

Grading *versus* Reliability: how Academia perspectives evaluation on MOOCs

Moura, Ana S.^a, and Natália D.S. Cordeiro, M.^a

^aREQUIMTE, Department of Chemistry and Biochemistry, University of Porto, Portugal

Abstract

Massive Open Online Courses (MOOCs) have experienced in recent years a significant growth in courses'offer and the number of enrolled students. Nevertheless, the controversy regarding if its quality is reliable, namely in student evaluation and assessment, has not found closure. In this study, we aim at establishing an initial prospection of the academic teaching professionals' perspective regarding the quality of the most common/usual evaluation methods and tools used in MOOCs. After the elaboration of a questionnaire and its implementation to an international sample of academic professors, the analysis of the answers allows perceiving which MOOC grading methods are acceptable in presential Higher Education courses and its eventual acceptable weight in the final grade. Further, within certain constraints, a large percentage of the inquired academics presented no problem with the inclusion of MOOC grading methods on their non-online courses. Overall, within those constraints, the academics felt the quality of the academic orthodox courses was maintained, a perspective that can contribute to change eventual suspicious attitudes regarding MOOCs evaluation methodologies and their student assessment.

Keywords: MOOC; Grading; Peer-review; Evaluation

1. Introduction

Massification of educational systems made its debut in the 19th century as education methodologies become the object of academic study, while the pedagogical goals and constraints were subject to political debate (Stray, 2001). Education, as an absolute value, became an increasingly perennial foundation of contemporary society and the desire to have such value reach every citizen, at any age and any time and any location, led to several attempts for its universal diffusion, as the 1922 decision of Higher Education institutions to broadcast lectures through radio exemplifies and media duly noted (Bliven, 1924). As stated by Marquez-Ramos *et al.* “Current and future learning environments involve both physical and virtual scenarios; [...] A remarkable growth in technology-based educational practices is accompanying the traditional educational model.” (Marquez-Ramos, 2017).

However, apart from the controversy and discussion regarding the evolution of Higher Education pedagogics practice in new contexts, since its primordial attempts, the question of assessment quality of such massive educational approaches *versus* those methods of more orthodox academic courses was raised, and at present, under the rise of Massive Open Online Courses (MOOCs), where massification reaches the domestic embedded nature of intimacy only Internet can offer, the question has increased its prominence.

Assessment and grading are always in question and under analysis, whether in orthodox and non-orthodox courses, but as Pessoa *et al.* wrote: “There is yet to be presented a grading or evaluation method that is pristine in principle and flawless in execution.” (Pessoa, 2015). Nevertheless, one can perceive some methods as less vulnerable to fraud and/or more trustworthy in evaluating specific skills. The questions that can be asked in such a reality are: (a) does this perception is sound or skewed?, and (b) to whom does this perception belongs? Regarding question (a), there have been several studies regarding grading methods and its dependability. Moreover, grading methodologies, namely orthodox grading methodologies, such as written tests, can be used as a pedagogical tool for learning in formative evaluation (Nicol and Macfarlane-Dick, 2005). But this study is focused on some ramified aspects of the question (b).

Perception and personal perspective are the basis of any given person’s reality and if, in particular, such person has education as leitmotif for a profession, then they not only are the person’s paradigms for reality but will also affect how the established education shapes the reality of others. The persons traditionally bestowed with the trust of society regarding Higher Education quality are the orthodox academic professors. As such, how they view MOOCs grading methods certainly affects how MOOC assessment is perceived, either as reliable or not. The results presented in this study consist in a preliminary stage for a future and deeper search in order to understand how orthodox educational professors perceive MOOC grading systems and how that perception affects not only their grading methods

(and those of Academia) but the perception of society, students, and employers from Academia and MOOC evaluation.

This work presents the results of a first stage, where an inquiry regarding the opinion of the academics on MOOCs most usual grading methods was performed and its results are presented. The following stage will be the construction and implementation of a new questionnaire, more detailed, in order to scrutinize even further their academic perception on this subject. The specifics of method and sample are presented in Section 2, while the results and their analysis are presented in Section 3. A final section of concluding remarks not only sums up the overall relevant aspects of the study but also indicates future perspectives regarding the following stages of research.

2. Method and sample details

The survey entitled *MOOCs and Evaluation: the POV of Professors*, was developed/written by the authors of this study and conducted using Google Surveys (2018), being delivered directly to the institutional emails of Higher Education teaching professionals from two countries, Portugal and Spain, which represent the authors's geographical and cultural context. The applied survey consisted of two parts, *Personal Details* and *MOOCs grading and evaluation*, with a total of seventeen questions. The questions ranged from multiple choice (10), short text (4), and selecting boxes (3). In *Personal Details*, the questions aimed at defining the sample and could be divided in individual absolute characteristics (e.g. gender, age, etc) and professional experience (e.g. years of teaching practice, etc). The specifics of the sample are indicated in the following subsections 2.1 and 2.2. In *MOOCs grading and evaluation*, the questions explore the opinions and perspective of the surveyed individuals regarding MOOCs grading methodologies. The results of this part will be presented and discussed in Section 3.

2.1. Sample: Personal Details

The total number of surveyees in this study is 32, an accepted value for minimal statistical significance, especially considering this is a prospection stage for developing a more detailed questionnaire and performing it in a wider community of academics. The sample presents a gender distