difference in the number of codified TUs are phases two and three; especially noteworthy are those identified under feedback functions.

The types of peer-feedback that were identified in the present course within both dimensions of feedback, aspects and functions, are presented and exemplified next.

Aspects and functions of feedback

Content aspects

The three assessment criteria defined for the analysed course focused on the content aspects, and they were: *Relevance*, *Substance, and Clarity and coherence.*

These three were the first types of feedback that were identified in the analysed content.

The descriptions of these types of feedback are based on the original descriptions provided in the course.

The code used to identify the type of feedback is displayed between parentheses "( )" next to the name.

Relevance (rel)

This is concerned with valid, well-grounded insights presented in the work. It highlights the relevance of an idea of the work for the case under study.

Example

> "Your analysis is substantially relevant since it builds a very clear vision of the CCA case! it provides a wide understanding of their values and their challenges!"

Substance (sub)

This considers whether the content of the work is meaningful in the sense that it has deeply engaged with the learning materials in an attempt to offer solutions, answer questions and solve problems.

Example

> "Part 5 offers loads of brilliant concrete and practical ideas for CCA to consider. Well done!"

Clarity and Coherence (cla)

This considers whether viewpoints presented in the work are clearly formulated and ideas connect in a coherent manner.

Example

> "The plan is very good explained, all the steps are connected to each other, it's clear to perceive it as a whole."

Within the analysed content, the presence of other types of feedback was revealed. These referred to the presentation aspects and to the motivational and cognitive functions of feedback.

Presentation aspects

In this course, participants provided feedback on the structure and style of the assessed work to highlight specific features of its presentation or make suggestions to improve the work presentation-wise.

Structure (str)

It offers comments or suggestions on the presentation of ideas, order of sections of the work, and points out missing components such as: resources list, literature, etc.

Example

> "You may should have changed the order of the steps or the general approach, since you already talk about the HAUBUS in your second step although you're only clarifying it in your third step."

Style (s)

It offers comments or suggestions on the outer form of the work such as: writing and visual style, language used, format of presentation.

Example

> "In terms of formatting, the complementary information (approach) I feel should have been put as attachment instead of left in the text, it does seem out of place there."

Within the analysed content, various comments aimed at showing agreement, disagreement or both with the assessed work or parts of it.

Motivational function

Within this dimension, subcategories of three types were identified: two that related to the work in terms of *content* and *presentation*, and one of *general* nature without a specific focus.

Praise (pr)

It positively highlights the work or specific features of it.

Examples

Content: “It's a good analysis and I really like how you introduced your words from the lexicon in the essay”

Presentation: “I liked your format, because you start with analysis and then give the key terms”

General: “well started with high degree of passion for details”

Mitigating Language (ml)

It compliments but also criticises (or vice versa) the work or parts of it. It uses specific words or signs (e.g. emoticons) to diminish the effect of negative comments.

Examples

Content: “The terms chosen are relevant but very wide and basic”

Presentation*:* “Very well structured, although maybe some graphics would have been great”

General: “good start but I think it is still unfinished”

Inflammatory Language (il)

It refers to comments that read as harsh, aggressive, sarcastic or ironic, and are not constructive.

Examples

Content: “The strategy is merely a rough idea”

Presentation: “The graphic is not only unnecessary but also poorly done”

General: “it seems more like a brainstorm than an actual submission”

Criticism (cr)

It negatively highlights the work or specific features of it.

Examples

Content: “This is not at all case scenario oriented. You don’t mention one time the CCA Lagos”

Presentation: “Unfortunately, your text is a little chaotic and hard to read”

General*:* “A bit unclear and random”

Furthermore, content in the analysed messages showed types of feedback reflecting deeper engagement with the evaluated work. Comments of this type, for instance, revised the work, discussed or challenged the ideas or positions presented in it in order to contribute to its improvement content-wise.

Cognitive function

Analysis (ana)

This includes comments discussing or summarising ideas presented in the work with the aim of understanding it. Inspired by this understanding, reviewers may reflect on their own practice or experience. Some comments search for an interaction, as they sound eliciting, wanting to continue a discussion.

Example

> "It is interesting to see how you considered the absence of a competition a strength, as other teams proposed a strategy for the BACC based on the creation of a competitive system which would enable the Centre to work more on some of its issues."

Evaluation (eva)

It revises the work or parts of it, thereby identifying missing arguments that may affect the consistency of the work. It questions the work or parts of it by offering critical comments, and challenges the authors to defend or justify the ideas presented in the work.

Example

> "You talked about "remodelling the image": to what extent would that be compatible with the goal not to lose the faithful audience, the ones that gather around Trafo's original identity?"

Revision (rev)

It identifies potential parts of the work that can be improved content-wise. Ideas, suggestions or advice are directly or indirectly offered for that purpose.

> "Maybe such a detailed analysis could have focused on a selection of the most relevant categories and variables, making the document easier to understand. This would have also allowed the reader to establish a clearer correlation between the matrix and the strategy proposal."

Comments that were off-topic and did not relate to the work itself in any way were categorised as *Not applicable* (NA). Examples of this are: "I really don't have any suggestions", "sorry - no helpful tips from my side", "I know that you can make it", "best wishes for the next submission".

Since this last category clearly does not meet the objectives of formative assessment as discussed in the previous chapters, it will be disregarded in the next sections when describing the results of the analyses.

4.2 Second level of analysis (Phase 3) - Describing the evolution of feedback in terms of quantity and of quality

This second part regards the feedback provided by a specific number of participants (n=235) who had earned the Popular Evaluator badge. The number of TUs identified in this group make up practically 70% of the total TUs that were identified from the population of 824 participants. Information from the analysed sample is displayed in Table 4.2.

Table 4.2 Number of participants with gender and roles (n=235).

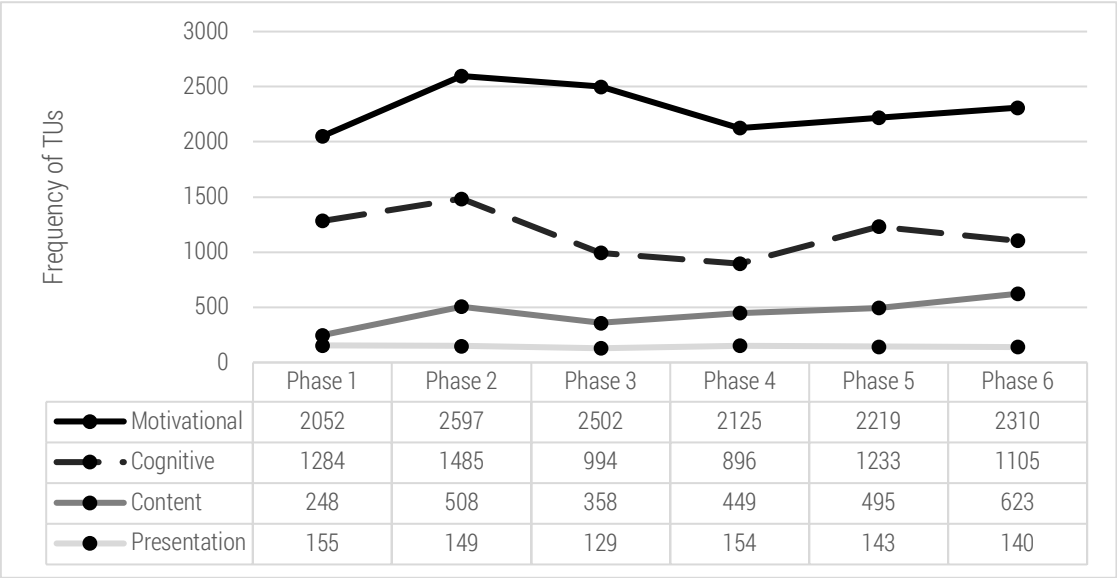
Role	Gender			Total
	female	hidden	male	
Student	104	13	30	147
Supporter	70	4	14	88
Total	174	17	44	235

The number of participants performing the student role was greater than those performing the role of supporter. In terms of gender, the number of female participants was ten times greater than participants who identified themselves as hidden and four times higher than male participants.

Below, the evolution of peer-feedback in terms of quantity will be addressed, followed by its evolution in terms of quality.

4.2.1 The evolution of peer-feedback in terms of quantity

In terms of quantity, Figure 4.2 shows an overview of the total number of TUs that were identified for feedback aspects (*content* and *presentation*), and functions (*motivational* and *cognitive*), in each learning phase. A total of 24,353 TUs were codified into the categories previously presented. Note that category NA is not included in this overview or other description, as explained previously.



	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Motivational	2052	2597	2502	2125	2219	2310
Cognitive	1284	1485	994	896	1233	1105
Content	248	508	358	449	495	623
Presentation	155	149	129	154	143	140

Figure 4.2 Evolution of feedback per learning phase.

In summary, the types of feedback within the motivational function of feedback were identified the most, followed by those within the cognitive function, and lastly by the types referring to feedback aspects. Within the aspects, types of feedback relating to content appeared with more frequency than those for presentation.

Having a closer look into the results under the motivational function, it can be said that the highest number of TUs was reached in phase 2 of the course with 2,597 TUs, and the lowest in learning phase 1 with 2,052. The learning phase with the second highest number of TUs was number 3, followed by phases 6 and 5 with 2,502, 2,310, and 2,219 TUs, respectively.

The frequency of appearance of the types of feedback within the cognitive function was highest in phase 2 with 1,485 TUs, followed by phases 1 and 5 with 1,284 and 1,233 TUs, respectively. Phase 4 displayed the lowest number of TUs identified for this area with 896 TUs.

In regard to the content aspects, phase 6 displayed the highest number during the course with 623 TUs, followed by phases 2 and 5 with 508 and 495 TUs, respectively. Phase 1 was the one in which the number of appearance was the lowest, with 248 TUs. As for the presentation aspect, it can be said that the number of TUs were considerably constant, oscillating between 155 and 129 TUs across the phases.

It is interesting to note that given the starting criteria for assessment which intendedly aimed at guiding feedback to be specifically related to content aspects, TUs referring to other types of feedback were identified and the frequency of appearance greatly surpassed this specific aspect. In the next sub-section, a closer look will be taken to describe the frequency of appearance, ranked from highest to lowest, of each type of feedback within the functions and aspects. The order will be the following: first, the motivational function, next cognitive function, then content aspects, and finally, presentation aspects.

4.2.1.1 Feedback functions

The frequency of TUs identified under the motivational function of feedback was two times higher than those identified under the cognitive function of feedback.

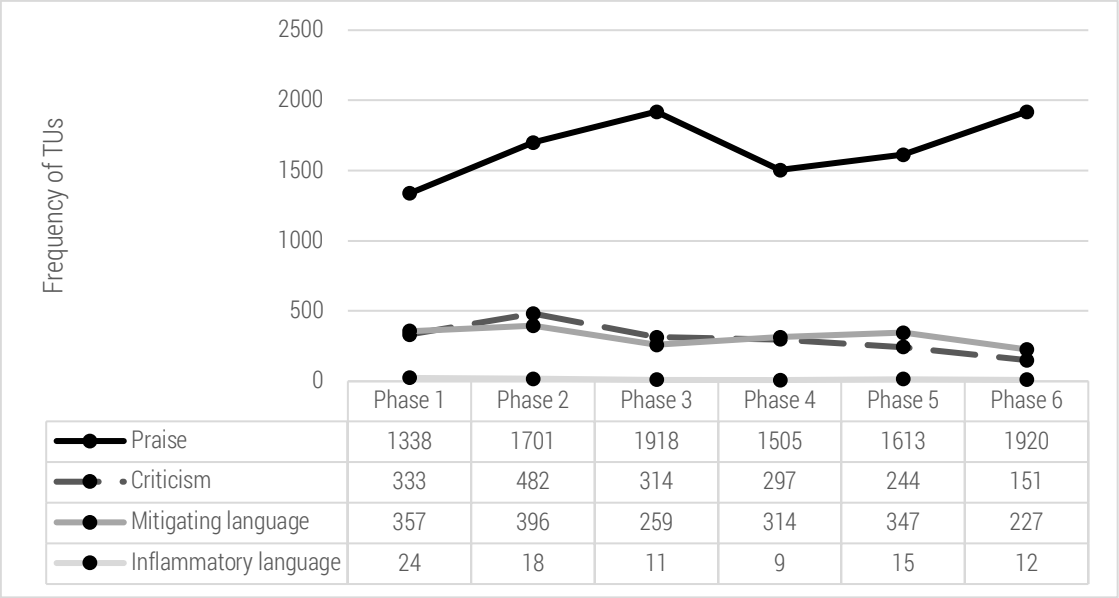
Motivational function of feedback

Four types of feedback were identified within the motivational function. Considering the frequency of TUs, the types are, in descending order: *Praise*, *Criticism*, *Mitigating language* and *Inflammatory language*. Figure 4.3 summarises the number of TUs identified in all three subdivisions corresponding to content, presentation and general. The figure suggests that *Praise* was the type of feedback with most TUs, and

Inflammatory language with least. The types of feedback *Criticism* and *Mitigating language* showed a comparable number of TUs.

Praise had the highest number of TUs in phases 6 and 3, and less in phase 1. The values were 1,920, 1,918, and 1,338 TUs, respectively.

The highest numbers of TUs identified for *Criticism* and *Mitigating language* were in phase 2, and the lowest in phase 6. In phase 2 the numbers were 482 and 396 TUs, whilst in phase 6 there were 151 and 227 TUs. The frequency of appearance of *Inflammatory language* was minimal compared to the other types of feedback. Phase 1 presented the highest value, with 24 TUs, whilst phase 4 had the fewest, with 9 TUs.



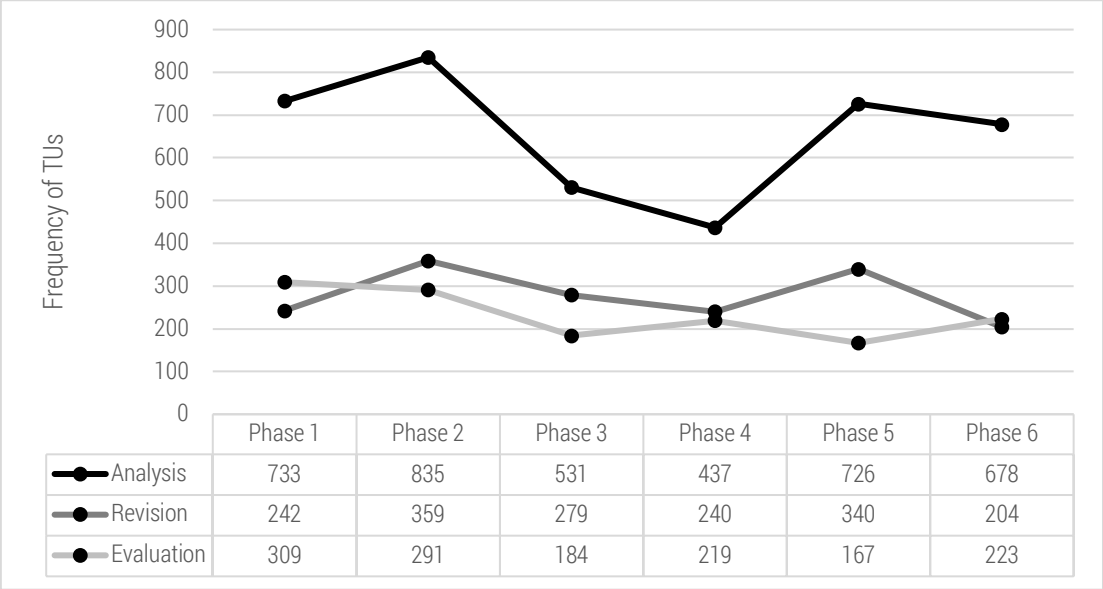
	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Praise	1338	1701	1918	1505	1613	1920
Criticism	333	482	314	297	244	151
Mitigating language	357	396	259	314	347	227
Inflammatory language	24	18	11	9	15	12

Figure 4.3 Feedback types within motivational function.

Cognitive function

Three types of feedback were identified under the cognitive function: *Analysis*, *Revision* and *Evaluation* (see Figure 4.4). In summary, 3,940 TUs were identified for *Analysis*, 1,664 for *Revision* and 1,393 for *Evaluation*. For *Analysis*, the highest number of TUs appeared in phase 2 (835 TUs), and the lowest number appeared in phase 4 (437 TUs). The number of TUs for the feedback types *Revision* and *Evaluation* oscillated comparably. The frequency of values for the former oscillated between 204 and 359 TUs, and the latter between 167 and 309 TUs. *Revision* had most TUs in phases 2 and 5 with 359 and 340, respectively, and fewer TUs in phase 6 with 204. *Evaluation* reached its highest value of TUs in phase 1 with 309 and fewer in phase 5 with 167.

From the figure (Figure 4.4), it can be appreciated that whilst *Analysis* and *Revision* follow a similar curve pattern throughout the phases, the pattern followed by *Evaluation* is in most of the phases inverted. In other words, where *Analysis* and *Revision* dip, *Evaluation* usually rises, and vice versa.



	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Analysis	733	835	531	437	726	678
Revision	242	359	279	240	340	204
Evaluation	309	291	184	219	167	223

Figure 4.4 Feedback types within Cognitive Function.

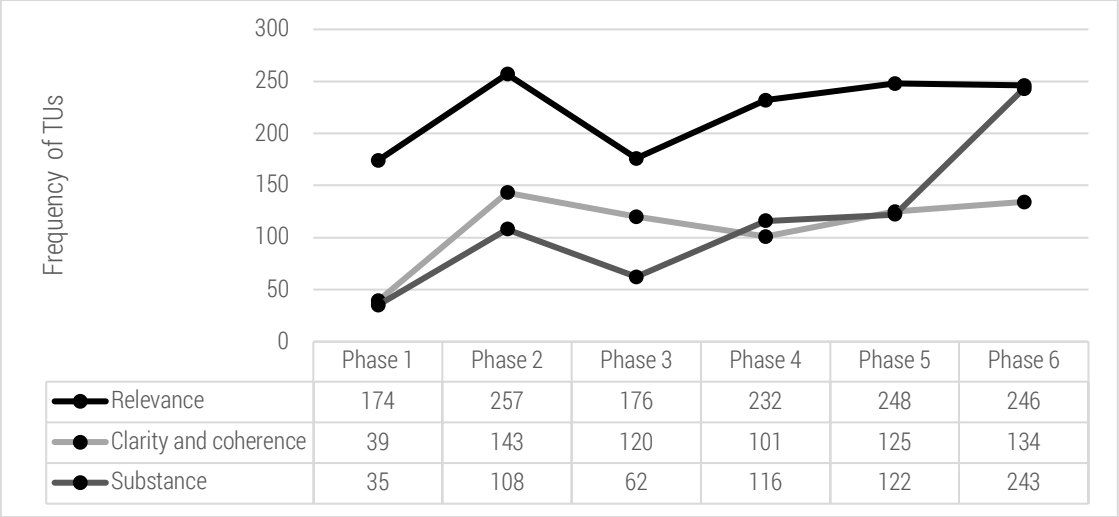
4.2.1.2 Feedback aspects

The frequency of TUs identified under the feedback aspects was in general lower than that under feedback functions. The use of the specific types of feedback for content and presentation are described next.

Content aspects

There are three aspects related to content. Taking into account the frequency of TUs, the aspect types in a descending order are: *Relevance*, *Clarity and coherence*, and *Substance*. As mentioned in the previous sub-section, these were the starting criteria offered by the course instructors for assessing participants' work. From the criteria, *Relevance* was used more than *Clarity and coherence*, and *Substance* (see Figure 4.5). The phases in which all these were referred to the most were phases 2, 5 and 6. The frequency of TUs in these phases was, respectively, 257, 248 and 246 for *Relevance*, 143, 125 and 134 for *Clarity and coherence*; and 108, 122 and 243 TUs for

Substance. The phase in which these aspects were referred to the least was in phase 1. *Relevance* had 174 TUs, *Clarity and coherence* 39, and *Substance* 35. The graphic clearly shows that a similar curve pattern is followed by all types of feedback throughout the phases. However, a significant jump can be noticed for *Substance* from learning phase 5 to 6, nearly reaching the same value as *Relevance*.



	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Relevance	174	257	176	232	248	246
Clarity and coherence	39	143	120	101	125	134
Substance	35	108	62	116	122	243

Figure 4.5 Feedback types within Content aspects.

Presentation aspects

There are two aspects related to presentation: *Style* and *Structure* (Figure 4.6). In comparison to the frequency of appearance of all types of feedback within the functions and aspects of feedback, the frequency of those addressing presentation was the lowest one. *Style* was used at a constant rate after phase 2, reaching a maximum of 90 TUs in phase 6. The frequency of appearance of *Structure* was highest in phase 1 with 100 TUs, and lowest in phase 3 with 48 TUs.

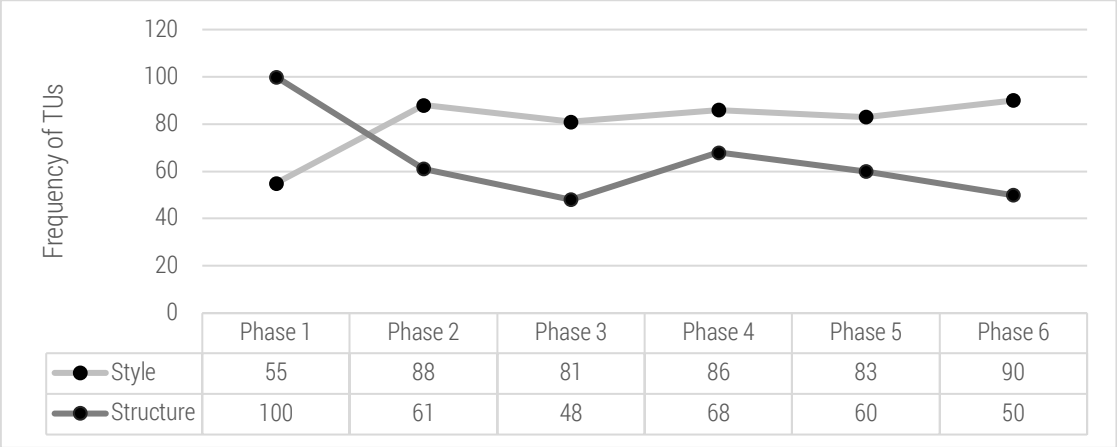


Figure 4.6 Feedback types within Presentation aspects.

After describing the evolution of feedback throughout the course quantitatively, the evolution of it in terms of quality will be described next.

4.2.2 The evolution of feedback in terms of quality

The evolution of the quality of feedback is regarded and presented from two perspectives. The first perspective considers the types of feedback that the literature has defined as fulfilling the objectives of formative assessment, and which are therefore of quality. The second perspective reflects participants' perception of quality, referring to the types of feedback that were perceived as useful. Considering the available data, the description of quality for both perspectives appeared to be more accurate, easier for the reader to visualize and understand if presented in a quantified manner. A qualitative description for each perspective would have required access to data that was not available, a precise recreation of the different happenings throughout the learning phases, and a confident interpretation of the associations between these. This was not only complex –as the course dynamics manifested in different spaces within the learning platform–, but also not possible.

Whilst the first perspective provided concrete types of feedback defining quality based on the type of task, the second perspective was subjective to the participants who took part in the course. For this reason, the quality of feedback as perceived from the second perspective was considered to be directly linked to the types of feedback found in the messages awarded with the Popular Evaluator badge. The process undertaken to address this second perspective required the definition of a quality criterion that 63 out of the 235 participants met. The description of the process can be reviewed in sub-section 3.4.2.

As discussed in the literature review chapter, peer-feedback fulfils cognitive, metacognitive and motivational functions (Narciss, 2008; Narciss & Huth, 2004b; Nicol & Macfarlane-Dick, 2006). For the first perspective, and following the understanding of formative quality of feedback from current studies (e.g. Alvarez et al., 2012; Guasch & Espasa, 2015; Guasch et al., 2010, 2013), the most meaningful types of feedback are those that question the work or parts of it, and those offering ideas on what could be improved. The types of feedback implied by these studies refer to the types *Evaluation* and *Revision* as identified in the first analysis.

For the second perspective, the types of feedback employed by 63 participants were studied. Like any message submitted by any other participant, the feedback provided may have been divided into different TUs. The average of the frequency of use of each type of feedback during the course is shown in Table 4.3. The types of feedback with the highest average of use are highlighted there and are used to report on the evolution of peer-feedback in terms of quality as perceived by the participants.

Table 4.3 Frequency of use means for participants meeting the criterion of quality.

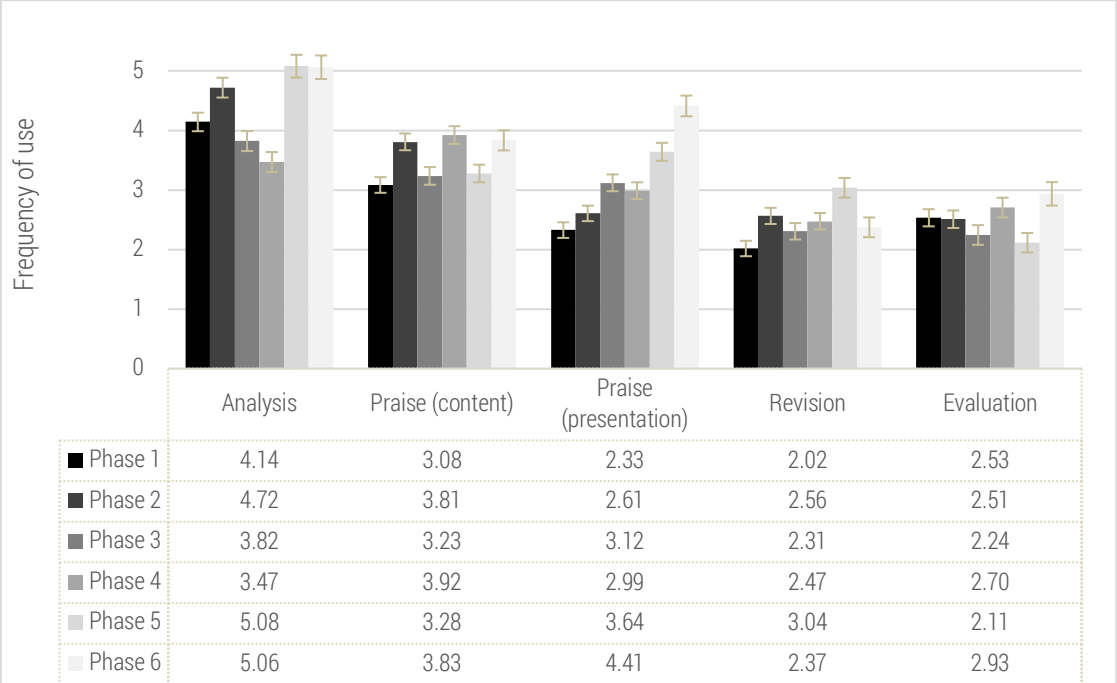
	Mean (SEM)
Feedback functions	
Cognitive	
Analysis	4.01 (0.19)
Revision	1.59 (0.09)
Evaluation	1.48 (0.09)
Motivational	
Praise	
content	2.66 (0.13)
general	0.94 (0.07)
presentation	2.09 (0.11)
Mitigating language	
content	0.80 (0.06)
general	0.13 (0.02)
presentation	0.44 (0.04)
Criticism	
content	0.85 (0.06)
general	0.09 (0.02)
presentation	0.51 (0.04)
Inflammatory language	
content	0.00 (0.18)
general	0.01 (0.39)
presentation	0.00 (0.23)
Feedback aspects	
Content	
Relevance	0.86 (0.06)
Clarity and coherence	0.52 (0.04)
Substance	0.30 (0.04)
Presentation	
Style	0.39 (0.03)
Structure	0.33 (0.03)

Note: Standard Error of the Mean (SEM) is indicated in parentheses after the means. *n*= 63 (female 47, male 15, hidden 1). The highlighted areas indicate the types of feedback with the highest average use.

From the information in table 4.3, it can be seen that the average use of the different types of feedback is divided into two groups: the types of feedback that were used with a frequency higher than one time, and the types below that average.

On average, each of the studied 63 participants wrote 20 messages and there were five types of feedback that they used more than once, in the following order: *Analysis*, *Praise - content*, *Praise - presentation*, *Revision*, and *Evaluation*. *Analysis* was used with an average of 4.01 times, followed by *Praise - content* (2.66 times), *Praise – presentation* (2.09), *Revision* (1.59) and *Evaluation* (1.48). The messages that contained these five types of feedback (as single TUs or combined TUs) are considered to have been most likely awarded with the Popular Evaluator badge. Thus, these will be used to report on the evolution of feedback quality from participants' perspectives.

For both perspectives, all identified types of feedback referring to the cognitive function of feedback were regarded as quality feedback; however, participants also perceived *Praise*, from the motivational function of feedback, as useful. Having identified these five types of feedback, their use and evolution will be closely observed and described considering the requirements of each of the learning phases. Figure 4.7 displays the average use of the five types of feedback by the participants in the study.



	Analysis	Praise (content)	Praise (presentation)	Revision	Evaluation
Phase 1	4.14	3.08	2.33	2.02	2.53
Phase 2	4.72	3.81	2.61	2.56	2.51
Phase 3	3.82	3.23	3.12	2.31	2.24
Phase 4	3.47	3.92	2.99	2.47	2.70
Phase 5	5.08	3.28	3.64	3.04	2.11
Phase 6	5.06	3.83	4.41	2.37	2.93

Figure 4.7 Average frequency of use (LS means ± SEM) for feedback types per learning phase (n=235).

To better understand the use of certain types of feedback, the learning context is briefly explained. The overarching type of task in the course was the case study. Participants in teams had the possibility to work in one out of four real-life case scenarios from cultural organizations around the globe and to explore the respective organizational cultures and their art practices. An overview of the learning tasks suggested for each of the learning phases is presented in Table 4.4.

Table 4.4 Learning phases in the MOOC "Managing the Arts: Marketing for Cultural Organizations"

Learning Phase	Name	Short Description	Dates
1	The cultural economy: Markets and marketing for arts organisations	Keywords: Marketing, Crisis/Change, Form, Intention, Structure, Resources, Relevance Assignment: written analysis of the conceptual framework/values of the respective arts institution	26 Feb - 12 Mar
2	Mapping the terrain:	Keywords: Cultural Market, Infrastructure,	12 - 26

	Art organisations and cultural institutions in context	Relationships, Networks, Interdependency, Affiliation, Adversity Assignment: map of institutional/relational environment using morphological box; positioning strategy sketch	Mar
3	Enabling creativity: Marketing cultural projects and managing artistic process	Keywords: project management, artistic/curatorial practices, process analysis, communications/ PR Assignment: process-centred evaluation of a cultural project considering four dimensions of sustainability	26 Mar - 16 Apr
4	Reaching across the fourth wall: Building audience relationships	Keywords: publics, audiences, diversity, outreach, community Assignment: audience analysis and strategies for appropriate audience building/development	16 – 30 April
5	Emerging identities: Co-creating and shaping digital brands	Keywords: co-creation, sharing, distributed and behavioural branding, incentives, identity-formation Assignment: value analysis & mission statement considering digital media for brand positioning and identity	30 Apr - 14 May
6[4]	Curating strategies: Artistic practices and sustainability	Keywords: arts markets, problem solving, creativity, scope, innovation management, failure, strategy Assignment: develop a project plan for integrated marketing campaign addressing challenge to organization	14 – 28 May

Note: An introductory week was planned for participants to become familiar with the learning community, the course concept and the learning platform.

In summary, the demand of the tasks was cognitively high and the complexity of them increased after each learning phase. Each task was built around a specific topic

[4] After the completion of the course an independent jury selected three team finalists and awarded one of them with a winning design. The winning team had the opportunity to visit personally the cultural organization of their focus and present their final assignments on-site.

and an associated perspective on managing arts organizations and marketing cultural projects. Besides the requirements posed in each of the learning assignments, the learning phases did not differ pedagogically. Once a learning phase started, all learning materials were made accessible in the learning platform. For the start of each phase the *MOOC Facilitator* published a message on the learning platform with a short and motivating introduction to the phase and its goals. Participants would then have access to a comprehensive assignment description (its learning goals and all requirements), to video-keynotes and literature. A forum thread was opened for each learning phase for any comments or questions. The assignment outputs were mainly written documents containing analyses, descriptions and proposals, among others. Some assignments lent themselves to the inclusion of other type of media. Whilst the length of the work had specific limitations, the presentation of the information did not. Each assignment followed different steps, and these needed to be documented by the students. Because of the different steps in each assignment, the feedback provided either referred to a specific step, some steps, or the work as a whole. The products submitted at the end of each phase were evaluated by mentors and could be evaluated by any other participant from the learning community. It is important to note that although teams could improve their work for their next submissions, a dialogue between assessors and assesses to discuss or clarify the feedback received did not take place. Technically, an option for answering or replying to a feedback was not provided by the learning platform.

Next, the learning goals of the tasks proposed for each learning phase and the required steps are presented[5], followed by a quantitative description of the average frequency of use of the types of feedback that represented quality in this analysis.

Phase 1

Learning goal: Participants debate different conceptual definitions of basic terminologies and gain understanding of contemporary discourses in strategic cultural marketing. Challenges facing arts organizations and cultural managers are introduced.

The task consisted of three steps:

Step 1: Make a list of the most relevant 10-12 key terms emerging from all materials.

Step 2: Develop a lexicon of definitions for 8-10 of the key terms.

Step 3: Using this lexicon, create a short value analysis of your case organization.

In this phase *Analysis* was used the most, and *Praise - content* the second most, with an average of 4.14 times for the former, and of 3.08 for the latter. *Evaluation* was used with a higher average than *Praise - presentation* and *Revision*, which were used 2.53, 2.33 and 2.02 times, respectively.

[5] The information is presented here as it was used during the course.

Phase 2

Learning goal: Participants learn to analyse an organization within relationships of the cultural market and understand how they sustain, develop and limit the radius of the institution's activities. They debate positioning an arts organization within the ever-changing environment under conditions of the cultural market and within a corresponding policy framework.

This phase also consisted of three steps:

Step 1: Create a morphological box (12-25 categories) to capture and classify the features of cultural organizations.

Step 2: Map your case organization's features within the morphological box[6] and identify strengths, potentials, risks.

Step 3: Create a concrete strategy proposal for strategic growth and development, no more than 500 words in length.

All types of feedback, except for *Evaluation*, were used more than in the first phase. *Analysis* was the most frequently used, with 4.72 times, *Praise - content* was second with 3.81, and *Praise - presentation* was third with 2.61 times. In comparison to the previous phase, the use of *Revision* increased to 2.56 times, and the use of *Evaluation* remained almost equal with 2.51 times.

[6] This is a tool that enables creative thinking for generating solutions to complex problems.

Phase 3

Learning goal: Participants focus on processes and practices pertaining to individual cultural projects. They learn about opportunities and challenges of enabling artistic processes and marketing cultural projects in process perspective as well as strategies for the creation of relevance by representing and responding to changes in their environment.

This phase consisted of two steps:

Step 1: create a model for project management to address specific aspects of decision-making (if creating a visual representation of the model with accompanying notes then use max. 1 page, if using best-practice protocols then use max. 400 words).

Step 2: Write a sustainability analysis considering the four dimensions of environmental, economic, social, and cultural factors (max 400 words).

Here, the average use of *Analysis* and *Praise - content* although still highest, decreased to 3.82 and 3.23 times respectively. Conversely, the use of *Praise - presentation* rose to 3.12 times, in comparison to the previous phases. The average use of *Revision* decreased to 2.31 times, in comparison to the previous phase, and the use of *Evaluation* was lower than the previous two phases with 2.24 times.

Phase 4

Learning goal: Introducing distinction of actual/desired audiences and the need to learn as much as possible about them as basis for segmentation and development

strategies. Learn about low-cost digital options for audience interaction and relationship-building and multi-tiered approaches to communicate with current and future publics.

The steps were the following:

Step 1: Create an Action Plan

Step two: Describe it in comprehensive programmes

Step three: Evaluate your Action Plan

Step four: Identify your audience profile

Step five: Manage your audience

There were five steps for solving the assignment. The expected final output was a written document containing an action plan, an evaluation matrix and a reflection addressing specific questions.

In this phase, the average use of *Praise - content* was higher than the rest of the types of feedback and the highest in all phases with 3.92 times. *Analysis* was used with less frequency (3.47 times). In comparison to the previous phase, the use of *Praise - presentation* decreased to 2.99 times, and the use of *Revision* and *Evaluation* increased to 2.47 and 2.70 times, respectively.

Phase 5

Learning goal: Brand creation and management within network culture, use of distributed branding and co-creation with external audiences/stakeholders and internal staff. Reflection on remixed usage of proprietary content, digital replication, peer-to-peer sharing, and limits for controlling independent brand initiatives while managing repercussions.

The task consisted of three steps:

Step 1: Synthesize the case organization's values and distil a value framework.

Step 2: Consider the use of digital content and strategies for capturing and shaping an organization's identity.

Step 3: Create a proposal for digital content & strategies to be used in brand positioning.

The average of use of *Analysis* increased, and so did *Praise - presentation* and *Revision*, with 5.08, 3.64 and 3.04 times, respectively. The values for *Analysis* and *Revision* were the highest reached during the whole course. The use of *Praise - content* and *Evaluation* decreased, and the average use was 3.28 and 2.11 times, respectively.

Phase 6

Learning goal: Questions of innovation, creativity, feasibility and sustainable engineering of institutions through design thinking to develop integrated marketing

approach. Identify and interpret resources and ideas underlying everyday practices into a coherent epistemology while thinking about (implicit) vision, stated mission and scale, impact and significance.

The task consisted of five steps:

Step 1: Identify a strategic goal with impact on the case organizations marketing practices (objective).

Step 2: Develop and define parameters of the goal (validation).

Step 3: Select and define the elements of your plan (operationalization).

Step 4: Choose your audience (formatting and structure).

Step 5: Proposing through practice (execution).

Here, all types of feedback were used with a higher frequency, except for *Revision*. The use of *Analysis* was the highest and its frequency of use remained almost the same as the previous phase with 5.06. *Praise - presentation* was used second most and it also reached the highest average compared to all previous learning phases, with 4.41 times. The average use of *Praise - content* increased to 3.83 times and the use of *Evaluation* to 2.93 times. *Evaluation* reached its highest average use in this last phase. As for *Revision*, the average of use decreased in comparison to the previous phase, and it was used 2.37 times.

Findings suggest that participants' perception of useful feedback concur with those types of feedback that have been defined by the literature as most meaningful. The

main requirement that needs to be fulfilled by the message containing feedback is to be directly related to the assessed product. Thus, another type of feedback that showed to be useful to participants was *Analysis*, which aims at understanding the text by either discussing or summarising ideas found in the work. In fact, this type reported the highest average use amongst all other types. Moreover, participants considered that when their work or certain features of it were positively highlighted, either by making a specific reference to its content or its presentation, then the feedback was also considered useful.

The development of the types of feedback throughout the different learning phases suggest that the type of peer-feedback acknowledged as useful, and in participant's perception as qualitative, may depend on different factors. Some factors could be related to the type of assignment that was required for each phase, the complexity and length (in documented steps) of the assignment, and most probably to the enjoyment or tension experienced by the participants during each phase while working on the assignments. For instance, in phases where the tasks became more complex an important value could have been placed on those types of feedback regarding the motivational dimension over the cognitive. An accurate recreation of important happenings during the course may have provided information that would have helped to interpret the frequencies of use of certain types of feedback.

The next section explores the types of feedback that were employed by participants in their roles as student or supporter. Thereby, differences in the use of types of feedback will be highlighted and described.

4.3 Third level of analysis (Phase 4) - Differences in the use of feedback between students and supporters

After having an overview of the development of peer-feedback in all six learning phases of the course quantitatively and qualitatively, the types of feedback used by the same sample of 235 participants (role student *n*= 147, role supporter *n*= 88) will be presented. Hereby, the differences in the use of feedback types that may be related to the role of the reviewer were investigated and are reported. First, an overview of the frequency of use means of the types of feedback grouped in functions and aspects is provided, and initial differences in feedback use are identified. Second, a detailed description of the frequency of use means for each type of feedback within those groups is presented and described. Finally, statistical differences in the use of specific types of feedback in relation to reviewer's role are reported.

Figure 4.8 displays the mean frequency of use of feedback functions and aspects in each of the roles and in all learning phases of the course.

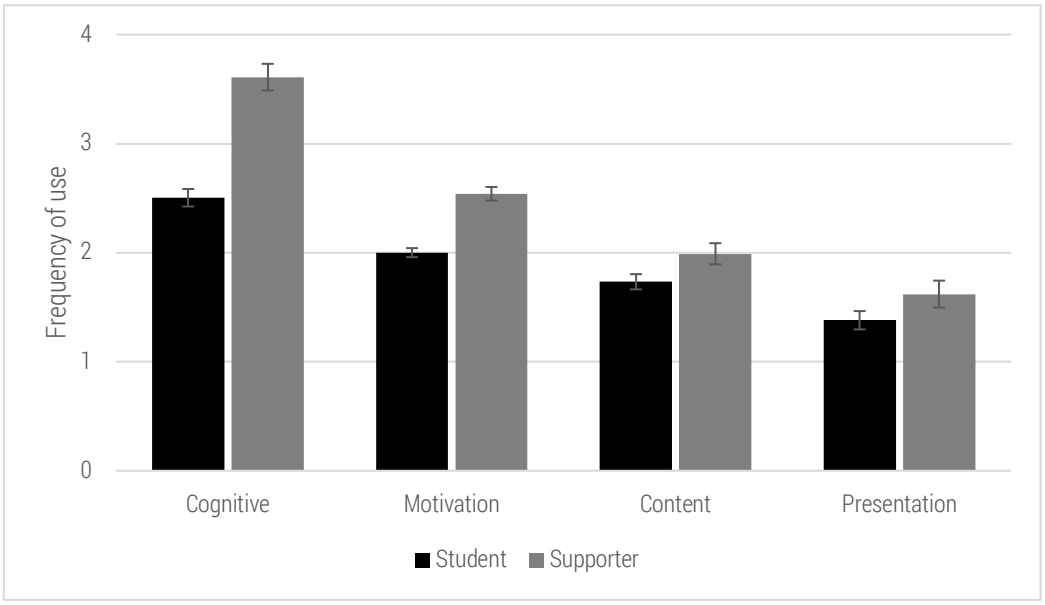


Figure 4.8 Average frequency of use (LS means ± SEM) for feedback functions and aspects in all phases - Roles: Student and Supporter (n=235).

In general terms, participants performing the role of supporter addressed the functions and aspects of feedback with more frequency than those performing the role of student. The highest mean value was displayed in the cognitive function of feedback for both roles. Students' mean value was of 2.5 and supporters' mean value was 3.61. The types of feedback within the motivational function of feedback were used the second most, with a mean value of 2 for students and 2.54 for supporters. As for the aspects of feedback, students referred to content on average 1.73 times, while supporters did so nearly 2 times. The aspect presentation was the least used by both roles. On average, students used types of feedback within this dimension 1.38 times and supporters 1.62 times.

The mean frequency of use between both roles were both higher in the cognitive and motivational functions of feedback and lower in the content and presentation aspects of feedback. Whilst the difference in the frequency of use in the

presentation aspects does not seem relevant, the differences for all other areas do. After running a Poisson regression, a statistically significant difference (p = .005) between students and supporters was established. The latter group reported a higher average of use. Odds ratios (OR) were calculated to understand the association between the differences encountered (see Table 4.5).

Table 4.5 Differences for feedback functions and aspects Student versus Supporter

	Students (n = 147)	Supporters (n = 88)			
Feedback	Frequency of use (± SEM)		p	*OR*	95% *CI OR*
Functions					
Cognitive	2.5 (± 0.08)	3.61 (± 0.12)	<0.001	0.69	0.63 - 0.76
Motivation	2 (± 0.04)	2.54 (± 0.06)	<0.001	0.78	0.73 - 0.83
Aspects					
Content	1.73 (± 0.07)	1.99 (± 0.10)	0.0293	0.87	0.76 - 0.98
Presentation	1.38 (± 0.08)	1.62 (± 0.12)	0.102	0.85	0.70 - 1.03

Note: Significant at the *p<0.05* level. *OR = odds ratio. CI = confidence interval.*

The cognitive function of feedback appeared to be more likely to have been used by supporters 1.44 times, 95% CI [1.31, 1.58] more than students. In a similar way, for the motivational function of feedback, supporters are more likely to use this dimension, 1.28 times, 95% CI [1.2, 1.36] more than students. The aspects related to content also showed a statistical significance. Also here, supporters seem more likely to refer to content aspects, 1.14 times 95% CI [1.02, 1.31] more than the other group of participants.

When zooming in on the frequency of use of the types of feedback by these two types of reviewers, specific types of feedback within the groups functions and aspects that stood out the most become apparent. Figure 4.9 provides an overview

of the frequency of use of all types of feedback employed by students and supporters during the course.

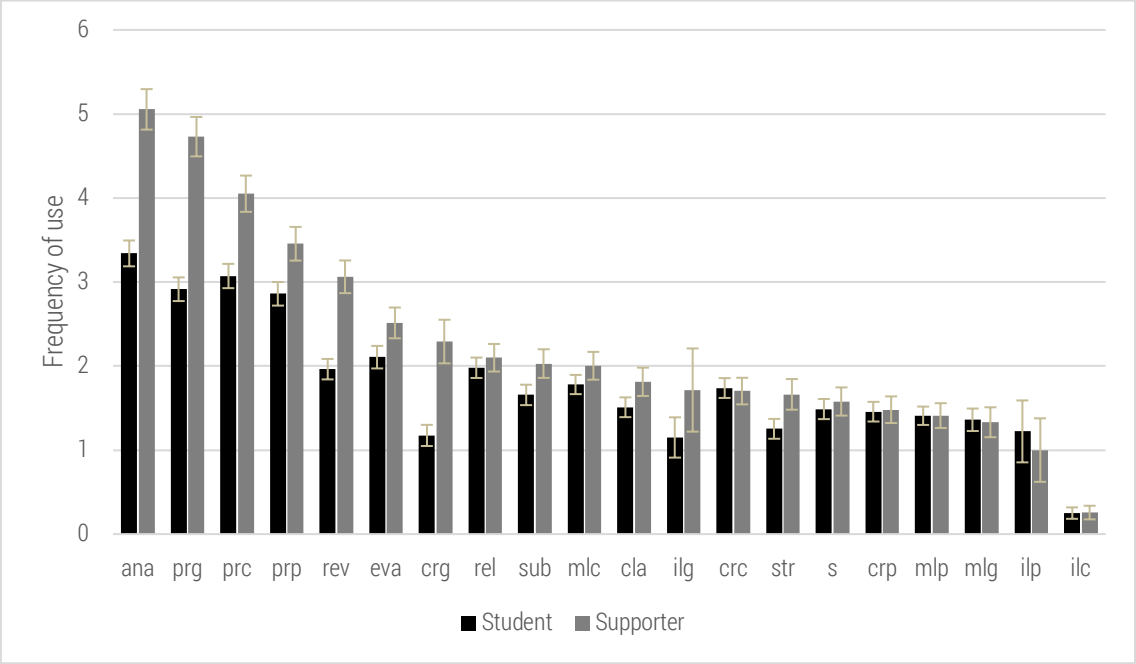


Figure 4.9 Average frequency of use (LS means ± SEM) for feedback types) in all phases - Roles: Student versus Supporter.

On average, the frequencies of use of *Analysis (ana)* and *Praise (pr)* from the cognitive and motivational functions of feedback were the highest among all other types of feedback. These were used the most by both roles. However, participants performing the role of supporter displayed a higher frequency of use than those performing the role of student.

Analysis (ana) from the cognitive function of feedback, for instance, was used 1.5 times more by supporters than by students. The mean value for the former was of 5.05 and of 3.34 for the latter. The second most used type of feedback was *Praise* from the motivational function of feedback in all its sub-types: *general, content* and *presentation (prg, prc, prp)*. Also here, all sub-types were used between 1.2 and 1.6

times more frequently by supporters than by students. While the frequency of use of the three sub-types remained almost constant for students (with a mean value of 2.91 for *Praise – general (prg)*, 3.07 for *Praise – content (prc)*, and 2.86 for *Praise – presentation (prp)*), the frequency of use by supporters varied. For this group, *Praise – general (prg)* reported a mean value of 4.73, *Praise – content (prc)* of 4.05, and *Praise – presentation (prp)* of 3.46.

Within the functions of feedback, the type *Revision (rev)* reported a mean value of 3.06 for supporters and of 1.96 for students. Conversely, the type *Evaluation (eva)* was used with a similar frequency by both roles. The mean value was of 2.51 for supporters and 2.11 for students.

As for the types of feedback within the motivational functions, *Criticism – content (crc)* and *Criticism – presentation (crp)* were used similarly by students and supporters. The mean values for content were of 1.74 and of 1.70 respectively, and the ones for presentation were 1.46 and 1.48. However, the average frequency of use of *Criticism – general (crg)* was higher in supporters than in students, with a mean value of 2.29 and 1.17, respectively.

The types of feedback included under *Inflammatory language (il)* were used similarly by both roles, yet the frequency of use by students was the lowest of all. Whilst the types referring to *Inflammatory language – content (ilc)*, and *Inflammatory language – presentation (ilp)* were used similarly by both roles, the type referring to the *Inflammatory language – general (ilg)* was higher by supporters. These participants reported a mean value of 1.71 and students a value of 1.15.

Within the sub-types of feedback of *Mitigating language (ml)*, the average use of each of them was greatly similar. The mean value for *Mitigating language – content (mlc)* was of 1.78 for students, and of 2 for supporters; the one for *Mitigating language – general (mlg)* of 1.36 and 1.33, respectively. Finally, the mean value for *Mitigating language – presentation (mlp)* was of 1.41 for both students and supporters.

In respect to the aspects of feedback, the frequencies of use of the types regarding content were higher than the types regarding presentation. In the former, *Clarity and coherence (cla)* displayed a mean value of 1.51 by students and of 1.81 by supporters; *Relevance (rel)* 1.98 and 2.10, respectively; and *Substance (sub)* 1.66 and 2.03, respectively.

For presentation, the mean value for the frequency of use of *Style (s)* was 1.49 for supporters and 1.58 for students. For *Structure (str)*, the mean values were 1.25 for students and 1.66 for supporters.

As previously described, significant statistical differences were established for the cognitive and motivational functions, as well as in the content aspects. However, the types of feedback marking the difference were not identified. A Poisson regression was run once again to find differences in the use of types of feedback within those three dimensions between students and supporters. Moreover, odds ratio was calculated to understand the associations (see Appendix Five).

Whilst individual types of feedback from the cognitive and motivational functions of feedback displayed a significant statistical difference (p = .005), the individual use of

the types of feedback from the content aspects did not. In conclusion, five types of feedback reported had statistical significant differences in its use. For the cognitive function of feedback these were *Analysis (ana)* and *Revision (rev)*, and for the motivational function of feedback *Criticism - general*, *Praise - general*, and *Praise – content (prc)*.

Analysis (ana) was more likely to have been used by supporters 1.51 times, 95% CI [1.29, 1.72] more than students, and *Revision (rev)* 1.56 times more, 95% CI [1.31, 1.85].

Criticism - general (crg) was more likely to have been used 1.95 times, 95% CI [1.43, 2.65] more than students, followed by *Praise – general (prg)* 1.62 times, 95% CI [1.41, 1.85], and by *Praise – content (prc)* 1.31 times, 95% CI [1.14, 1.51].

The use of these five types of feedback will be discussed in the following chapter.

The following section explores further the use of the types of feedback, with an emphasis on their gender. It also reports on the differences found.

4.4 Fourth level of analysis (Phase 5) - Differences in the use of feedback between female and male participants in their respective role

This section reports on the last analysis undertaken in this study, which aimed at exploring and describing the differences in the use of feedback types that may be related to the gender of the reviewer in their respective role. The analysis zoomed in on the previous analysis by integrating the variable gender. The number of participants that were studied in this analysis were 218. From this number 48% were female students, 32% female supporters, 14% male students and 6% male

supporters. As mentioned in the presentation of the case, as participants filled their profile for the learning platform with information, three options were provided to define their gender: female, hidden and male. For the purpose of the analysis, only both gender options, female and male have been included. It is worth mentioning that besides participants' roles, no other characteristics regarding their socio-cultural background were recorded or compared. The analysis was purely based on the content of the feedback provided and its meaning. Table 4.6 specifies the sample studied.

Table 4.6 Participants' gender and roles.

Role	Gender		Total
	female	male	
Student	104	30	134
Supporter	70	14	84
Total	174	44	218

The presentation of the findings follows a similar order as the previous analysis. In a first stage, an overview of the frequency of use means of the functions and aspects of feedback by both female and male participants are provided and commented on. In a next step, reviewers' gender in their respective role will be explored and described by reporting on the frequency of use means of the types of feedback grouped in functions and aspects. Finally, statistical differences in the use of specific types of feedback in relation to reviewer's gender within the performed roles will be described.

Information from Figure 4.10 suggests that both genders addressed the functions and aspects of feedback with a similar average. In fact, the mean values for both

genders in both dimensions was practically equivalent. For instance, the mean values in the cognitive function of feedback were 2.88 for female participants and 3.03 for male participants. Furthermore, in the motivational dimension of feedback, the mean value for female participants was 2.25 and 2.12 for male participants. Concerning the aspects of feedback, female participants presented a mean value of 1.91 for *Content*, and males a mean value of 1.64. Lastly, *Presentation* reports the lowest mean values of all. Female participants report a mean value of 1.44 and male participants 1.51. Statistical differences in the frequency of use of types of feedback addressing the functions and aspects of feedback between female and male participants could not be established (see Appendix Six).

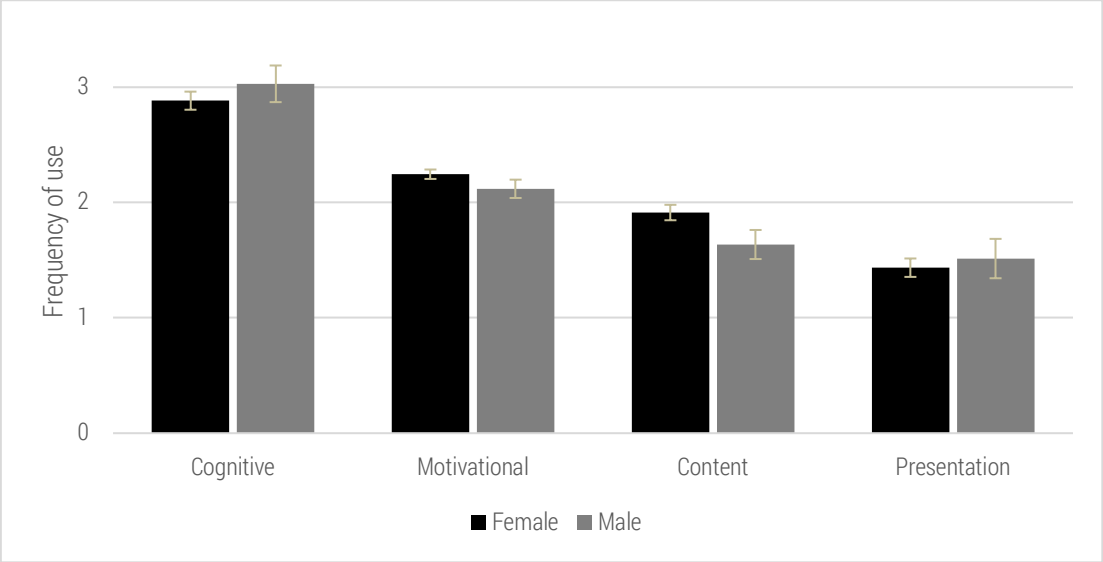


Figure 4.10 Average frequency of use (LS means SEM) for feedback functions and aspects in all phases - Gender: female and male (n=218).

However, when exploring the role of each gender, the frequency of use means for the functions and aspects of feedback show variations (see Figure 4.11).

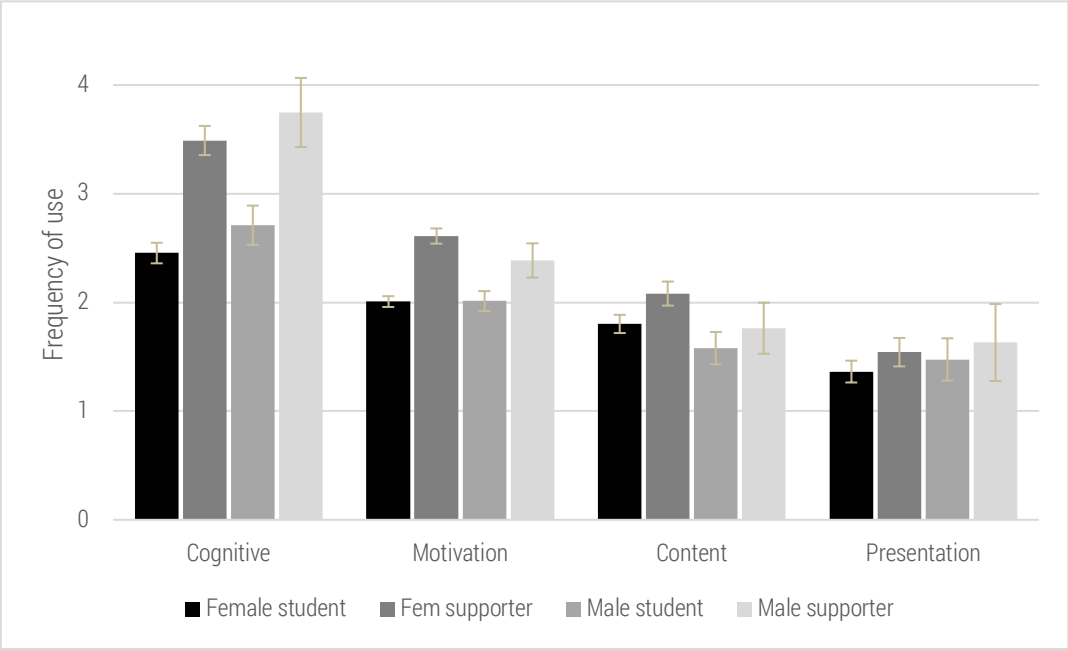


Figure 4.11 Average frequency of use (LS means ± SEM) for feedback functions and aspects in all phases - Gender in roles (n=218).

Overall, the dimensions that were addressed most by the four groups were the cognitive and the motivational ones. In all dimensions, supporters displayed higher

mean values than students. For the cognitive function, male and female supporters had comparable mean values of 3.75 and 3.49, respectively. Also comparable were the mean values reached by female and male students, which were 2.5 and 2.71, respectively.

Likewise, in the motivational function of feedback, the mean values for female and male supporters were comparable, with 2.61 and 2.39 respectively. Female and male students displayed the same mean value of 2.01.

The aspects of feedback were less addressed by participants in both genders and roles; however, the mean values for content were higher than for presentation. In terms of content, female supporters reported the highest mean value of 2.08, followed by female students with 1.80, male supporters with 1.76, and finally male students with 1.58. In terms of presentation, male supporters displayed the highest value with 1.63, closely followed by female supporters with 1.54, male students with 1.48, and female students with 1.36.

From the description and from Figure 4.10, it can be discerned that the differences on the average of use of the presentation aspects are not great between the groups of participants, and with 95% certainty, it can be confirmed that no significant statistical differences were encountered for any of the groups of participants in that area. However, for the cognitive and motivational functions as well as content aspects, significant statistical differences were identified (p = .005) (see Appendix Seven). For instance, in the cognitive function of feedback, it was found that male supporters addressed this dimension 1.56 times more than female students 95% CI

[1.28, 1.85], and 1.38 times more than male students, 95% CI [1.12, 1.72]. Moreover, female supporters were found to be more likely to use the types of feedback within this function of feedback, 1.42 times, 95% CI [1.28, 1.58] more than female students, and 1.28 times, 95% CI [1.1, 1.49] more than male students.

Within the motivational function of feedback, the differences encountered between students and supporters in their respective roles are the following: female supporters were more likely to employ feedback types from this dimension, 1.31 times, 95% CI [1.21, 1.40] more than female students and 1.29 times, 95% CI [1.16, 1.43] more than male students. On the other hand, male supporters were more likely to refer to this area 1.19 times, 95% CI [1.02, 1.38] more than male students and equally 1.19 times more than female students, 95% CI [1.04, 1.36].

As for the content aspects, female supporters are reported to be more likely to address this dimension more times than both female and male students. They report a value of 1.31 times, 95% CI [1.06, 1.63] more than male students, and 1.16 times, 95% CI [1.01, 1.33] more than female students.

Figure 4.12 provides a detailed view of the frequency of use of the different types of feedback by all four groups during all learning phases of the course.

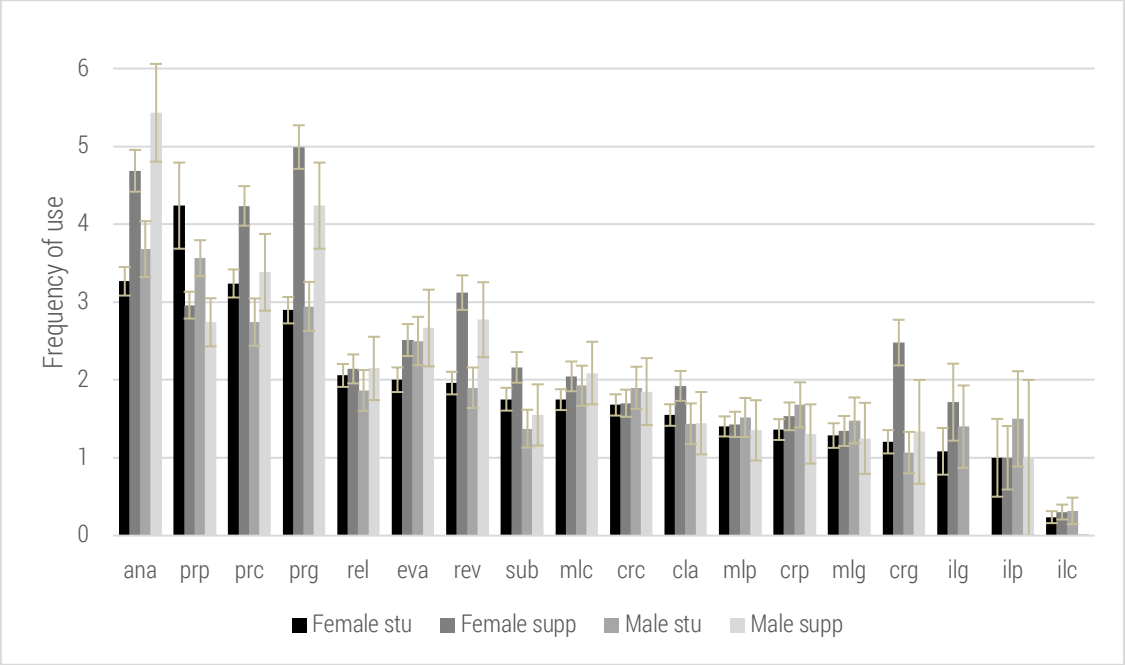


Figure 4.12 Average frequency of use (LS means ± SEM) for feedback types in all phases - Gender in roles (n=218).

As reported in the previous section, on average, the frequencies of use of *Analysis (ana)* and *Praise (pr)* from both cognitive and motivational functions of feedback were the highest among all other types of feedback for both roles. The results of this analysis allow the identification of the gender within the role that reported the frequencies.

Within the group of supporters, males showed the highest average of use of *Analysis (ana)* (5.43 times), followed by females (4.69 times). In the students group, males used *Analysis (ana)* 3.68 times, and females did so 3.27 times.

The average use of this type of feedback by male supporters was the highest from all other types of feedback used by all other groups of participants, and the highest one ever used by this group.

The second type of feedback that was used with a high average was *Praise* in its different sub-types, with *Praise – general (prg)* being the one that was used the most on average by female supporters (4.99 times), and male supporters (4.24 times). Female students and male students used this type of feedback at a similar rate. The former used it 2.9 times and the latter 2.94 times. For female supporters, the average use of this type of feedback was the highest they reached. The second most used type of feedback was *Praise – content (prc)*. Also here, female supporters reported the highest average use, with 4.23 times, male supporters the second most with 3.38, followed by female students with 3.24, and lastly by male students with 2.74 times. Finally, *Praise – presentation (prp)* was the type of feedback that was most used by female students (4.24 times), and second most used by male students (3.56 times). Female and male supporters used it on average comparably (2.96 and 2.74 times, respectively). For the group of female students, the average use of this type of feedback represented the highest within their group.

Evaluation (eva) was used more than twice in each of the groups. Male supporters reported the highest average use with 2.67, followed by female supporters and male students with 2.51 and 2.50 respectively, and lastly by female students with 2. The last type of feedback within the cognitive function is *Revision (rev)*, and its average use was highest among female supporters with 3.12, second highest among male supporters with 2.77; and these were followed by female and male students who used it on average 1.96 and 1.90 respectively.

The types of feedback regarding *Criticism* and *Mitigating language* specific to the sub-type *presentation (crp, mlp)* were used similarly by all groups of participants, an

average of 1.5 times. When feedback of these two types referred to the sub-type *general*, then all groups reached an average use of above one time. However, the group using *Criticism – general (crg)* the most was female supporters with an average of 2.48 times. All groups of participants referred to the content aspect in both types of feedback at least 1.5 times, and female and male supporters used it on average 2 times.

The least used types of feedback were those regarding *Inflammatory language (il)*. Whilst the average use of *Inflammatory language – content (ilc)* was below 0.5 times in all groups, the average of use referring to *Inflammatory language – presentation (ilp)* was 1 for most of the groups, except for male students who used it on average 1.5 times. *Inflammatory language – general (ilg)* was used the most by female supporters and male students, 1.7 times and 1.4 times, respectively, whereas female students reported to use it 1.08 times, and male supporters did not report any use.

As for the aspects of feedback, findings show that although the use of the types of feedback within Content and Presentation were comparable and the differences were minimal, the former was used more than the latter. The average use for *Style (s)* and *Structure (str)* under the aspect Presentation, was 1.56 and 1.42 times respectively. For Content, the average use of *Substance (sub)* was 1.71 times, with female supporters using it an average of 2 times. *Clarity and coherence (cla)* was used on average 1.58 times. Also here, female supporters reported the highest use with nearly 2 times. Finally, the average use of *Relevance (rel)* was highest with 2.05 times. In this case, all groups except male students addressed this type of feedback at least 2 times.

The difference on the averages of use of the types of feedback by the groups seems minimal. However, after comparing the use of each type between all groups, statistically significant differences (p = .005) were encountered in both functions and aspects of feedback in specifically eight types of feedback. In the cognitive function these were: *Analysis (ana)*, *Evaluation (eva)* and *Revision (rev)*; in the motivational function: *Criticism – general (crg)*, and *Praise – content (prc)*, *general (prg)* and *presentation (prp);* lastly, in content aspects: *Substance (sub)*. (see Appendices Eight, Nine and Ten).

In the cognitive function of feedback, it appeared that both female and male supporters addressed the types of feedback *Analysis (ana)* more times than both female and male students. Female supporters were more likely to have addressed *Analysis (ana)* 1.42 times, 95% CI [1.23, 1.62] more than female students, and 1.27 times, 95% CI [1.02, 1.58] more than male students. Male supporters reported slightly higher levels of use, and were more likely to have employed this type of feedback 1.66 times, 95% CI [1.29, 2.12] more than female students, and 1.47 times, 95% CI [1.09, 2] more than male students.

In regard to *Evaluation (eva)*, a statistically significant difference was only found between female supporters and female students. The former used this type of feedback 20% more than the latter OR = 1.25, 95% CI [1, 1.56].

The last type of feedback in this dimension is *Revision (rev)*. For this type, statistically significant differences between female supporters and students in both genders were established. Female supporters were more likely to employ *Revision (rev),* 1.64

times, 95% CI [1.21, 2.22] more than male students, and 1.58 times, 95% CI [1.29, 1.96] more than female students.

In the dimension regarding motivational functions, differences were found in the use of the types *Criticism (cr)* and *Praise (pr)*. Whilst the former only reported differences in its sub-type *general (crg)*, the latter reported differences in all its sub-types: *content*, *general* and *presentation (prc, prg, prp)*.

Criticism - general (crg) was found to have been used more by female supporters than all other groups. They were more likely to employ this type of feedback 2.04 times, 95% CI [1.47, 2.85] more than female students, 1.82 times, 95% CI [1.35, 2.46] more than male students, and 1.62 times, 95% CI [1.17, 2.24] more than male supporters.

The differences in the use of *Praise – content (prc)* were found to be significant between female supporters, and both female and male students. Female supporters were more likely to use this sub-type 1.54 times, 95% CI [1.21, 1.97] more than male students, and 1.31 times, 95% CI [1.12, 1.53] more than female students.

Praise – general (prg) was used more times by supporters than by students. For instance, female supporters were shown to have used it more, 1.72 times, 95% CI [1.47, 2] more than female students and 1.70 times, 95% CI [1.34, 2.14] more than male students. In a similar way, although slightly less times, male supporters appear to have used this sub-type 1.47 times, 95% CI [1.11, 1.92] more than female students, and 1.44 times, 95% CI [1.03, 2] more than male students.

Praise – presentation (prp) appeared to be used more by female supporters when compared to students in both genders. They used it 1.2 times, 95% CI [1.02, 1.42] more than female students, and 1.3 times, 95% CI [1.01, 1.67] more than male students.

Finally, the last type of feedback in which significant statistical differences were found was in feedback aspects, specifically in the area of Content and the use of the feedback type *Substance (sub)*. Herein, female students were shown to have employed it 1.57 times, 95% CI [1.06, 2.33] more than male students.

In conclusion, statistically significant differences were encountered in the use of eight types of feedback that belong to the cognitive and motivational functions, and to the content aspects of feedback. Findings suggest that female supporters used five types of feedback more times than female students. These types were: *Analysis (ana)*, *Evaluation (eva)*, *Revision (rev)*, *Criticism – general (crg)*, and *Praise* in all its sub-types *(prc, prg, prp)*. Furthermore, findings display that the first group used five types of feedback more times than male students. The types were: *Analysis (ana)*, *Revision (rev)*, *Criticism – general (crg)*, *Praise* in all its sub-types (*prc*, *prg*, *prp*), and *Substance (sub)*. However, that same group used only one type of feedback more than male supporters, which was *Criticism – general (crg)*.

Findings also indicate that male supporters used two types of feedback more times than female and male students. These were *Analysis (ana)*, and *Praise – general (prg)*.

The differences encountered in the frequency of use of these eight types of feedback in the different groups in this last analysis are discussed in the following chapter.

Chapter 5 Discussion

This chapter discusses the main findings of the previous analyses. The principal objective of this study was to describe the nature of peer-feedback in the analysed MOOC. In order to meet the research objective, different analyses were undertaken guided by three research questions: how does feedback evolve over time in terms of quantity and quality?; how does the nature of feedback differ between students and supporters?; and finally, how does the nature of feedback differ between female and male participants in their respective roles as students or supporters?

5.1 Summary of results and interpretation

5.1.1 Description of the evolution of feedback in terms of quantity and quality

The evolution of feedback in terms of quantity and quality throughout the course was examined in 235 participants who had earned the Popular Evaluator badge, the badge that acknowledged the quality of the feedback provided. With regard to the evolution of feedback in terms of quantity, a fluctuation of messages was expected based on the assessed type of task. In respect to the evolution of feedback quality, the premise was that independent from the quantity of feedback provided, its quality would tend to increase over time.

According to the description of the tasks, these were different in various aspects: their nature (theory, practice, a combination); the number of steps required to complete it (up to 5); the allowed length (up to 800 words sometimes in one step); their presentation type (text or enhanced with visual media), and their requirements in terms of output (e.g. analysis, proposal, model). Despite these differences, all tasks –but not all steps– were similar in the sense of requiring at least the last two

stages of critical thinking –*independent* and *contextual* (Baxter Magolda, 2007)– and the skills related to it. Thus, the expected types of feedback reflecting this were *probing* and *collaborative* (Lockhart & Ng, 1995) represented by the types of feedback *Evaluation* and *Revision*.

Based on this information, the main findings for the evolution of peer-feedback in terms of quantity and quality are presented next.

Evolution in terms of quantity

In this regard, three findings are presented. First, a fluctuation in the quantity of feedback (identified TUs) throughout the learning phases could be confirmed, but not its clear association to the types of tasks. Second, although a fluctuation in the quantity of feedback could be confirmed, it displayed a positive evolution. Third, the frequency of use of types of feedback regarding the motivational functions –specifically *Praise*– prevailed over those within the cognitive functions, and aspects of feedback.

According to Neubaum et al. (2014), the tasks, the length and the content of the product to be assessed have a direct influence on the quantity and also quality of feedback. However, as per the first finding it appears that all other factors related to this specific learning setting –besides the already known (e.g. the enjoyment or tension of a given learning phase, time available, attitude, external motivations)– cannot be disregarded and need to be brought into consideration when planning such an analysis, anticipating and interpreting results. The study presented by Neubaum et al. (2014) differed from the present one in many respects: the MOOC

was semi-open, participants were studying for a Bachelor's or Master's degree, they received credits after completing the course (all participants performed the student role), they were of a similar age, academic levels and had a common language. Moreover, the study analysed only one out of eleven learning phases, and even though providing elaborated feedback on the work of their peers was not mandatory (as in the present study), participants were required to grade the works. The assessment process was anonymous and not accessible to all.

These differences are highly significant and the conditions in both studies are not, therefore, comparable. Because of the different factors affecting each individual at different stages of the course, a variation in the quantity of feedback independent from the task specifications seems understandable. This opinion is supported by Waite, Mackness, Roberts and Lovegrove (2013). In their study, 'transformative shifts' were reported as experienced by MOOC participants. Throughout the course they realised that their participation in any activity of the course was actually voluntary. This would explain why drop-out and completion rates are ongoing research topics in the MOOC field (e.g. Clow, 2013).

A study carried out by Coffrin, Corrin, de Barba and Kennedy (2014) measured – among others– participation in assessment activities over the different stages of two MOOCs and reported a progressively decreasing tendency. This is contradictory to the second finding of the present study, which displayed a positive tendency. Even if the number of participants who offered feedback may have fluctuated throughout the course, the quantity of feedback provided did not dip. This finding concurs with Waite et al.'s (2013) observation that participants taking part in assessment activities

tended to remain active throughout the MOOC. In fact, at those points where Coffrin et al. (2014) report a decrease in participation, in the present study an increase is displayed. Beyond the different factors that may have affected the quantity and quality of the feedback provided and that were previously mentioned, this positive development is likely associated with the combination of a set of external factors. To begin with, the role played by the different supportive actors in the course, especially that of *the MOOC Facilitator,* seemed crucial. The role of this actor was to keep up the motivation in the learning community throughout the different stages by being attentive and observant to all developments, and by posting encouraging, provocative and interesting messages in a timely manner. Additionally, the motivation of being awarded with badges and the Statement of Accomplishment (SA) could have played a role, as well as the design of the learning tasks and the topics treated therein.

The third finding regards the frequency of use of the type of feedback *Praise* above all other types of feedback, which concur with other studies and learning settings (e.g. Cho et al, 2006; Patchan, Charney & Schunn, 2009, López-Benavides, 2015). Patchan et al. (2009) suggest that the reasons for this finding can be related to three factors: 1) being truly impressed by the assessed work; 2) understanding praise as an important component of feedback; 3) giving praise to receive praise in return. In their study, these authors suggested having evidence for the third factor. However, here it is argued that the principal reason lies on the first factor, and Meek, Blakemore and Marks (2017) provide some evidence supporting this notion. Diversity (e.g. in terms of age, gender, level of education, knowledge and experience,

interests, language and cultural background) was one of the principal characteristics of the MOOC (as discussed in Chapter 1). This element assured a rich and varied spectrum of approaches to the problems that were suggested in the different tasks. Thus, it is believed that *Praise* was used to celebrate the diversity of ideas regarding approaches, proposals, analyses, models and solutions that arose out of similar problem statements, and the diversity of ideas in finding ways for presenting their work in a creative and artistic way. Participants differentiated *Praise* for *content* and for *presentation*. In phases 1 to 4, the former was used with more frequency than the latter, and the latter was used with more frequency in the last two learning phases. Whilst a direct connection to the type of task could not be established, and this is explained in sub-section 5.1.1.1, it is thought that the topic of the fourth task (Co-Creation and Digital Brand Development) and the incentive for the fifth task (winning a prize to travel to their specific case institution and present their ideas on-site) resulted in extra eye-appealing submissions that explain the high frequency of this type of feedback.

Evolution in terms of quality

As concluded in Chapter 1, feedback of quality involves aspects from the cognitive, metacognitive and motivational dimensions (Narciss & Huth, 2004; Nicol & Macfarlane-Dick, 2006). The quality of feedback for this analysis was considered from two perspectives: quality as described by the literature (feedback that helps students to further develop) and quality as perceived by the participants. The types of peer-feedback defining quality were: *Revision*, *Evaluation*, *Analysis*, *Praise - content* and *Praise - presentation*. The two first ones are perceived as markers of

quality by both the literature and participants, and the other three only by participants.

Two important findings can be presented for the evolution of quality of peer-feedback. First, independent from the quantity of peer-feedback provided (identified TUs), the quality of peer-feedback remained constant throughout the course. This evolution is based on the expected types of feedback for the suggested types of tasks: *Revision* and *Evaluation.* Second, the evolution of quality for the other three types of feedback (*Analysis*, *Praise - content*, *Praise -presentation*) shifted respectively for each type.

The first finding contradicts the initial assumption. Apparently, information based on the number of written messages, TUs identified and analysed, and the number of awarded badges in the course, suggested a positive development in the quality of feedback. However, after analysing evidence from different sources, the quality of feedback was consistent. Revisiting Shute (2008), the function of formative assessment is that of enhancing learners' knowledge, skills and comprehension of a topic studied, and the types of feedback suggested to fulfil this function are *Evaluation* and *Revision*. The use of these two requires a great degree of engagement with the assessment criteria and the work itself, as well as a high degree of critical thinking. It also implies that decontextualised feedback does not meet the previously mentioned expectation.

The constant evolution of *Evaluation* and *Revision* verifies two aspects in respect to the characteristics of the tasks and the participants. On the one hand, it confirms the

high cognitive demand required in the type of task used for this course: the case study (Cohen et al., 2000; Falchikov, 2005) and its respective sub-tasks. On the other hand, it shows that participants engaging in the assessment activity displayed drive, initiative, high motivation but specially great confidence (Bouchard, 2009). Confidence, as the foundation for peer-assessment to take place (Falchikov, 2005), suggests that the stages of critical thinking of those who made use of these two types of feedback, stood between *independent* and *contextual* (Baxter Magolda, 2007). This leads us to think that those participants were capable of expressing themselves clearly and precisely (Strijbos et al., 2010) and that language or cultural differences seemed actually not to have been an impediment for offering written feedback. Since decontextualised feedback does not support the function of formative assessment, the value that participants placed on *Analysis*, *Praise - content*, *Praise - presentation*, becomes evident.

The second finding supports partly the initial assumption. Whilst *Analysis* and *Praise - presentation* were the types of feedback showing a positive evolution throughout the course, the use of *Praise - content* remained clearly constant. An explanation for the positive trend in the evolution of *Analysis* can be associated with three factors: first, the learning tasks grew in complexity over time; second, feedback providers may have wanted to avoid the series of misunderstandings that have been reported in the literature and that are directly associated with language proficiency (Liyanagunawardena et al., 2013; Nkuyubwatsi, 2014); and third, participants may have been keen to fulfil the requirements associated with the SA. This third factor

appears to have had a much greater impact in the evolution of *Praise - presentation*, and is therefore presented later on.

Primarily, *Analysis* aims at understanding the assessed work by summarising or discussing its ideas. This type has been identified by Nelson and Schunn (2009) as *Summarization.* The authors highlight its importance as it is a way to share a common understanding of the ideas displayed in the works, and from which both parties can profit (assessors and assessees). In this course, the use of *Analysis* went beyond this description. In some cases, and based on the understanding of the work or parts of it, reviewers reflected on their personal experiences from their practice and shared them with their peers, similar to the findings of Krogstie et al. (2015).

It is important to note that this type of feedback fulfils a cognitive function and a formative purpose primarily for reviewers themselves. Here, their engagement with understanding the work(s) becomes evident, as do other important skills they need to put into practice (Costello & Crane, 2013; Nicol et al., 2014; Strijbos et al., 2010; Tuzi, 2004; van den Berg et al., 2006).

The positive evolution of *Praise - presentation* is explained by two factors. First, by the form of the submitted works; second, by the requirements set to be awarded with the SA. Although a prescription on the way assignments needed to be presented was not specified, it appears that participants gained more confidence throughout the learning phases and that the task requirements left them with room to care for the presentation of their submissions. It is likely that participants interest and possible background in arts are the main reasons for this development.

Participants may have felt urged to create eye-appealing products and reviewers may have felt urged to address that dimension.

Moreover, the requirements for earning the SA were: 1) assess 18 final submissions (equal to 6 Active Evaluator badges); and 2) have earned 3 Popular Evaluator badges throughout the course. This fact offers a possible explanation about the frequency of use of this type of feedback especially in the last phase of the course. However, it does not provide an explanation on why the use of this and not other type of feedback.

As for *Praise - content*, it can be said that its constant use along the learning phases confirms that –in the eyes of the reviewers– the assessed works or their parts met the expectations of the tasks. This finding is thought to provide evidence to support the first factor that explains the high use of *Praise* (Patchan et al., 2009) as a way to truly express admiration for the work, and which was presented previously.

Limitations, practical implications and suggestions for future research

This analysis presents four limitations. First, despite the known fact that each assignment consisted of different steps, this analysis did not distinguish feedback provided for each step principally because a reference was not always clear or specific in the message. In other words, participants were not required to focus their evaluation on a specific part of the work. Having had feedback on specific parts of the work would have helped to accurately identify the types of feedback that were used for each part of a given task. This limitation becomes clear in the following example: if a task consisted of 4 steps, and 1 of them required a visual

representation, it was likely to find feedback related to that particular step than to the rest of the steps. Accordingly, the type of feedback used would most likely regard its presentation rather than its content. As an effect, the types of feedback employed and the frequency of feedback used only reflect the requirements of feedback for steps that cannot be specified. The results of the analysis reflect participants' preferences at the moment of evaluation. One way to remedy this without discouraging participants from evaluating as they wish is to ask participants to mark the area on the assessed document where they would like to write feedback. A similar idea like the 'Annotation system' proposed and implemented in van der Pol et al. (2008) could be helpful. Offering such a system could support and promote more interaction between assessors and assessees, which appears to be lacking in this type of assessment strategy.

Also, although specific assessment criteria were defined for the course and manifold guiding questions were formulated under each criterion to assist participants when providing feedback, they appeared not to conform to it. Thus, two concrete recommendations are proposed here. The first one is to offer teams the possibility of creating own assessment criteria, by allowing them to formulate concrete questions in the areas they feel are of concern in their work. This way, the informal assessment process becomes a more dynamic process for all: assessors are directed to provide feedback that will be welcomed by their peers, and peers will receive a variety of comments to consider. This type of MOOCs has mentors who are responsible for giving clear expectations of the works that need to be submitted and students can rely on them. The second recommendation is that if teams do not wish to create an

own set of assessment criteria, then assessors are free to apply either their own criteria or the suggested ones. Hereby, teams need to be aware of this implication.

The second limitation concerns the Popular Evaluator badges. These were awarded after the ratings received by all participants in the community, and not only by those to whom the feedback was intended. Therefore, although the messages may have been considered of quality by those rating them, it was not clear whether the messages were found useful by the students and their teams in order to improve their work. Although the types of feedback identified as useful in the analysis concur with the ones presented in the literature (e.g. Cho et al., 2006; Ferguson, 2011; Guasch et al., 2013; Lizzio & Wilson, 2008; Tseng & Tsai, 2007; van der Pol et al., 2008), it would have been valuable to have participants' direct opinions. If doing this, and if in a similar context where tasks consist of different steps, it would be valuable if participants could refer to the perception of usefulness according to the type of assignment or specific step in an assignment.

Third, based on the literature and on the results of the analyses undertaken by the researcher of the present study, it has been inferred that those participants who used *Evaluation* and *Revision* wrote clear and precise messages, and also that language or cultural differences seemed not to have been a limiting factor. However, this conclusion did not include the perspective of either the participants who received the written feedback, or those who wrote it. Thus, further research could bear this in mind and consider it.

Fourth, taking into account that feedback type *Analysis* was used beyond the summarisation of ideas contained in a work, an extra categorisation may have been useful. Currently, it is assumed that most participants who used the type of feedback *Analysis* may have used it only to check whether they had understood the ideas correctly, without taking a further step, even if they would have wished to. However, this cannot be inferred by the current analysis. Distinguishing between these two possible correlations between the participants who employed *Analysis*, will enable further understanding of its use and the topics contained in specific tasks or its steps.

5.1.2 Differences in the use of feedback between students and supporters

The roles performed by each type of participant were clearly different. Whilst students were required to fulfil a set of expectations, supporters were not. The obligations for students were to study the learning materials, collaborate effectively within their teams, complete and submit the assignments. Successful completion of this process would result in obtaining a university certificate that represents 5 ECTS (European Credit Transfer System). Supporters, however, were encouraged to study the course materials and familiarise themselves with the assignments. Both types of participants were motivated to assess the work of the teams after each of the learning phases and provide peer-feedback.

Based on the role differences the principal assumption was that those participants who had spent the same amount of time completing the same assignment under the same conditions would provide deep and critical feedback, compared to other learners that did not undergo the same situation. Since the type of task was a case study that included assignments of different degrees of complexity, the expected

types of feedback reflecting deep and critical feedback were *probing* and *collaborative* (Lockhart & Ng, 1995) represented by the categories identified as *Evaluation* and *Revision* –as commented on in the previous section.

Because of what students were required to do in their role, they had to undergo processes of cognitive analysis of each task (Narciss & Huth, 2004). This fact led to think that if students were to provide feedback, this would focus more on the types of feedback regarding the cognitive functions than those types regarding the motivational functions. Conversely, it was expected that supporters' feedback would focus more on the motivational functions of feedback than the cognitive ones. There were no further assumptions regarding the differences between both roles.

Findings revealed three important insights. Firstly, the initial assumption could not be confirmed in its entirety. On the one hand, it was confirmed that students' use of feedback types within the cognitive functions prevailed over those types within the motivational functions. However, the same result could also be confirmed for supporters. Secondly, supporters addressed the functions and aspects of feedback with more frequency than students. Thirdly, supporters differed from students in the use of five types of feedback: *Analysis*, *Praise - general*, *Praise - content*, *Revision* and *Criticism - general*.

The first findings suggest that both roles engaged in providing peer-feedback, which was a voluntary activity. Some authors explain that when participants have the possibility to engage in such a cognitively demanding activity and think like experts, they are able to gain a better understanding of the topics presented (Falchikov,

1986; Roscoe & Chi, 2007; Topping, 1998). When this understanding has been gained, then the experience may be categorised as enriching, encouraging participants to continue engaging in the activity. However, in order to engage with the process at all participants must have a certain level of confidence (Falchikov, 2005). It is assumed that participants' levels of expertise in this or other courses (e.g. level of expertise on the subject matter, on MOOCs, language proficiency, on peer-assessment) may have equipped them with a certain level of confidence that allowed them to undertake different activities within the course. Yet, participants' uneven levels of expertise confirms that peers were actually not true peers (Meek et al., 2017). This said, this finding suggests that students and supporters shared certain commonalities independent of the role performed in the course.

The second and third findings appear to be associated to the circumstances under which participants could join the course. Many participants registered to receive a university certificate but student places were limited to 800. As a result, many interested participants could only join the course as supporters. The association with the findings becomes clearer in the explanations that follow.

The second finding indicates that supporters addressed the functions and aspects of feedback with more frequency than students. This is interpreted based on a specific factor that may have played a key role in the case of supporters: time. According to information in the post-course survey, the average time that was invested by supporters per peer-evaluation was 30 minutes, 10 minutes more than students. This time may have allowed supporters to carefully work on the elaboration of their messages, in which different types of feedback could be identified. The elaboration

in the messages is inferred when noticing that the number of messages written by supporters was lower than the number written by students for each phase of the course.

Having in mind the types of tasks and the length of the works that supporters chose to assess, it only seems logical that good time was required in order to get the most out of this activity. Having access and being exposed to new ideas and projects allowed reviewers to gain new insights that surely supported their own learning and development (Topping, 1998).

The time factor can be associated with the circumstances under which some participants joined as supporters (as explained previously). Those highly motivated participants that were asked to participate as supporters seemed to have invested the time they had planned for the obligations as students on the expected activities that were also foreseen for supporters, such as peer-assessment.

This situation may also explain why mostly supporters were keen on being eligible for the SA. Although this document was not equivalent to any formal credit, it acknowledged the efforts of co-learning and was awarded by both course organisers, the Goethe Institute and the Leuphana University. In other MOOCs, besides having the possibility of receiving a certificate after completing a course by paying a fee, participants may be awarded with the SA if achieving specific requirements, which usually take the form of tests (e.g. Engle, Mankoff, & Carbrey, 2015). There are two main differences between those MOOCs and the analysed one. Firstly, in this course not all participants were students. Secondly, the requirements for earning an SA

were specifically associated to peer-assessment. As previously mentioned, participants were required to: 1) assess 18 final submissions; and 2) have earned 3 Popular Evaluator badges by the end of the course. Because of the characteristics and requirements of this course specifically, receiving an acknowledgement seemed to be important for those engaging in the assessment activity. Thus, this challenges the finding by Liu et al. (2014) about the little importance that participants seem to give to certificates awarded in courses of this kind.

The third finding revealed that supporters differed from students in the use of 5 types of feedback: *Analysis*, *Praise - general*, *Praise - content*, *Revision* and *Criticism - general*. This difference appears to partly support the initial assumption that supporters feedback would focus more on types of feedback within the motivational functions than the cognitive functions. Nevertheless, this may not hold true and a situational context can explain this difference.

Whilst *Praise - content* requires the reviewer to have engaged with the content in order to be able to specify a focus for praise, *Praise - general* and *Criticism - general* do not require much, but a brief look at the work. This finding can possibly be associated with the requirement that was set for those aiming at receiving the SA, specifically that of assessing 18 final submissions. Although it is not implied that the degree of quality of elaboration decreased specially for the last phase, it can mean that the structure of the feedback provided in that phase was different to the previous ones. This interpretation, however, presents some limitations that will be discussed in sub-section 5.1.2.1.

The use of *Analysis* suggests two issues. Firstly, that there is an actual difference between both roles. Secondly, that there is an association with the requirements for the SA. Not having been required to undergo the process of cognitive analysis of each task (Narciss & Huth, 2004) indicates that supporters may have missed contextual elements that could have helped them understand the works better. Therefore, the use of *Analysis* can be understood as a means to bridge this information gap. If assessing the same team(s) over time, this information gap may have decreased. This understanding supports the second issue. Because of the requirements for eligibility for the SA, supporters were required to increase the amount of assessment of final submissions. Thus, they were required to assess other teams whose works they were not familiar with.

Finally, the use of *Revision* suggests that although modifications are proposed, reviewers agree with the overall ideas that have been presented. The use of this type of feedback can reflect familiarity with the works by the teams. It may mean that supporters provided feedback to the same teams throughout the course, and that they were familiar with their previous work and therefore did not require clarifications, as these possibly had been explained throughout the phases.

Despite the aforementioned differences, feedback type *Evaluation* was used with a similar frequency between both roles. This reflects a corresponding level of commitment and engagement with the course independent of the role description. This refutes the premise presented at the beginning, namely that deep and critical feedback would only be provided by learners undergoing the same process

(completing the same assignment under the same conditions). This also confirms that both roles shared some commonalities.

5.1.2.1 Limitations, practical implications and suggestions for future research

This analysis presents the following important limitation. Results reflect the use of types of feedback used by those in the population who had earned at least a Popular Evaluator badge, a badge that required great efforts and acknowledgment from their peers to be earned. This said, the sample does not represent the population of all those who provided feedback. However, on the other hand, the analysis does take into consideration the feedback provided by those participants who seem to have fulfilled the purpose of formative feedback with their assessments, and which their peers considered meaningful. One of the main efforts that is required to earn this acknowledgement is that of writing clearly and precisely (Ferguson, 2011; Lizzio & Wilson, 2008; Moon, 2005; Prins et al., 2005), which evidently requires not only a good level of proficiency in English, but also skilfulness in its use. This said, this analysis did not take into account the feedback that was provided by participants who may have been less able, less resourced, with less academic experience, or less proficient in English (Meek et al., 2017). This is problematic, as it supports the claim made by Liyanagunawardena (2015) and Liyanagunawardena et al. (2013) that MOOCs are indeed not accessible to all because of the requirements they impose on their participants.

In this case, it would have been important to take into account the whole population (N=824) and other variables that may reflect an association with the peer-feedback provided. These could be, for instance, whether participants' country of residence

had English as a national language, participants' self-assessed level of English, or participants' career or educational level. A study carried out by Meek et al. (2017) that took into account the language variable in the analysis of peer-review in a MOOC, reported a slight difference in the quality of peer-review provided by participants who resided in countries where English was the national language, compared to those who did not. That study, however, presented limitations and those participants taking part in that part of the analysis had to meet two requirements: having submitted a summary assignment and undertaken peer review.

Currently, little qualitative research has been carried out regarding participants' experiences in MOOCs, and the few studies that have done so present two common limitations: the sample size, and the sample selection. However, many of these studies highlight the important role played by language and culture in diverse online learning communities (Popov et al., 2012, 2014), also in MOOCs (Aharony & Bar-Ilan, 2016; Cho & Byun, 2017; Colas, Sloep, & Garreta-Domingo, 2016; Engle, & Carbrey, 2015; Liu et al., 2014; Liyanagunawardena, et al., 2013; Nkuyubwatsi, 2014; Reilly et al., 2016; Sanchez Gordon & Luján Mora, 2014). Thus, further research should take into account different samples, including participants from countries that do not represent a majority. This could give course designers a better overview of the reach of the course, the diversity of participants and enable them to identify possible needs or limitations they may encounter. It is, however, clear that this approach may not be feasible in courses with hundreds of thousands of participants, but perhaps it

can be done in those of smaller range like the one reported here, or other emerging types as those reported by Sanchez Gordon and Luján Mora (2014).

Although the purpose of this analysis was fulfilled, a further step would be analysing the structure of feedback provided. This way, possible patterns could be identified.

For instance, it would be intriguing to understand the usage pattern of *Analysis* and *Revision* (within a message and throughout the phases). If these were used together in a message (provided they referred to the same idea) it could mean that participants did not want to misspend any time waiting for a clarification, instead they took the following step, and they provided a suggestion –using *Revision*– based on their own understanding. This interpretation would support what Saunders (1989) and van der Pol et al. (2008) assert about peer assessment as a more limited form of collaborative learning, in which a lower degree of interactivity is offered. This results in the limited possibility for interactive construction of meaning and collaborative knowledge construction (Saunders, 1989). If these two types of feedback were used in the same message (and referring to the same idea), it could reflect the need for a quicker way of correspondence, even if later at some point their understanding could have been confirmed to be incorrect.

Further research could add another layer to the analysis, which regards a differentiation for the use of types of feedback between both roles for each of the learning phases. Although this information was available for this study, it was not required for answering the research question. However, as the topic related to the requirements associated to the SA arose during discussion, it would have been useful

to have this information in order to understand the use of types of feedback that appear easy to use and that seem not to require much time, such as *Praise - general* or *Criticism - general*. It was tempting to assume that these two were used the most in the last phase by supporters in order to fulfil the requirements; however, there is no evidence for this. This is to say, though, that adding another variable to the analysis also adds another layer of complexity, implying that different statistical models would need to be employed.

5.1.3 Differences in the use of feedback between female and male participants in their respective role

In this MOOC, participants had three options to categorise their gender when filling out their profile information: female, hidden and male. This analysis only focused on the options female and male.

The aim of this last analysis was to describe possible differences between both genders in their roles as they took place in the course. Despite different studies in online learning environments where differences between genders had been reported, the main hypothesis of this analysis was that there were likely to be no differences found between the nature of feedback provided by female or male participants. In fact, similarly to the study by Read et al. (2005), no significant statistical differences could be established between both sex categories. However, after exploring the role performed by the participants in their respective gender, differences could be established. In other words, the gender of the participants providing specific types of feedback becomes important in this analysis, completing the picture of this case study.

It is worth noting that the statistical analyses that were employed for this and the previous phase considered the clustering in the feedback provided by participants. Thus, although the number of male participants was significantly lower than that of female participants, data from both were brought to a standard condition.

The previous analysis reported that supporters overtook students in the use of five types of feedback. Based on this, there are three main findings of this last analysis. First, *Analysis* and *Praise - general* were used by both female and male supporters without any significant difference. Secondly, it was female supporters who employed the types *Praise - content*, *Revision* and *Criticism – general* more significantly. Third, female supporters overtook female students in the use of *Evaluation* and male students in the use of *Substance.*

These results reveal that females (here in their role as supporters) highly engaged in the peer-assessment activity, contradicting Meek et al.'s (2017) findings in a MOOC. Moreover, these results support those studies reporting on female participants generally posting more messages in online learning environments than males (Bostock & Lizhi, 2010; Gunn & McSporran, 2003; Rovai & Baker, 2005; Wishart & Guy, 2009), possibly resulting from their preference for online discussions against face-to-face discussions (Bostock & Lizhi, 2010).

Whilst in many MOOCs the number of women enrolled is rather low (e.g. Christensen et al., 2013; Dillahunt et al., 2014), the number of women enrolled in this particular course was considerably higher. As described in the context, this type of MOOC is less conventional. It focuses on collaboration in small groups, and relies

on the connection across a network of peers. These characteristics appear to go in line with women's interaction styles that have been reported by Price (2006) and Yukselturk and Bulut (2009). According to these authors, women's style tends to be oriented towards networking, collaboration and community.

The first finding appears to contradict what has been reported for men in the context of online courses. Blum (1999), and Rovai and Baker (2005), for instance, argue that men's voices in interactions adopt a disinterested, authoritative and inquiring form. Although the use of *Analysis* may require an inquiring form, it expresses mainly an actual interest in understanding a work or parts of it. Its use does not seem compatible with authoritative forms of interaction, as the tone would most probably be categorised as other types of feedback (either within the cognitive or motivational dimensions). Similarly, according to Herring's (1993) findings, the use of *Praise - general*, would most likely speak for its use by women. This author suggests that the features used by men in their language include strong assertions, rhetorical questions, self-promotion, humour or sarcasm, among others. Conversely, according to the author, women's features are attenuated assertions, questions, apologies, personal orientation and support, among others. This last set of features may back up the use of *Analysis* by women. Rovai and Baker (2005) point out that – despite the characteristics of men's and women's voices in the motivational dimension– most of the messages that they analysed for their study were task-oriented (cognitive dimension). Unfortunately, for that dimension no characteristics were provided for any gender.

These studies represent a handful of examples in the literature that have studied gender differences in communication patterns prior to the emergence of MOOCs, and their findings are considered as the foundation for several investigations across disciplines. Whilst some current studies (e.g. Merchant, 2012) continue reinforcing gender stereotypes without questioning the characteristics of the context and the period of time in which these were investigated, some others challenge previous findings based on their own results (Wishart & Guy, 2009), despite their cultural context where specific gendered behaviours are expected (e.g. Netshitangani, 2008). Still, it appears that although extensive research has been done on culture in the field of distance and online education (Al-Harthi, 2014), gender issues in online learning remain little explored. This investigation supports the second type of studies by raising awareness of the manifold factors that may have contributed to differences in the case at stake, starting by the learning context itself, the type of MOOC and its topic. With this, it is argued that a generalisation on gender differences is neither possible nor meaningful, as it would ignore the individual characteristics of a person, which have been strongly influenced by their cultural context and their own experiences. In this specific case, the use of *Analysis* and *Praise - general* has been previously addressed and associated to different factors.

The second finding shows an exclusive use by females of feedback types *Praise - content*, *Revision* and *Criticism - general*. The use of *Praise - content* appears to be supported by the study by Leung, Chan, Maxwell and Poon (2010). The authors reported that in a Wiki-supported Chinese language class, women provided comments with a specific focus, such as content, structure and style. Interestingly,

these authors observed a pattern in the use of feedback by male and female students. These included the combination of *Praise* and *Criticism.* Whilst men used it in the order as presented, women tended to be more critical and analytical before offering any *Praise.* Although the identification of patterns was not the focus of the present study, findings concur in that female participants in the MOOC focused on content for providing *Praise.* However, when offering *Criticism,* the comments remained general.

In the Wiki-supported class, students provided feedback using a pseudonym. This characteristic would seem more common in a MOOC than in an environment with true peers (students with homogeneous characteristics). Although in this MOOC all participants provided an identity, it is worth considering that: 1) the identity provided by supporters may not have been the true one, in which case they would have not profited from the SA (if earned); 2) the identity provided by supporters may have been true, but they were still anonymous within the large and diverse community. The latter leads to the idea that anonymity may play a certain role in the way participants provide feedback, independent from their gender. In fact, a study carried out by López-Benavides (2015) in which types of peer-feedback in a similar MOOC were analysed, mentions that anonymity can be a factor empowering the use of certain tones in participants messages. However, this was not further explored. Being in and also exposed to a diverse community could result in empowering or supressing situations for any participant independent of their own socio-cultural context. The patterned behaviours that may be encouraged or restricted for women or men in the socio-cultural context they are familiar with, are likely not to be found

in this type of learning context. Thus, findings of this study in this matter appear to be less likely comparable to studies that were carried out in other settings with different characteristics.

In Leung et al.'s (2010) study, the use of three types of *Revision* were identified: 'Direction for Suggestion', 'Direct Error Correction', and 'Specific Suggestion'. The first two were reported to be used more by female participants, and the latter by male participants. In this MOOC, the type *Revision* contained all comments that offered ideas, suggestions or advice with a focus on content, and findings suggest that female supporters used this type significantly more. Nonetheless, understanding a specific type of *Revision* may offer a connection to participants' stage of critical thinking, as the ones presented by Baxter Magolda (2007), which were discussed in Chapter 2.

Finally, the last finding revealed that female participants overtook female students in the use of *Evaluation* and male students in the use of *Substance.* The use of the former evidently speaks for a difference in the role performed, therefore it partially concurs with the observations made by Wishart and Guy's (2009) over three years on gender patterns in online discussions in online Master Business Administration (MBA) International Business courses. In that study, it was reported that men employed evaluative and inquiring statements, yet women were more likely to be critical and challenging. The use of *Evaluation* required reviewers to identify those pieces of information that put into question the work, decisions or conclusions within a work. *Evaluation* is considered to be the most influential type of feedback that may make a group reconsider the whole purpose or line of argument in their

work. In this context, Leung et al. (2010) reported that women were more able to locate problems in a work than men. In other words, and as mentioned previously in another feedback type, women provided comments after having identified a specific focus. On the other hand, the same authors reported that men were more able to identify good points in a work than women. However, the high use of *Substance* by women, which requires reviewers to consider the usefulness of a work or parts of it, contradicts the findings of those authors.

The use of both types of feedback confirm that participants employing it (in this specific analysis female supporters versus female and male students) must have reached all levels of high order thinking skills from Bloom's amended taxonomy (Krathwohl, 2002), as well as the contextual stage of critical thinking (Baxter Magolda, 2007). Whilst participants' characteristics for using these types of feedback are clear, the factors that had an influence over this difference are not.

5.1.3.1 Limitations, practical implications and suggestions for future research

Understanding that different factors may have affected the way in which participants engaged in the assessment activity and the elaboration of their messages, a series of issues arise. The issues are not seen here as limitations per se, but as important topics that need to be considered in further studies. These present practical implications and provide ideas for future work. Although the issues apply to the different analyses undertaken in the study, their significance was recognised as the findings of this final analysis were reported.

The tones used in the messages varied among the different participants. A positive tone was commonly supported by positive words, whereas a negative tone may not have always been reflected by negative words. The first issue identified was: interpretation. As expected, interpretation is linked to intercultural awareness and proficiency in the language of the person reading or analysing a message. This study was carried out by a single researcher, and although a consistent procedure was followed during analysis, the result of the interpretation was based on the researcher's own intercultural awareness and proficiency in the language. Taking into account that the focus of analysis was placed on the content of the TUs and not on the other codes (learning phase, reviewer's role and gender), the result of this interpretation revealed that the tone used by female supporters was significantly more negative than by students in both genders. Yet, a question remains: why?

The focus of this and all previous analyses was to describe behaviours as these happened in the course, and not to find out the reason for those behaviours. However, after having reported on the second and third findings, specifically on the use of *Criticism* and *Evaluation* among female supporters and female students, it is intriguing to understand the reasons influencing this behaviour. Understanding it can have important implications in the pedagogical design of the course and the requirements that are set for each type of participant, or the requisites for earning certain incentives.

On the one hand, it appeared that time could have played a role here and there are some arguments supporting this notion (discussed in previous sections); however, there is no further evidence confirming it, other than the information provided in the

post-course survey. Thus, if other investigations report similar behaviours, it seems worth analysing other variables and whether correlations exist. Information shared by participants when enrolling in the course may be taken into account, such as age, career- and educational-level, level of familiarity with peer-assessment activities, as well as the level of proficiency in which the course is offered, or the national language spoken in the country where participants reside, as requested in other courses (Meek et al., 2017). All information that can be gathered raises another issue: inaccuracy of information provided by participants. The nature of a MOOC appears to support this. Participants in this environment may choose not to provide veridical information, nor provide information for all areas of interest of the researcher, making assessment of information difficult. Although techniques for reducing this type of bias are commonly known in qualitative methods such as interviews (Warner, 1965), it appears unlikely that biases can neither be reduced in a voluntary survey in a MOOC, nor in the information participants provide when filling in their profiles. Moreover, even if certain requirements (i.e. enrolling as a student with the aim of receiving a certificate) may expect participants to reveal veridical information, this is only valid for one type of participant and not necessarily for the other. However, if in this context only qualitative methods are employed (such as interviews), as other studies have done, then another issue emerges: continuing adding studies to the literature that present two common drawbacks related to the sample, namely its size and selection.

Findings of this last analysis offer ideas for further research, in which the questions to be investigated could be: Do individuals tend to employ certain types of feedback

each time they assess? Under which circumstances do individuals tend to modify the use of feedback types?

5.2 Conclusion

Considering that the trend of including peer-feedback strategies in MOOCs will continue to increase (O'Toole, 2013; Pilli & Admiraal, 2017) and that little is actually known about current practices, the need to explore this field further becomes more evident. Taking into account that MOOC participants are diverse and that their participation in the courses is voluntary, the present study aimed at inquiring into the nature of peer-feedback in the MOOC "Managing the Arts: Marketing for Cultural Organizations". The method of content analysis and a regression model with a Poisson distribution were employed to explore and describe three different aspects of feedback in the course: the quantity and quality of feedback provided throughout the course; the feedback provided by students and supporters; and the feedback provided by females and males in their roles as students and supporters.

To this end, three specific research questions were formulated:

1. How does feedback evolve over time in terms of quantity and quality?

The quantity of feedback provided in the MOOC was inconstant but displayed a positive trend. The quantity of badges awarded or the types of badges awarded do not offer a reliable association with the quality of feedback provided, especially if participants rating comments were not the ones for which these were intended. In fact, and in agreement with Prins et al. (2006) and Read et al. (2005), the value of feedback appears to lie in the perception of the person receiving it or reading it, and the moment in learning in which feedback was received or read. This perception might be intensified by individual characteristics of the learners, which as expected

are directly associated with their sociocultural context and their previous collected experiences. According to the findings of this study, the types of feedback perceived as quality feedback include a combination of types within the dimensions of feedback introduced by Narciss and Huth (2004), and Nicol and Macfarlane-Dick (2006): cognitive, metacognitive and motivational. The evolution of the expected types of feedback required for the suggested types of tasks was constant, whereas the other types displayed alternating shifts.

2. How does the nature of feedback differ between students and supporters?

Differences were encountered in the use of five types of feedback contained in the cognitive and motivational dimensions of feedback, whereby supporters overtook students. These differences are believed to be strongly associated with the contextual situations of the learning phases of the course and to how each type of role experienced them (in relation to individual aims for the course, possibilities and limitations). Despite the differences, both types of participants taking part in this voluntary activity shared a specific commonality, related to a certain level of confidence (Falchikov, 2005), thus, allowing them to engage in this cognitively demanding activity. Furthermore, the level of confidence appears to combine linguistic and academic aspects, as well as specific knowledge on the topic treated.

3. How does the nature of feedback differ between female and male participants in their respective roles as students or supporters?

Despite the fact that no differences could be established for the use of peer-feedback types between both genders, differences were established in the roles that

they performed. The differences encountered challenge findings from other studies in which gender stereotypes are reinforced. This suggests that the characteristics of the learning context, the dynamics developed therein, as well as the empowering or supressing situations experienced by the participants, to name a few, greatly impact participants' behaviour during learning and participation. Moreover, the differences encountered recognise the individuality in each participant and acknowledge that sociocultural factors, personal skilfulness and the situational context may have played a decisive role regarding the nature of feedback provided.

Despite the fact that the number of participants who took part in the peer-assessment activity was significantly small compared to the number of enrolled participants, it can be suggested that their diversity did not represent a risk to the purpose and value of the very last step of the cycle of the evaluation process: assessment, which was expressed as a concern in the introduction of this study. Their participation was key in understanding the nature of feedback in such a course and in confirming that manifold variables underpin the intricacies of peer-feedback. Although these variables can be researched independently, it appears impractical to address many of them in one study, as well as study their correlations. Limitations of various types emerge, as presented previously. This case did not present pedagogical differences between the learning phases, and if there would have been any, these would have only represented important but not exclusive indicators for the nature of peer-feedback provided. It is to consider that the characteristics of the learning setting (participants' diversity and voluntary participation) combined with the individual lived experiences are equally important, increasing the complexity of the

case. A number of contextual and individual factors may play a great role in the nature of peer-feedback identified in the courses. These factors are not assumed to be present simultaneously, but to emerge depending on situations experienced at an individual and collective level. Examples of factors are: time, mood, engagement, access in its different dimensions, internal and external incentives and the works to be assessed. It is believed that because every MOOC happens in a unique way, different variables of interest will arise based on those different situational contexts. This case study provides a unique view of peer-feedback in a less conventional type of MOOC that is team-based, offers two different modes of participation, and where different supporting actors are present. In contrast to other MOOCs reported in the literature, this course assigns an equal value to all three steps of the learning cycle: teach, learn and assess, and the way this is accomplished is reflected in the different elements integrated in its pedagogical design. This study pioneers the exploration of the field of peer-feedback in MOOCs, and sheds light on some of the hitherto unexplored variables that are: feedback quantity and quality, and feedback in relation to participant's role and gender.

It is evident that although a specific set of assessment criteria was provided in the analysed course, participants found their own ways to express the feedback that they felt was important, or that was more accessible for them in terms of language, among others. This is reflected in the emergence of 20 types of feedback. Consequently, it appears that assessment criteria should be clearly presented as a suggestion and not as an imposition. The latter may have a negative impact in the quantity and also quality of feedback provided, which can lead to disappointment for

those providing feedback but also to those receiving it. In the course analysed, these criteria were fixed items that were in an automatically generated template.

At the same time, mechanisms for motivating learners in taking part in this assessment strategy need to be further researched, to be able to offer and support a lively dynamic in a learning setting that interacts with and depends on its community.

The inclusion of peer-feedback practices in such context is undoubtedly challenging, as the degree of expectations vary along with the topic of the course, participants' diversity and own lived experiences. Thus, future research could analyse peer-feedback in relation to the general topic of the course and a common background of the participants, as this can evidently have a direct impact on the types of feedback that are used and that seem to fulfil the expectations of those particular groups. This understanding would help instructional designers to focus on encouraging those types of feedback that should be provided, thus supporting what Kop and Fournier (2010) suggest:

> In order to develop empowering learning environments that foster active learning, designers and developers of such environments first need to understand the factors that influence people's attitudes, intentions and behaviours. They must also understand the prerequisites for people to thrive in such environments in order to create favourable components and conditions. (p. 5)

5.3 Study implications

The present investigation studied a specific single case with a less conventional pedagogical approach. Because of its characteristics it represents a unique or an atypical case (Merriam, 1998) that – although of great value– can have implications for practitioners and researchers who have been involved in MOOCs offered by commercial learning platforms (e.g. Coursera, edX, Udacity, FutureLearn) or in MOOCs that do not place a focus on peer-collaboration and does not rely on peer-feedback. One of the basic differences between most of the courses offered by those platforms and the course that has been studied is that although collaboration between peers is encouraged, participants can complete the courses independently. Thus, practitioners and researchers may be more interested in mechanisms for motivating individual learners to complete courses. Because of the peer-to-peer focus of this course, the level of complexity for creating and maintaining a lively dynamic among participants of a learning community is greatly higher and requires the integration and coordination of different pedagogical elements. Accordingly, the requirements for some parts of the conception, design and implementation of this type of course may differ as well and may appear not transferable at first sight.

Moreover, whilst a great effort in the organisation and coordination of the courses is expected when creating any MOOC, this type of course requires an extra level of flexibility, attentiveness and readiness during the running of the course. Here, a special team (i.e. *MOOC facilitator*, *mentors and tutors*) is required to immerse in the course and exchange continuously regarding different happenings (e.g. communication going on in any of the spaces provided within the learning platform,

overall understanding of the requirements of the assignments and the introduced concepts, wellbeing of the teams) that may affect the healthy development of a lively dynamic of a learning community. Flexibility of this type can imply immediate modifications at technological or pedagogical levels. These are all challenges that may not necessarily emerge in other types of MOOC, making this case informative but less transferable. Despite the implications mentioned previously which can lead to limitations for some practitioners and researchers working on different types of MOOC, the characteristics of this case are considered to be eye-openers. It is through different pedagogical approaches that practitioners can enhance the pedagogical design of their courses and researchers can expand their understanding about the intricacies underpinning the learning process in a challenging learning setting.

References

Admiraal, W., Huisman, B., & van De Ven, M. (2014). Self- and Peer Assessment in Massive Open Online Courses. *International Journal of Higher Education*, *3*(3), 119–128. http://doi.org/10.5430/ijhe.v3n3p119

Aharony, N., & Bar-Ilan, J. (2016). Students' Perceptions on MOOCs: An Exploratory Study. *Interdisciplinary Journal of E-Skills and Lifelong Learning*, *12*, 145–162. http://doi.org/10.28945/3540

Al-Harthi, A. (2014). An Analysis of Culture-Focused Articles in Open, Distance and Online Education Journals. In I. Jung & C. N. Gunawardena (Eds.), *Culture and online learning: Global perspectives and research* (pp. 37–54). Virginia: Stylus Publishing, LLC.

Alfred, M. V. (2002). The Promise of Sociocultural Theory in Democratizing Adult Education. *New Directions for Adult and Continuing Education*, *96*, 3–14. http://doi.org/10.1002/ace.74

Alvarez, I., Espasa, A., & Guasch, T. (2012). The value of feedback in improving collaborative writing assignments in an online learning environment. *Studies in Higher Education*, *37*(4), 387–400. http://doi.org/10.1080/03075079.2010.510182

Balfour, S. (2013). Assessing writing in MOOCS: Automated essay scoring and Calibrated Peer Review. *Research & Practice in Assessment*, *8*(1), 40–48. Retrieved from http://www.rpajournal.com/dev/wp-content/uploads/2013/05/SF4.pdf

Bartholet, J. (2013). Free online courses bring "magic" to Rwanda. Retrieved from https://www.scientificamerican.com/article/free-online-classes-bring-magic-rwanda/

Baxter Magolda, M. B. (2007). Self-authorship: The foundation for twenty-first-century education. *New Directions for Teaching and Learning*, *2007*(109), 69–83. http://doi.org/10.1002/tl.266

Bayne, S., & Ross, J. (2014). *The pedagogy of the Massive Open Online Course: the UK view*. York, UK: The Higher Education Academy. Retrieved from https://www.heacademy.ac.uk/sites/default/files/HEA_Edinburgh_MOOC_WEB_240314%0A_1.pdf

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. http://doi.org/10.1080/0969595980050102

Blum, K. D. (1999). Gender differences in asynchronous learning in higher education: Learning styles, participation barriers and communication patterns. *Journal of*

Asynchronous Learning Networks, *3*(1), 46–66.

Bostock, S., & Lizhi, W. (2010). Gender in Student Online Discussions. *Innovations in Education and Teaching International*, *42*(1), 37–41. http://doi.org/10.1080/14703290500048978

Bouchard, P. (2009). Some Factors to Consider When Designing Semi- Autonomous Learning Environments. *Electronic Journal of E-Learning*, *7*(2), 93–100.

Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, *22*(2), 151–167. http://doi.org/10.1080/713695728

Boyer, S. L., Edmondson, D. R., Artis, A. B., & Fleming, D. (2014). Self-Directed Learning. *Journal of Marketing Education*, *36*(1), 20–32. http://doi.org/10.1177/0273475313494010

Budd, R. W., Thorp, R. K., & Donohue, L. (1967). *Content analysis of communications*. New York, NY: The Macmillan Co.

Busuttil-Reynaud, G., & Winkley, J. (2006). *JISC e-Assessment Glossary (Extended)*. Retrieved from http://www.webarchive.org.uk/wayback/archive/20140615085353/http://www.jisc.ac.uk/media/documents/themes/elearning/eassess_glossary_extendedv101.pdf

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281. http://doi.org/10.2307/1170684

Charalambos, V., Michalinos, Z., & Chamberlain, R. (2004). The Design of Online Learning Communities: Critical Issues. *Educational Media International*, *41*(2), 135–143. http://doi.org/10.1080/09523980410001678593

Chen, N. S., Wei, C. W., Wu, K. T., & Uden, L. (2009). Effects of high level prompts and peer assessment on online learners' reflection levels. *Computers and Education*, *52*(2), 283–291. http://doi.org/10.1016/j.compedu.2008.08.007

Cheng, W., & Warren, M. (1997). Having second thoughts: Student perceptions before and after a peer assessment exercise. *Studies in Higher Education*, *22*(2), 233–239. http://doi.org/10.1080/03075079712331381064

Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, *20*(4), 328–338. http://doi.org/10.1016/j.learninstruc.2009.08.006

Cho, K., Schunn, C., & Charney, D. (2006). Commenting on writing typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication*, *23*(3), 260–294.

http://doi.org/doi:10.1177/0741088306289261

Cho, M.-H., & Byun, M. (2017). Nonnative English-Speaking Students' Lived Learning Experiences With MOOCs in a Regular College Classroom. *International Review of Research in Open and Distributed Learning*, *18*(5), 173–190.

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. J. (2013). *The MOOC Phenomenon: Who Takes Massive Open Online Courses and Why? SSRN Electronic Journal*. Retrieved from https://pdfs.semanticscholar.org/43f5/ab1690524d291ae21946683ae63b221564e4.pdf

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

Clow, D. (2013). MOOCs and the funnel of participation. In *Third Conference on Learning Analytics and Knowledge - LAK' 13* (pp. 185–189). Leuven, Belgium. http://doi.org/10.1145/2460296.2460332

Coffrin, C., Corrin, L., de Barba, P., & Kennedy, G. (2014). Visualizing patterns of student engagement and performance in MOOCs. *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14*, 83–92. http://doi.org/10.1145/2567574.2567586

Cohen, L., Manion, L., & Morrisson, K. (2000). *Research methods in education*. London: Routledge Falmer.

Colas, J. F., Sloep, P. B., & Garreta-Domingo, M. (2016). The effect of multilingual facilitation on active participation in MOOCs. *International Review of Research in Open and Distance Learning*, *17*(4), 280–313. http://doi.org/10.1017/CBO9781107415324.004

Costello, J., & Crane, D. (2013). Technologies for Learner-Centered Feedback. *Open Praxis*, *5*(3), 217–225.

Creswell, J. W. (1998). *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

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

Davies, P. (2006). Peer assessment: judging the quality of students' work by comments rather than marks. *Innovations in Education and Teaching International*, *43*(1), 69–82. http://doi.org/10.1080/14703290500467566

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A

review. *Computers and Education*, *46*(1), 6–28. http://doi.org/10.1016/j.compedu.2005.04.005

Denscombe, M. (2003). *The Good Research Guide For small-scale social research projects* (2nd ed.). Bukingham: Open University Press.

Denscombe, M. (2010). *The Good Research Guide For small-scale social research projects* (4th ed.). Maidenhead: McGraw-Hill. http://doi.org/10.1371/journal.pone.0017540

Dillahunt, T., Wang, Z., & Teasley, S. D. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *International Review of Research in Open and Distance Learning*, *15*(5), 177–196. http://doi.org/10.19173/irrodl.v15i5.1841

Dippold, D. (2009). Peer Feedback Through Blogs: Student and teacher perceptions in an advanced German class. *ReCALL*, *21*(1), 18–36. http://doi.org/10.1017/S095834400900010X

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, *24*(3), 331–350. http://doi.org/10.1080/0307507991233137993 5

Donnelly, R., & Gardner, J. (2011). Content analysis of computer conferencing transcripts. *Interactive Learning Environments*, *19*(4), 303–315. http://doi.org/10.1080/10494820903075722

Emanuel, E. (2013). MOOCs taken by educated few. *Nature*, *503*(7476), 342–342. http://doi.org/10.1038/503342a

Engle, D., Mankoff, C., & Carbrey, J. (2015). Coursera's introductory human physiology course: Factors that characterize successful completion of a MOOC. *International Review of Research in Open and Distance Learning*, *16*(2), 46–68.

Eun, B. (2010). From learning to development: a sociocultural approach to instruction. *Cambridge Journal of Education*, *40*(4), 401–418. http://doi.org/10.1080/0305764X.2010.526593

European Commission. (2014). *Report on web skills survey. Support services to foster Web Talent in Europe by encouraging the use of MOOCs focused on web talent*. Retrieved from https://www.openeducationeuropa.eu/sites/default/files/news/MOOCs-for-web-skills-survey-report.pdf

Evans, C. (2013). Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research*, *83*(1), 70–120. http://doi.org/10.3102/0034654312474350

Evans, C., & Waring, M. (2011). Exploring students' perceptions of feedback in

relation to cognitive styles and culture. *Research Papers in Education*, *26*(2), 171–190. http://doi.org/10.1080/02671522.2011.561976

Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment and Evaluation in Higher Education*, *11*(2), 146–166. http://doi.org/10.1080/0260293860110206

Falchikov, N. (2005). *Improving assessment through student involvement: practical solutions for aiding learning in higher and further education*. London: Routledge Falmer.

Falchikov, N., & Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, *70*(3), 287–322. http://doi.org/10.3102/00346543070003287

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education*, *36*(1), 51–62. http://doi.org/10.1080/02602930903197883

Fini, A. (2009). The Technological Dimension of a Massive Open Online Course: The Case of the CCK08 Course Tools. *International Review of Research in Open and Distance Learning*, *10*(5), 1–26.

Flick, U. (2009). *An introduction to qualitative research* (4th ed.). Thousand Oaks, CA: Sage Publications.

Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, Diagnosis, and the Strategies of Revision. *College Composition and Communication*, *37*(1), 16–55.

Frederiksen, J. R., & Collins, A. (1989). A Systems Approach to Educational Testing. *Educational Researcher*, *18*(9), 27–32. http://doi.org/10.3102/0013189X018009027

Fund, Z. (2010). Effects of communities of reflecting peers on student-teacher development - including in-depth case studies. *Teachers and Teaching*, *16*(6), 679–701. http://doi.org/10.1080/13540602.2010.517686

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, *20*(4), 304–315. http://doi.org/10.1016/j.learninstruc.2009.08.007

Goodman, J., Wood, R., & Hendrickx, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology*, *89*(2), 248–262. http://doi.org/10.1037/0021-9010.89.2.248

Guasch, T., & Espasa, A. (2015). Collaborative Writing Online: Unravelling the Feedback Process. In M. Deane & T. Guasch (Eds.), *Learning and Teaching Writing Online* (pp. 13–30). Leiden, The Netherlands: Brill.

Guasch, T., Espasa, A., & Alvarez, I. (2010). Formative e-feedback in collaborative writing assignments: the effect of the process and time. *ELC Research Paper Series*, *1*, 49–59.

Guasch, T., Espasa, A., Alvarez, I. M., & Kirschner, P. A. (2013). Effects of feedback on collaborative writing in an online learning environment. *Distance Education*, *34*(3), 324–338. http://doi.org/10.1080/01587919.2013.835772

Gunn, C., & McSporran, M. (2003). Dominant or different: Gender issues in computer supported learning. *Journal of Asynchronous Learning Networks*, *7*(1), 14–30. Retrieved from http://cs.lamar.edu/faculty/osborne/COSC1172/v7n1_gunn.pdf

Gutierrez, K. D., & Rogoff, B. (2003). Cultural Ways of Learning: Individual Traits or Repertoires of Practice. *Educational Researcher*, *32*(5), 19–25. http://doi.org/10.3102/0013189X032005019

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. http://doi.org/10.3102/003465430298487

Herring, S. C. (1993). Gender and democracy in computer-mediated communication. *Electronic Journal of Communication*, *3*(2), 476–489. Retrieved from http://portal.acm.org/citation.cfm?id=233411

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development*, *27*(1), 55–67. http://doi.org/10.1080/07294360701658765

Hsieh, H.-F., & Shannon, S. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277–1288. http://doi.org/10.1177/1049732305276687

Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science*, *39*(3), 387–406. http://doi.org/10.1007/s11251-010-9133-6

Kim, M. (2005). *The effects of the assessor and assessee's roles on preservice teachers' metacognitive awareness, performance, and attitude in a technology-related design task*. Retrieved from http://purl.flvc.org/fsu/fd/FSU_migr_etd-3051

Knowles, M. (1975). *Self-directed learning*. New York, NY: Association Press.

Kop, R., & Fournier, H. (2010). New Dimensions of Self-Directed Learning in an Open-Networked Learning Environment. *International Journal of Self-Directed Learning*, *7*(2), 2–20.

Kop, R., Fournier, H., & Mak, J. F. S. (2011). A Pedagogy of Abundance or a Pedagogy to Support Human Beings? Participant support on Massive Open Online Courses. *The International Review of Reshearch in Open and Distributed*

Learning, *12*(7), 74–93. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1041/2042

Krathwohl, D. (2002). A Revision of Bloom's Taxonomy: An Overview David. *Theory Into Practice*, *41*(4), 212–218. http://doi.org/10.1207/s15430421tip4104_2

Krause, S. D. (2013). MOOC response about "listening to world music." *College Composition and Communication*, *64*(4), 688–695.

Krogstie, B. R., Horgen, S. A., & Hjeltnes, T. (2015). Criteria for evaluating peer feedback comments in a MOOC. *Norsk Konferanse for Organisasjoners Bruk Av IT*, *23*(1). Retrieved from http://ojs.bibsys.no/index.php/Nokobit/article/view/285

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, *1*(4), 279–308. http://doi.org/10.1007/BF01320096

Kulkarni, C. (2014). Making distance matter. In *Proceedings of the adjunct publication of the 27th annual ACM symposium on User interface software and technology - UIST'14 Adjunct* (pp. 25–28). http://doi.org/10.1145/2658779.2661169

Kulkarni, C., Bernstein, M., & Klemmer, S. (2015). PeerStudio: Rapid Peer Feedback Emphasizes Revision and Improves Performance. In *L@S '15 Proceedings of the Second (2015) ACM Conference on Learning @ Scale* (pp. 75–84). http://doi.org/10.1145/2724660.2724670

Kulkarni, C., Cambre, J., Kotturi, Y., Bernstein, M., & Klemmer, S. (2016). Talkabout: Making Distance Matter with Small Groups in Massive Classes. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design Thinking Research: Making Design Thinking Foundational* (pp. 67–92). Cham, Switzerland: Springer International Publishing. http://doi.org/10.1007/978-3-319-19641-1_6

Kulkarni, C., Wei, K. O. H. P., Le, H. U. Y., Chia, D., Cheng, J., Koller, D., & Klemmer, S. (2013). Peer and Self Assessment in Massive Online Classes. *Acm Transactions On Computer-Human Interaction*, *20*(6), Art. 33. http://doi.org/10.1145/2505057

Leber, J. (2013). In the developing world, MOOCs start to get real. Retrieved from http://www.technologyreview.com/news/512256/in-the-developing-world-moocs-start-to-get-real

Leung, K., Chan, M., Maxwell, G., & Poon, T. (2010). A Qualitative Analysis of Sub-degree Students Commentary Styles and Patterns in the Context of Gender and Peer e-Feedback. In P. Tsang, S. Cheung, V. Lee, & R. Huang (Eds.), *Third International Conference, ICHL 2010* (Vol. 6248 LNCS, pp. 149–159). Beijing, China: Springer-Verlag Berlin Heidelberg. http://doi.org/10.1007/978-3-642-14657-2_28

Lewis, D., & Allan, B. (2004). *Virtual learning communities: A guide for practitioners*. New York, NY: McGraw-Hill Education.

Lincoln, Y., & Guba, E. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications.

Liu, M., Kang, J., Cao, M., Lim, M., Ko, Y., Myers, R., & Schmitz Weiss, A. (2014). Understanding MOOCs as an Emerging Online Learning Tool: Perspectives From the Students. *American Journal of Distance Education*, *28*(3), 147–159. http://doi.org/10.1080/08923647.2014.926145

Liu, N.-F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279–290. http://doi.org/10.1080/13562510600680582

Liyanagunawardena, T. (2015). Massive Open Online Courses. *Humanities*, *4*, 35–41. http://doi.org/10.1080/15228959.2012.730415

Liyanagunawardena, T. R., Williams, S., & Adams, A. (2013). The Impact and Reach of MOOCs: A Developing Countries' Perspective | eLearning. *ELearning Papers*, (May), 1–8. http://doi.org/10.1016/j.jbusres.2007.11.013

Liyanagunawardena, T., Williams, S., & Adams, A. (2013). The Impact and Reach of MOOCs: A Developing Countries' Perspective. *ELearning Papers*, 38–46. http://doi.org/10.1016/j.jbusres.2007.11.013

Lizzio, A., & Wilson, K. (2008). Feedback on assessment: students' perceptions of quality and effectiveness. *Assessment & Evaluation in Higher Education*, *33*(3), 263–275. http://doi.org/10.1080/02602930701292548

Lockhart, C., & Ng, P. (1995). Analyzing Talk in ESL Peer Response Groups: Stances, Functions, and Content. *Language Learning*, *45*(4), 605–651. http://doi.org/10.1111/j.1467-1770.1995.tb00456.x

López-Benavides, D. (2014). *Assessing Leuphanas' Digital School MOOC course design from a multicultural perspective* (Unpublished assignment for ED.S825 Interculturality and Globalisation in Technology Enhanced Learning in Doctoral Programme in E Research and Technology Enhanced Learning). Lancaster: Lancaster University.

López-Benavides, D. (2015). *Exploring types of peer-feedback in a cMOOC* (Unpublished assignment for ED.S823 Researching Technology Enhanced/Networked Learning, Teaching and Assessment in Doctoral Programme in E Research and Technology Enhanced Learning). Lancaster: Lancaster University.

Luo, H., Robinson, A., & Park, J.-Y. (2014). Peer Grading in a MOOC: Reliability, Validity, and Perceived Effects. *Journal of Asynchronous Learning Networks*, *18*(2). Retrieved from http://eric.ed.gov/?id=EJ1036269

Mak, S. F. J., Williams, R., & Mackness, J. (2010). Blogs and forums as communication and learning tools in a MOOC. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, & T. Ryberg (Eds.), *Proceedings of the 7th International Conference on Networked Learning 2010* (pp. 275–284). Aalborg, Denmark.

Meek, S. E. M., Blakemore, L., & Marks, L. (2017). Is peer review an appropriate form of assessment in a MOOC? Student participation and performance in formative peer review. *Assessment & Evaluation in Higher Education*, *42*(6), 1000–1013. http://doi.org/10.1080/02602938.2016.1221052

Merchant, K. (2012). *How men and women differ: Gender differences in communication styles, influence tactics, and leadership styles*. *CMC senior Theses*. Retrieved from http://scholarship.claremont.edu/cmc_theses/513/

Merriam, S. B. (1988). *Case study research in education: A qualitative approach*. San Francisco, CA: Jossey-Bass.

Merriam, S. B. (1998). Case Study as Qualitative Research. In *Qualitative Research And Case Study Applications in Education*. San Francisco, CA: Jossey-Bass.

Mills, A., Durepos, G., & Wiebe, E. (Eds.). (2010). *Encyclopedia of Case Study Research*. Thousand Oaks, CA: Sage Publications. http://doi.org/10.4135/9781412957397

Moon, J. (2005). We seek it here: a New perspective on the elusive activity of critical thinking: a theoretical and practical approach. *The Higher Education Academy*. Retrieved from www.escalate.ac.uk/2041

Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktion spsychologischer Erkenntnisse (Informative Tutoring Feedback)*. Münster, Germany: Waxmann.

Narciss, S. (2008). Feedback Strategies for Interactive Learning Tasks. In J. M. Spector, M. D. Merrill, J. Van Merrienboer, & M. P. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 125–144). Mahwah, NJ: Lawrence Erlbaum.

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multi-media learning. In H. M. Niegemann, D. Leutner, & R. Brunken (Eds.), *Instructional Design for Multimedia Learning* (pp. 181–195). Münster, Germany: Waxmann.

Nasir, N. I. S., Rosebery, A. S., Warren, B., & Lee, C. D. (2005). *The Cambridge Handbook of the Learning Sciences*. (R. K. Sawyer, Ed.)*The Cambridge handbook of the learning sciences*. Cambridge: Cambridge University Press. Retrieved from http://ebooks.cambridge.org/ref/id/CBO9780511816833

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: how different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375–

401. http://doi.org/10.1007/s11251-008-9053-x

Netshitangani, T. (2008). Gender differences in communication styles : The impact on the managerial work of a woman school principal. In *ANZCA08 Conference, Power and Place* (pp. 1–20).

Neubaum, G., Wichmann, A., Eimler, S. C., & Krämer, N. C. (2014). Investigating Incentives for Students to Provide Peer Feedback in a Semi-Open Online Course. In *Proceedings of The International Symposium on Open Collaboration, OpenSym'14*. New York, NY: ACM. http://doi.org/10.1145/2641580.2641604

Nicol, D. (2008). *Quality enhancement themes: Transforming assessment and feedback: Enhancing integration and empowerment in the first year*. Mansfield: Quality Assurance Agency.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. http://doi.org/10.1080/03075070600572090

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: a peer review perspective. *Assessment & Evaluation in Higher Education*, *39*(1), 102–122. http://doi.org/10.1080/02602938.2013.795518

Niewolny, K. L., & Wilson, A. L. (2006). (Re) Situating Cognition: Expanding Sociocultural Perspectives in Adult Education. In *Adult Education Research Conference*. Retrieved from http://newprairiepress.org/aerc/2006/papers/46

Nkuyubwatsi, B. (2014). Cultural Translation in Massive Open Online Courses (MOOCs). *ELearning Papers*, *37*, 23–32. Retrieved from https://www.openeducationeuropa.eu/en/article/Cultural-Translation-in-Massive-Open-Online-Courses-%28MOOCs%29

O'Toole, R. (2013). *Pedagogical strategies and technologies for peer assessment in Massively Open Online Courses (MOOCs)* (Unpublished discussion paper). Coventry: University of Warwick. Retrieved from http://wrap.warwick.ac.uk/54602/7/WRAP_O'toole_ROToole Peer Assessment in MOOCs (1).pdf

Patchan, M. M., Charney, D., & Schunn, C. D. (2009). A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor. *Journal of Writing Research*, *1*(2), 124–152.

Piech, C., Chen, Z., Koller, D., & Ng, A. (2013). Tuned Models of Peer Assessment in MOOCs. In *Proceedings of The 6th International Conference on Educational Data Mining (EDM 2013)* (pp. 153–160). Retrieved from https://arxiv.org/abs/1307.2579

Pilli, O., & Admiraal, W. (2017). Students' Learning Outcomes in Massive Open Online Courses (MOOCs): Some Suggestions for Course Design. *Yükseköğretim*

Dergisi, *7*(1), 46–71. http://doi.org/10.2399/yod.17.001

Popov, V., Biemans, H. J. A., Kuznetsov, A. N., & Mulder, M. (2014). Use of an interculturally enriched collaboration script in computer-supported collaborative learning in higher education. *Technology Pedagogy and Education*, *23*(3), 349–374. http://doi.org/10.1080/1475939X.2014.945474

Popov, V., Brinkman, D., Biemans, H. J. A., Mulder, M., Kuznetsov, A., & Noroozi, O. (2012). Multicultural student group work in higher education. An explorative case study on challenges as perceived by students. *International Journal of Intercultural Relations*, *36*(2), 302–317. http://doi.org/10.1016/j.ijintrel.2011.09.004

Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: the students' perspective. *Assessment & Evaluation in Higher Education*, *33*(2), 143–154. http://doi.org/10.1080/02602930601127869

Price, L. (2006). Gender differences and similarities in online courses: Challenging stereotypical views of women. *Journal of Computer Assisted Learning*, *22*(5), 349–359. http://doi.org/10.1111/j.1365-2729.2006.00181.x

Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles, and preferences. *Advances in Health Sciences Education*, *11*(3), 289–303. http://doi.org/10.1007/s10459-005-3250-z

Prins, F. J., Sluijsmans, D. M. A., Kirschner, P. A., & Strijbos, J. (2005). Formative peer assessment in a CSCL environment: a case study. *Assessment & Evaluation in Higher Education*, *30*(4), 417–444. http://doi.org/10.1080/02602930500099219

Read, B., Francis, B., & Robson, J. (2005). Gender, 'bias', assessment and feedback: analyzing the written assessment of undergraduate history essays. *Assessment & Evaluation in Higher Education*, *30*(3), 241–260. http://doi.org/10.1080/02602930500063827

Reeder, K., Macfadyen, L. P., Roche, J., & Chase, M. (2004). Negotiating cultures in cyberspace: Participation patterns and problematics. *Language Learning & Technology*, *8*(2), 88–105. http://doi.org/10.14288/1.0058429

Reilly, E. D., Williams, K. M., Stafford, R. E., Corliss, S. B., Walkow, J. C., & Kidwell, D. K. (2016). Global times call for global measures: Investigating automated essay scoring in linguistically-diverse MOOCs. *Journal of Asynchronous Learning Network*, *20*(2). http://doi.org/10.24059/olj.v20i2.638

Risman, B. (2004). Gender as a social structure - Theory Wrestling with Activism. *Gender & Society*, *18*(4), 429–450. http://doi.org/10.1177/0891243204265349

Risman, B., & Davis, G. (2012). The birth and evolution of biological theories for sex difference. *Current Sociology*, *61*(5–6), 733–755. http://doi.org/10.1177/205684601271

Rodriguez, O. C. (2012). MOOCs and the AI-Stanford Like Courses: Two Successful and Distinct Course Formats for Massive Open Online Courses. *European Journal of Open, Distance and E-Learning*, *15*(2). Retrieved from http://www.eurodl.org/materials/contrib/2012/Rodriguez.pdf

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding Tutor Learning: Knowledge-Building and Knowledge-Telling in Peer Tutors' Explanations and Questions. *Review of Educational Research*, *77*(4), 534–574. http://doi.org/10.3102/0034654307309920

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Methodological Issues in the Content Analysis of Computer Conference Transcripts. *International Journal of Artificial Intelligence in Education (IJAIED)*, *12*(1), 8–22.

Rovai, A., & Baker, J. (2005). Gender differences in online learning. *Quarterly Review of Distance Education*, *6*, 31–44.

Rovai, A. P. (2001). Building classroom community at a distance: A case study. *Educational Technology Research and Development*, *49*(4), 33–48. http://doi.org/10.1007/BF02504946

Rushton, C., Ramsey, P., & Rada, R. (1993). Peer Assessment in a Collaborative Hypermedia Environment: A Case Study. *Journal of Computer-Based Instruction*, *20*(3), 75–80.

Sadler, D. R. (1998). Formative Assessment: revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77–84. http://doi.org/10.1080/0969595980050104

Sanchez Gordon, S., & Luján Mora, S. (2014). Moocs gone wild. In *Proceedings of the 8th International Technology, Education and Development Conference (INTED 2014)* (pp. 1449–1458). Retrieved from http://library.iated.org/view/SANCHEZGORDON2014MOO

Saunders, W. M. (1989). Collaborative writing tasks and peer interaction. *International Journal of Educational Research*, *13*(1), 101–112. http://doi.org/10.1016/0883-0355(89)90019-0

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, *78*(1), 153–189. http://doi.org/10.3102/0034654307313795

Sluijsmans, D. M. A., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, *27*(5), 443–454. http://doi.org/10.1080/026029302200000931

Stake, R. (1995). *The Art of Case Study Research*. Thousand Oaks, CA: Sage Publications.

Stan, L. (2010). Content Analysis. In A. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of Case Study Research* (pp. 226–230). Thousand Oaks, CA: Sage Publications. http://doi.org/10.4135/9781412957397.n85

Strijbos, J.-W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning and Instruction*, *20*(4), 291–303. http://doi.org/10.1016/j.learninstruc.2009.08.008

Strijbos, J.-W., Pat-El, R. J., & Narciss, S. (2010). Validation of a (peer) feedback perceptions questionnaire. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, & T. Ryberg (Eds.), *Proceedings of the 7th International Conference on Networked Learning 2010* (pp. 378–386). Retrieved from http://www.networkedlearningconference.org.uk/past/nlc2010/abstracts/PDFs/Strijbos.pdf

Strijbos, J.-W., & Sluijsmans, D. (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction*, *20*(4), 265–269. http://doi.org/10.1016/j.learninstruc.2009.08.002

Suen, H. K. (2014). Peer assessment for massive open online courses (MOOCs). *International Review of Research in Open and Distance Learning*, *15*(3), 312–327.

Topping, K. (1998). Peer Assessment Between Students in Colleges and Universities. *Review of Educational Research*, *68*(3), 249–276.

Topping, K. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, *20*(4), 339–343. http://doi.org/10.1016/j.learninstruc.2009.08.003

Topping, K., & Ehly, S. (1998). *Peer assisted learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tsai, C.-C., Lin, S. S. ., & Yuan, S.-M. (2002). Developing science activities through a networked peer assessment system. *Computers & Education*, *38*(1–3), 241–252. http://doi.org/10.1016/S0360-1315(01)00069-0

Tseng, S.-C., & Tsai, C.-C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education*, *49*(4), 1161–1174. http://doi.org/10.1016/j.compedu.2006.01.007

Tuzi, F. (2001). E-Feedback's impact in ESL writers' revisions. In *Annual meeting of Teachers of English Speakers and Other Languages (TESOL)*. St. Louis, MO. http://doi.org/ED419696

Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition*, *21*(2), 217–235. http://doi.org/10.1016/j.compcom.2004.02.003

Urrutia, M., Fielding, S., & White, S. (2016). Professional Development Through MOOCs in Higher Education Institutions: Challenges and Opportunities for PhD Students Working as Mentors. *Journal of Interactive Media*, *1*, 1–10. http://doi.org/10.5334/jime.427

van den Berg, I., Admiraal, W., & Pilot, A. (2006). Designing student peer assessment in higher education: analysis of written and oral peer feedback. *Teaching in Higher Education*, *11*(2), 135–147. http://doi.org/10.1080/13562510500527685

van der Pol, J., van den Berg, B. A. M., Admiraal, W. F., & Simons, P. R. J. (2008). The nature, reception, and use of online peer feedback in higher education. *Computers and Education*, *51*(4), 1804–1817. http://doi.org/10.1016/j.compedu.2008.06.001

van Zundert, M., Sluijsmans, D., & van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, *20*(4), 270–279. http://doi.org/10.1016/j.learninstruc.2009.08.004

Veletsianos, G. (2013). *How do learners experience open online learning?* (G. Veletsianos, Ed.)*Learner Experiences with MOOCs and Open Online Learning*. Madison, WI: Hybrid Pedagogy. Retrieved from http://learnerexperiences.hybridpedagogy.com

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

Veletsianos, G., & Shepherdson, P. (2016). A Systematic Analysis and Synthesis of the Empirical MOOC Literature Published in 2013 – 2015. *International Review of Research in Open and Distributed Learning*, *17*(2), 198–221. http://doi.org/10.19173/irrodl.v17i2.2448

Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning? *Assessment & Evaluation in Higher Education*, *34*(2), 221–230. http://doi.org/10.1080/02602930801955986

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Waite, M., Mackness, J., Roberts, G., & Lovegrove, E. (2013). Liminal participants & skilled orienteers: A case study of learner participation in a MOOC for new lecturers. *Journal of Online Learning & Teaching*, *9*(2), 200–215. http://doi.org/10.2991/icssr-14.2014.24

Warner, S. (1965). Randomized Response : A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, *60*(309), 63–69.

Warschauer, M. (2003). *Technology and Social Inclusion: Rethinking the Digital Divide*. Cambridge, MA: MIT Press.

Wenger, E., McDermott, R., & Snyder, W. (2002). *Cultivating communities of practice: a guide to managing knowledge*. Boston, Mass.: Harvard Business School Press.

Wertsch, J. V. (1991). *Voices of the mind: a sociocultural approach to mediated action*. Brighton: Harvester.

West, C., & Zimmermann, D. H. (1987). Doing Gender. *Gender and Society*, *1*(2), 125–151.

Wishart, C., & Guy, R. (2009). Analyzing Responses, Moves, and Roles in Online Discussions. *Interdisciplinary Journal of E-Learning and Learning Objects*, *5*, 129–144.

Yin, R. (1994). *Case study research: Design and methods*. *Design and methods*. Thousand Oaks: CA: Sage Publications.

Yin, R. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Yuan, L., & Powell, S. (2013). *MOOCs and open education: Implications for higher education*. *Cetis White Paper*. Bolton: University of Bolton: Centre for Educational Technology, Interoperability and Standards & JISC. Retrieved from http://publications.cetis.org.uk/wp-content/uploads/2013/03/MOOCs-and-Open-Education.pdf

Yukselturk, E., & Bulut, S. (2009). Gender differences in self-regulated online learning environment. *Journal of Educational Technology & Society*, *12*(3), 12–22.

Zheng, S., Rosson, M. B., Shih, P. C., & Carroll, J. M. (2015). Understanding Student Motivation, Behaviors and Perceptions in MOOCs. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15* (pp. 1882–1895). http://doi.org/10.1145/2675133.2675217

Appendix One Supportive actors within LDS' MOOCs

The MOOC Facilitator: undertakes the moderation of the course at a content and/or at an organisational level, acting as a community manager.

Mentors: are knowledgeable in the topics of the course, are familiar with the course's didactical concept, and are responsible for encouraging, continuously supporting and evaluating the work of a specified number of groups. They are expected to participate, initiate and moderate discussions in the forum.

Tutors: are students' first contact persons, whose role is to support a specified number of groups in all matters related to organisation, operation of the course, and issues around online collaboration.

Teachers: They prepare and provide the content for the course and take part actively in the forum discussions.

Speakers: are scholars and experts who provide inspirational input in the form of short video lectures. They may join discussions initiated in the forum.

Appendix Two Descriptions and examples of original coding system

Appendix Table 1 Types of peer-feedback in a cMOOC - Coding system (López-Benavides, 2015)

Feedback aspects		
Name	Description	Example
Content	Includes the relevance of information, the clarity of the problem, the argumentation, and the explanation of concepts.	"The in depth background information sometimes distracts the reader from the main subject"
Structure	Means the inner consistency of a text, for example the relation between the main problem and the specified research questions, or between the argumentation and the conclusion.	"Your assignment flows quite well and you followed the requirements of the assignment"
Style	Refers to the 'outer' form of the text, which includes use of language, grammar, spelling and layout.	"Methodical use of visuals and appropriate textual style"
Feedback functions		
Cognitive		
Analysis	Includes comments aimed at understanding the text	"I'm not sure how this relates to graphene or the other derived innovation"
Evaluation	Refers to all explicit and implicit quality statements	"… your innovation is not mentioned at all in the text. What kind of innovation in airline company do you have in mind?"
Explanation	Refers to arguments supporting the evaluation	"Some points did not link together such as what are the skills needed to produce graphene windshields"
Revision	Refers to suggested measures for improvement	"… start with a short introduction of your product before discussing the skills"

Table 1 Types of peer-feedback in a cMOOC - Coding system (López-Benavides, 2015) (continued)

Motivational		
Name	Description	Example
Praise	Positive and emotive comments related to features in the text	"This is the best submission I have ever read!"
Mitigating Language	Includes both positive and negative comments, and uses mitigation language to make negative criticism sound more sensitive	"Just probably a bit too long, otherwise excellent"
Inflammatory Language	Comprises sarcastic, ironic or offensive comments that are not necessarily constructive	"it [the submission] actually shows that you are bloody amateurs in the field of medical devices"
Not Applicable	Comments that do not relate to either the motivational or cognitive functions of feedback.	I really don't have any suggestions".

Appendix Three Examples of the types of feedback used in the set of guidelines.

Content Aspects

Following the three criteria as presented in the course:

Relevance: Especially those parameters mentioned in the first part of the Mbox reflect the risks of the situation of the case analysed.

Substance: The parameters in your Mbox are practical and can be directly used in the current situation of the case analysed.

Clarity and coherence: The descriptions of the parameters in your Mbox are clear and easy to follow.

Presentation aspects

To improve the work presentation-wise:

Style: The use of colours would help to organise the parameters in your Mbox

Structure: A short introduction could be placed before presenting the Mbox

Motivational function

Praise

general: I really like your work

content: I really like the description of the parameters in the Mbox

presentation: I really like the colours used to highlight the parameters in the Mbox

Mitigating Language

general: The work is good but long

content: The parameters in your Mbox are well described, but a bit outdated

presentation: The parameters in your Mbox are well placed, but don't look

organised.

Criticism

general: It seems incomplete

content: I think the parameters mentioned in your Mbox don't make any sense

presentation: I think the colours you used to list the parameters in your Mbox don't match.

Inflammatory language

general: What is that?

content: Don't tell me that this is your finished Mbox

presentation: This looks like a draft submission

Cognitive function

To improve the work content-wise:

Evaluation: Why did you place parameter XX in the second part of the Mbox and not in the first one, where a reference to the case is done?

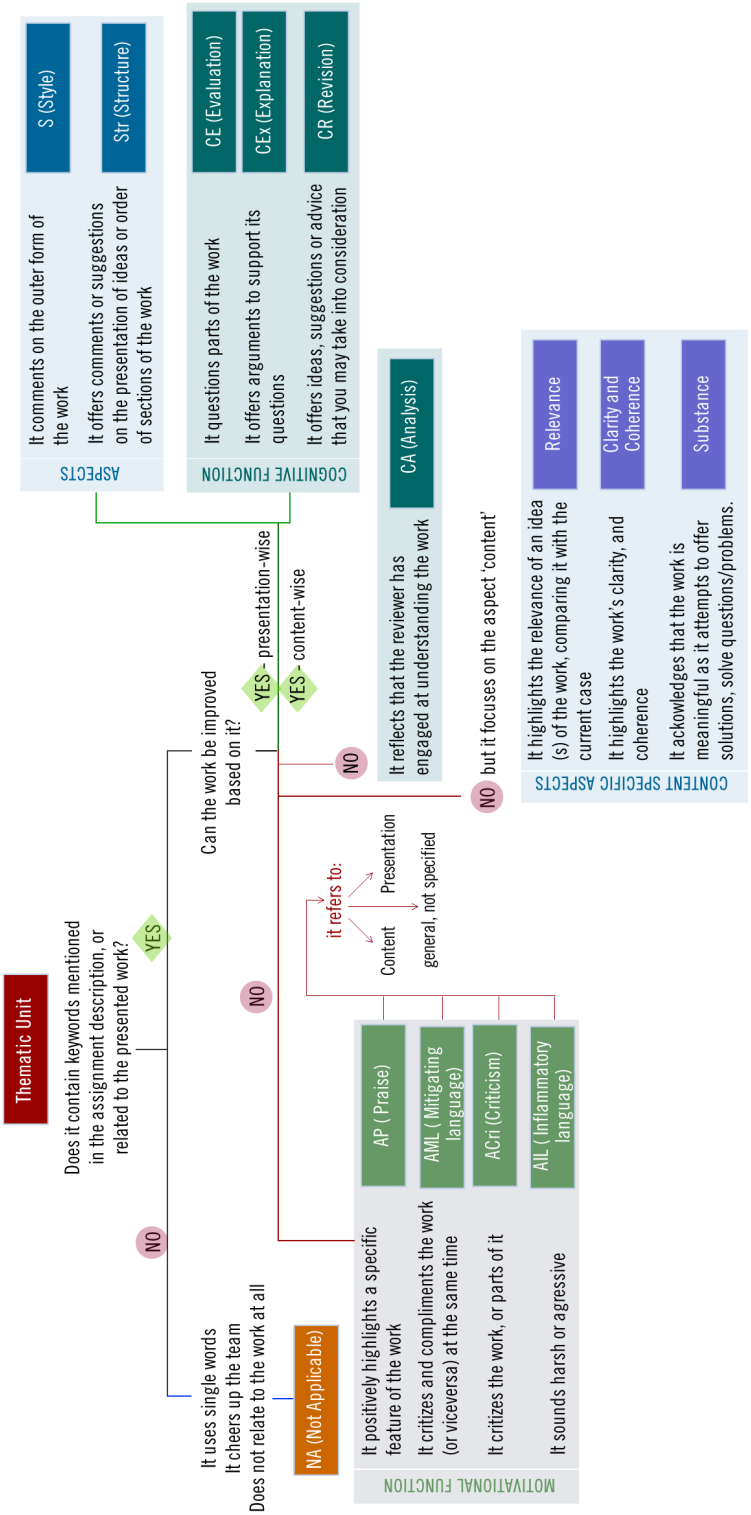
Explanation: If parameter XX had been placed in the first part of the box, then the relation to the case would become clearer.

Revision: You might want to include the fourth parameter called YYY.

Analysis: I now understand why parameter X needed to be mentioned before parameter YY, and also how the conclusion of including the risks in the Mbox was important.

Note: *Explanation* was not included as an independent category in the final coding system. Analysis showed that in most cases when comments referred to evaluation an explanation supporting that evaluation was provided. Therefore, a separation did not appear to be useful or meaningful.

Appendix Four Flow diagram created to support evaluators during the process of inter-rater reliability



Appendix Figure 1 Flow diagram created to support evaluators during the process of inter-rater reliability

Appendix Five Differences for feedback types. Student versus Supporter

Appendix Table 2 Differences for feedback types Student versus Supporter

	Students (*n* = 147)	Supporters (*n* = 88)			
Feedback	Frequency of use (± SEM)		*p*	*OR*	95% *CI OR*
Cognitive					
Analysis	3.34 (± 0.08)	5.05 (± 0.24)	**<0.001**	0.66	0.58 - 0.75
Evaluation	2.11 (± 0.13)	2.51 (± 0.18)	0.0967	0.83	0.69 - 1.01
Revision	1.96 (± 0.12)	3.06 (± 0.19)	**<.0001**	0.64	0.53 - 0.76
Motivation					
Criticism					
content	1.74 (± 0.12)	1.70 (± 0.16)	**0.0209**	1.02	0.81 - 1.27
general	1.17 (± 0.13)	2.29 (± 0.26)	**<.0001**	0.51	0.37 - 0.69
presentation	1.46 (± 0.12)	1.48 (± 0.16)	0.1339	0.98	0.75 - 1.27
Inflammatory Language					
content	.25 (± 0.07)	.26 (± 0.08)	0.1472	0.97	0.43 - 2.20
general	1.15 (± 0.24)	1.71 (± 0.49)	0.3561	0.67	0.33 - 1.34
presentation	1.46 (± 0.12)	1 (± 0.38)	0.4835	1.22	0.47 - 3.15
Mitigating Language					
content	1.78 (± 0.11)	2 (± 0.17)	0.1047	0.88	0.72 - 1.09
general	1.36 (± 0.13)	1.33 (± 0.18)	0.1660	1.02	0.73 - 1.41
presentation	1.41 (± 0.11)	1.41 (± 0.15)	0.1301	0.99	0.77 - 1.28
Praise					
content	3.07 (± 0.14)	4.05 (± 0.22)	**0.0001**	0.75	0.65 - 0.87
general	2.91 (± 0.14)	4.73 (± 0.23)	**<.0001**	0.61	0.53 - 0.70
presentation	2.86 (± 0.14)	3.46 (± 0.20)	0.0760	0.82	0.71 - 0.96
Aspects					
Content					
Relevance	1.98 (± 0.12)	2.10 (± 0.16)	0.5560	0.94	0.77 - 1.14
Clarity and Coherence	1.51 (± 0.12)	1.81 (± 0.17)	0.1312	0.83	0.65 - 1.05
Substance	1.66 (± 0.12)	2.03 (± 0.17)	0.0687	0.81	0.65 - 1.01
Presentation					
Style	1.49 (± .12)	1.58 (± 0.17)	0.1333	0.94	0.72 - 1.22
Structure	1.25 (± .12)	1.66 (± 0.18)	0.1446	0.75	0.56 - 1.00

Note: Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval.*

Appendix Six Differences for feedback functions and aspects. Female versus Male

Appendix Table 3 Differences for feedback functions and aspects. Female versus Male.

	Female (*n* = 174)	Male (*n* = 44)			
Feedback	Frequency of use (± SEM)		*p*	*OR*	95% *CI OR*
Functions					
Cognitive	2.88 (± 0.08)	3.03 (± 0.16)	0.4025	0.95	0.84 - 1.06
Motivation	2.25 (± 0.04)	2.12 (± 0.08)	0.1639	1.05	0.97 - 1.14
Aspects					
Content	1.91 (± 0.07)	1.64 (± 0.13)	0.0644	1.16	0.99 - 1.38
Presentation	1.44 (± 0.08)	1.51 (± 0.17)	0.6689	0.94	0.74 - 1.21

Note: Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval.*

Appendix Seven Differences for feedback functions and aspects. Gender within roles

Appendix Table 4 Differences for feedback functions and aspects in gender within roles.

Feedback	Freq. of use (± SEM)		Freq. of use (± SEM)	*p*	*OR*	95% *CI OR*
Functions						
Cognitive						
Fem stu	2.45 (± 0.09)	Fem sup	3.49 (± 0.13)	**<.0001**	0.7	0.63 - 0.78
Fem stu	2.45 (± 0.09)	Male stu	2.71 (± 0.18)	0.1985	0.9	0.77 - 1.05
Fem stu	2.45 (± 0.09)	Male sup	3.75 (± 0.32)	**<.0001**	0.65	0.54 - 0.78
Fem sup	3.49 (± 0.13)	Male stu	2.71 (± 0.18)	**0.001**	1.28	1.1 - 1.49
Fem sup	3.49 (± 0.13)	Male sup	3.75 (± 0.32)	0.4449	0.93	0.77 - 1.11
Male stu	2.71 (± 0.18)	Male sup	3.75 (± 0.32)	0.0027	0.72	0.58 - 0.89
Motivational						
Fem stu	2.01 (± 0.05)	Fem sup	2.61 (± 0.07)	**<.0001**	0.76	0.71 - 0.82
Fem stu	2.01 (± 0.05)	Male stu	2.01 (± 0.09)	0.964	0.99	0.9 - 1.1
Fem stu	2.01 (± 0.05)	Male sup	2.39 (± 0.16)	**0.0145**	0.84	0.73 - 0.96
Fem sup	2.61 (± 0.07)	Male stu	2.01 (± 0.09)	**<.0001**	1.29	1.16 - 1.43
Fem sup	2.61 (± 0.07)	Male sup	2.39 (± 0.16)	0.2064	1.09	0.95 - 1.25
Male stu	2.01 (± 0.09)	Male sup	2.39 (± 0.16)	**0.034**	0.84	0.72 - 0.98
Aspects						
Content						
Fem stu	1.80 (± 0.08)	Fem sup	2.08 (± 0.11)	**0.0416**	0.86	0.75 - 0.99
Fem stu	1.80 (± 0.08)	Male stu	1.58 (± 0.15)	0.2095	1.14	0.92 - 1.4
Fem stu	1.80 (± 0.08)	Male sup	1.76 (± 0.23)	0.8767	1.0	0.77 - 1.34
Fem sup	2.08 (± 0.11)	Male stu	1.58 (± 0.15)	**0.0109**	1.31	1.06 - 1.63
Fem sup	2.08 (± 0.11)	Male sup	1.76 (± 0.23)	0.2472	1.18	0.89 - 1.56
Male stu	1.58 (± 0.15)	Male sup	1.76 (± 0.23)	0.4993	0.89	0.65 - 1.23
Presentation						
Fem stu	1.36 (± 0.10)	Fem sup	1.54 (± 0.10)	0.2736	0.88	0.70 - 1.1
Fem stu	1.36 (± 0.10)	Male stu	1.48 (± 0.19)	0.6015	0.92	0.68 - 1.24
Fem stu	1.36 (± 0.10)	Male sup	1.63 (± 0.35)	0.4356	0.83	0.53 - 1.31
Fem sup	1.54 (± 0.10)	Male stu	1.48 (± 0.19)	0.778	1.04	0.76 - 1.42
Fem sup	1.54 (± 0.10)	Male sup	1.63 (± 0.35)	0.811	0.94	0.59 - 1.49
Male stu	1.48 (± 0.19)	Male sup	1.63 (± 0.35)	0.694	0.9	0.54 - 1.48

Note: Fem stu (female student, *n* = 104), Fem sup (female supporter, *n* = 70), Male stu (male student, *n* = 30), Male sup (male supporter, *n* = 14). Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval.*

Appendix Eight Differences for feedback types - cognitive function. Gender within roles

Appendix Table 5 Differences for feedback types - cognitive function in gender within roles.

Feedback	Freq. of use (± SEM)		Freq. of use (± SEM)	*p*	*OR*	95% *CI OR*
Functions						
Cognitive						
Analysis						
Fem stu	3.27 (± 0.18)	Fem sup	4.69 (± 0.27)	**<.0001**	0.70	0.60 - 0.81
Fem stu	3.27 (± 0.18)	Male stu	3.68 (± 0.36)	0.2853	0.89	0.71 - 1.10
Fem stu	3.27 (± 0.18)	Male sup	5.43 (± 0.63)	**<.0001**	0.60	0.47 - 0.77
Fem sup	4.69 (± 0.27)	Male stu	3.68 (± 0.36)	**0.0305**	1.27	1.02 - 1.58
Fem sup	4.69 (± 0.27)	Male sup	5.43 (± 0.63)	0.247	0.86	0.67 - 1.11
Male stu	3.68 (± 0.36)	Male sup	5.43 (± 0.63)	**0.0095**	0.68	0.50 - 0.91
Evaluation						
Fem stu	2.00 (± 0.16)	Fem sup	2.51 (± 0.21)	**0.0457**	0.80	0.64 - 1.00
Fem stu	2.00 (± 0.16)	Male stu	2.50 (± 0.31)	0.1317	0.80	0.60 - 1.07
Fem stu	2.00 (± 0.16)	Male sup	2.67 (± 0.49)	0.1541	0.75	0.51 - 1.11
Fem sup	2.51 (± 0.21)	Male stu	2.50 (± 0.31)	0.9727	1.01	0.75 - 1.34
Fem sup	2.51 (± 0.21)	Male sup	2.67 (± 0.49)	0.7681	0.94	0.63 - 1.40
Male stu	2.50 (± 0.31)	Male sup	2.67 (± 0.49)	0.7716	0.94	0.61 - 1.45
Revision						
Fem stu	1.96 (± 0.18)	Fem sup	3.12 (± 0.22)	**<.0001**	0.63	0.51 - 0.77
Fem stu	1.96 (± 0.18)	Male stu	1.90 (± 0.26)	0.8423	1.03	0.76 - 1.40
Fem stu	1.96 (± 0.18)	Male sup	2.77 (± 0.48)	0.0655	0.71	0.49 - 1.02
Fem sup	3.12 (± 0.22)	Male stu	1.90 (± 0.26)	**0.0013**	1.64	1.21 - 2.22
Fem sup	3.12 (± 0.22)	Male sup	2.77 (± 0.48)	0.527	1.13	0.78 - 1.62
Male stu	1.90 (± 0.26)	Male sup	2.77 (± 0.48)	0.087	0.69	0.44 - 1.06

Note: Fem stu (female student, *n* = 104), Fem sup (female supporter, *n* = 70), Male stu (male student, *n* = 30), Male sup (male supporter, *n* = 14). Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval*.

Appendix Nine Differences for feedback types - motivational function. Gender within roles

Appendix Table 6 Differences for feedback types - motivational function in gender within roles.

Feedback	Freq. of use (± SEM)		Freq. of use (± SEM)	*p*	*OR*	95% *CI OR*
Functions						
Motivational Criticism content						
Fem stu	1.68 (± 0.14)	Fem sup	1.70 (± 0.17)	0.9213	0.99	0.76 - 1.28
Fem stu	1.68 (± 0.14)	Male stu	1.90 (± 0.27)	0.4528	0.88	0.64 - 0.77
Fem stu	1.68 (± 0.14)	Male sup	1.85 (± 0.43)	0.694	0.91	0.56 - 1.47
Fem sup	1.70 (± 0.17)	Male stu	1.90 (± 0.27)	0.5294	0.90	0.64 - 1.26
Fem sup	1.70 (± 0.17)	Male sup	1.85 (± 0.43)	0.7409	0.92	0.56 - 1.51
Male stu	1.90 (± 0.27)	Male sup	1.85 (± 0.43)	0.9233	1.03	0.60 - 1.75
general						
Fem stu	1.21 (± 0.15)	Fem sup	2.48 (± 0.29)	**<.0001**	0.49	0.35 - 0.68
Fem stu	1.21 (± 0.15)	Male stu	1.07 (± 0.27)	0.6613	1.13	0.65 - 1.96
Fem stu	1.21 (± 0.15)	Male sup	1.33 (± 0.67)	0.8452	0.90	0.33 - 2.48
Fem sup	2.48 (± 0.29)	Male stu	1.07 (± 0.27)	**0.0001**	1.82	1.35 - 2.46
Fem sup	2.48 (± 0.29)	Male sup	1.33 (± 0.67)	**0.0038**	1.62	1.17 - 2.24
Male stu	1.07 (± 0.27)	Male sup	1.33 (± 0.67)	0.6898	0.80	0.27 - 2.39
presentation						
Fem stu	1.36 (± 0.13)	Fem sup	1.53 (± 0.18)	0.4417	0.89	0.66 - 1.20
Fem stu	1.36 (± 0.13)	Male stu	1.68 (± 0.29)	0.2928	0.81	0.55 - 1.20
Fem stu	1.36 (± 0.13)	Male sup	1.31 (± 0.38)	0.8892	1.04	0.57 - 1.91
Fem sup	1.53 (± 0.18)	Male stu	1.68 (± 0.29)	0.6599	0.91	0.61 - 1.37
Fem sup	1.53 (± 0.18)	Male sup	1.31 (± 0.38)	0.6094	1.17	0.63 - 2.17
Male stu	1.68 (± 0.29)	Male sup	1.31 (± 0.38)	0.4569	1.29	0.66 - 2.50
Inflammatory language Content						
Fem stu	0.24 (± 0.08)	Fem sup	0.30 (± 0.10)	0.5907	0.78	0.32 - 1.91
Fem stu	0.24 (± 0.08)	Male stu	0.32 (± 0.17)	0.64	0.75	0.22 - 2.54
Fem stu	0.24 (± 0.08)	Male sup	0.00 (± 0.00)	0.9685	77309.00	0.00 - ∞
Fem sup	0.30 (± 0.10)	Male stu	0.32 (± 0.17)	0.9374	0.95	0.28 - 3.22
Fem sup	0.30 (± 0.10)	Male sup	0.00 (± 0.00)	0.9678	98639.00	0.00 - ∞
Male stu	0.32 (± 0.17)	Male sup	0.00 (± 0.00)	0.9677	103571.0 0	0.00 - ∞

Note: Fem stu (female student, *n* = 104), Fem sup (female supporter, *n* = 70), Male stu (male student, *n* = 30), Male sup (male supporter, *n* = 14). Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval*.

Table 8 Differences for feedback types - motivational function in gender within roles (continued).

Feedback	Freq. of use (± SEM)		Freq. of use (± SEM)	*p*	*OR* *OR*	95% *CI*
Functions						
Motivational						
Inflammatory language general						
Fem stu	1.08 (± 0.30)	Fem sup	1.71 (± 0.49)	0.2517	0.63	0.29 - 1.39
Fem stu	1.08 (± 0.30)	Male stu	1.40 (± 0.53)	0.5844	0.77	0.31 - 1.94
Fem stu	1.08 (± 0.30)	Male sup	-	0.6703	1.22	0.48 - 3.11
Fem sup	1.71 (± 0.49)	Male stu	1.40 (± 0.53)	0.3506	0.71	0.55 - 5.32
Fem sup	1.71 (± 0.49)	Male sup	-	-	-	
Male stu	1.40 (± 0.53)	Male sup	-	-	-	
presentation						
Fem stu	1.00 (± 0.50)	Fem sup	1.00 (± 0.41)	1	1.00	0.28 - 3.55
Fem stu	1.00 (± 0.50)	Male stu	1.50 (± 0.61)	0.53	0.67	0.19 - 2.36
Fem stu	1.00 (± 0.50)	Male sup	1.00 (± 1.00)	1	1.00	0.11 - 8.95
Fem sup	1.00 (± 0.41)	Male stu	1.50 (± 0.61)	0.4826	0.67	0.21 - 2.07
Fem sup	1.00 (± 0.41)	Male sup	1.00 (± 1.00)	1	1.00	0.12 - 8.31
Male stu	1.50 (± 0.61)	Male sup	1.00 (± 1.00)	0.7074	1.50	0.18 - 12.4
Mitigating language content						
Fem stu	1.75 (± 0.13)	Fem sup	2.05 (± 0.19)	0.1917	0.85	0.67 - 1.08
Fem stu	1.75 (± 0.13)	Male stu	1.93 (± 0.26)	0.5287	0.91	0.67 - 1.23
Fem stu	1.75 (± 0.13)	Male sup	2.09 (± 0.40)	0.3865	0.84	0.56 - 1.25
Fem sup	2.05 (± 0.19)	Male stu	1.93 (± 0.26)	0.7099	1.06	0.77 - 1.46
Fem sup	2.05 (± 0.19)	Male sup	2.09 (± 0.40)	0.9214	0.98	0.64 - 1.49
Male stu	1.93 (± 0.26)	Male sup	2.09 (± 0.40)	0.7267	0.92	0.58 - 1.46
general						
Fem stu	1.29 (± 0.16)	Fem sup	2.05 (± 0.16)	0.813	0.96	0.66 - 1.39
Fem stu	1.29 (± 0.16)	Male stu	1.93 (± 0.26)	0.548	0.87	0.55 - 1.38
Fem stu	1.29 (± 0.16)	Male sup	1.25 (± 0.46)	0.9421	1.03	0.48 - 2.19
Fem sup	2.05 (± 0.16)	Male stu	1.93 (± 0.26)	0.6959	0.91	0.56 - 1.47
Fem sup	2.05 (± 0.16)	Male sup	1.25 (± 0.46)	0.8528	1.08	0.50 - 2.32
Male stu	1.93 (± 0.26)	Male sup	1.25 (± 0.46)	0.6848	1.18	0.52 - 2.68

Note: Fem stu (female student, *n* = 104), Fem sup (female supporter, *n* = 70), Male stu (male student, *n* = 30), Male sup (male supporter, *n* = 14). Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval*.

Table 8 Differences for feedback types - motivational function in gender within roles (continued).

Feedback	Freq. of use (± SEM)		Freq. of use (± SEM)	*p*	*OR*	95% *CI OR*
Functions						
Motivational						
Mitigating language presentation						
Fem stu	1.40 (± 0.13)	Fem sup	1.43 (± 0.16)	0.8984	0.98	0.74 - 1.31
Fem stu	1.40 (± 0.13)	Male stu	1.52 (± 0.25)	0.6792	0.92	0.64 - 1.34
Fem stu	1.40 (± 0.13)	Male sup	1.35 (± 0.39)	0.9028	1.04	0.57 - 1.87
Fem sup	1.43 (± 0.16)	Male stu	1.52 (± 0.25)	0.7662	0.94	0.64 - 1.40
Fem sup	1.43 (± 0.16)	Male sup	1.35 (± 0.39)	0.8571	1.06	0.58 - 1.94
Male stu	1.52 (± 0.25)	Male sup	1.35 (± 0.39)	0.7277	1.12	0.59 - 2.15
Praise content						
Fem stu	3.24 (± 0.18)	Fem sup	4.23 (± 0.25)	**0.0008**	0.76	0.65 - 0.89
Fem stu	3.24 (± 0.18)	Male stu	2.74 (± 0.30)	0.1771	1.18	0.93 - 1.50
Fem stu	3.24 (± 0.18)	Male sup	3.38 (± 0.49)	0.7802	0.96	0.71 - 1.30
Fem sup	4.23 (± 0.25)	Male stu	2.74 (± 0.30)	**0.0005**	1.54	1.21 - 1.97
Fem sup	4.23 (± 0.25)	Male sup	3.38 (± 0.49)	0.1513	1.25	0.92 - 1.70
Male stu	2.74 (± 0.30)	Male sup	3.38 (± 0.49)	0.2512	0.81	0.57 - 1.16
Praise general						
Fem stu	2.90 (± 0.17)	Fem sup	4.99 (± 0.28)	**<.0001**	0.58	0.50 - 0.68
Fem stu	2.90 (± 0.17)	Male stu	2.94 (± 0.31)	0.8895	0.98	0.78 - 1.25
Fem stu	2.90 (± 0.17)	Male sup	4.24 (± 0.55)	**0.0073**	0.68	0.52 - 0.90
Fem sup	4.99 (± 0.28)	Male stu	2.94 (± 0.31)	**<.0001**	1.70	1.34 - 2.14
Fem sup	4.99 (± 0.28)	Male sup	4.24 (± 0.55)	0.2455	1.18	0.89 - 1.55
Male stu	2.94 (± 0.31)	Male sup	4.24 (± 0.55)	0.0301	0.69	0.50 - 0.97
presentation						
Fem stu	4.24 (± 0.55)	Fem sup	2.96 (± 0.17)	**0.0307**	0.83	0.70 - 0.98
Fem stu	4.24 (± 0.55)	Male stu	3.56 (± 0.23)	0.5401	1.08	0.84 - 1.38
Fem stu	4.24 (± 0.55)	Male sup	2.74 (± 0.31)	0.7872	0.96	0.69 - 1.33
Fem sup	2.96 (± 0.17)	Male stu	3.56 (± 0.23)	**0.042**	1.30	1.01 - 1.67
Fem sup	2.96 (± 0.17)	Male sup	2.74 (± 0.31)	0.41	1.15	0.82 - 1.61
Male stu	3.56 (± 0.23)	Male sup	2.74 (± 0.31)	0.5268	0.88	0.61 - 1.29

Note: Fem stu (female student, *n* = 104), Fem sup (female supporter, *n* = 70), Male stu (male student, *n* = 30), Male sup (male supporter, *n* = 14). Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval*.

Appendix Ten Differences for feedback types in aspects. Gender within roles

Appendix Table 7 Differences for feedback types within feedback aspects in gender within roles.

Feedback	Freq. of use (± SEM)		Freq. of use (± SEM)	*p*	*OR*	95% *CI OR*
Aspects						
Content Relevance						
Fem stu	2.06 ± 0.15	Fem sup	2.14 ± 0.19	0.7229	0.96	0.77 - 1.20
Fem stu	2.06 ± 0.15	Male stu	1.86 ± 0.26	0.5323	1.10	0.81 - 1.50
Fem stu	2.06 ± 0.15	Male sup	2.15 ± 0.41	0.8328	0.96	0.64 - 1.42
Fem sup	2.14 ± 0.19	Male stu	1.86 ± 0.26	0.4025	1.15	0.83 - 1.59
Fem sup	2.14 ± 0.19	Male sup	2.15 ± 0.41	0.9891	1.00	0.66 - 1.50
Male stu	1.86 ± 0.26	Male sup	2.15 ± 0.41	0.549	0.87	0.55 - 1.38
Clarity and Coherence						
Fem stu	1.55 ± 0.14	Fem sup	2.14 ± 0.19	0.1078	0.81	0.62 - 1.05
Fem stu	1.55 ± 0.14	Male stu	1.44 ± 0.26	0.7117	1.08	0.72 - 1.60
Fem stu	1.55 ± 0.14	Male sup	1.44 ± 0.40	0.8104	1.07	0.61 - 1.90
Fem sup	2.14 ± 0.19	Male stu	1.44 ± 0.26	0.1609	1.34	0.89 - 2.01
Fem sup	2.14 ± 0.19	Male sup	1.44 ± 0.40	0.3318	1.33	0.75 - 2.37
Male stu	1.44 ± 0.26	Male sup	1.44 ± 0.40	0.9878	0.99	0.52 - 1.91
Substance						
Fem stu	1.75 ± 0.15	Fem sup	2.16 ± 0.20	0.0898	0.81	0.64 - 1.03
Fem stu	1.75 ± 0.15	Male stu	1.37 ± 0.24	0.216	1.28	0.87 - 1.88
Fem stu	1.75 ± 0.15	Male sup	1.55 ± 0.39	0.6469	1.13	0.67 - 1.91
Fem sup	2.16 ± 0.20	Male stu	1.37 ± 0.24	**0.0233**	1.57	1.06 - 2.33
Fem sup	2.16 ± 0.20	Male sup	1.55 ± 0.39	0.2178	1.39	0.82 - 2.37
Male stu	1.37 ± 0.24	Male sup	1.55 ± 0.39	0.6967	0.89	0.48 - 1.63
Presentation Style						
Fem stu	1.44 ± 0.14	Fem sup	1.52 ± 0.17	0.7293	0.95	0.70 - 1.28
Fem stu	1.44 ± 0.14	Male stu	1.68 ± 0.28	0.4395	0.86	0.59 - 1.26
Fem stu	1.44 ± 0.14	Male sup	1.61 ± 0.51	0.7396	0.89	0.46 - 1.73
Fem sup	1.52 ± 0.17	Male stu	1.68 ± 0.28	0.6333	0.91	0.61 - 1.36
Fem sup	1.52 ± 0.17	Male sup	1.61 ± 0.51	0.8634	0.94	0.48 - 1.84
Male stu	1.68 ± 0.28	Male sup	1.61 ± 0.51	0.9143	1.04	0.51 - 2.12

Note: Fem stu (female student, *n* = 104), Fem sup (female supporter, *n* = 70), Male stu (male student, *n* = 30), Male sup (male supporter, *n* = 14). Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval*.

Table 9 Differences for feedback types within feedback aspects in gender within roles. (continued).

Feedback	Freq. of use (± SEM)		Freq. of use (± SEM)	*p*	*OR*	95% *CI OR*
Aspects						
Presentation structure						
Fem stu	1.26 ± 0.14	Fem sup	1.56 ± 0.19	0.2123	0.81	0.59 - 1.13
Fem stu	1.26 ± 0.14	Male stu	1.23 ± 0.26	0.911	1.03	0.64 - 1.64
Fem stu	1.26 ± 0.14	Male sup	1.64 ± 0.48	0.414	0.77	0.42 - 1.43
Fem sup	1.56 ± 0.19	Male stu	1.23 ± 0.26	0.3391	1.26	0.78 - 2.05
Fem sup	1.56 ± 0.19	Male sup	1.64 ± 0.48	0.8759	0.95	0.51 - 1.78
Male stu	1.23 ± 0.26	Male sup	1.64 ± 0.48	0.4335	0.75	0.37 - 1.53

Note: Fem stu (female student, *n* = 104), Fem sup (female supporter, *n* = 70), Male stu (male student, *n* = 30), Male sup (male supporter, *n* = 14). Significant at the *p<0.05* level. The values in bold under *p* indicate statistical significance. *OR = odds ratio. CI = confidence interval*.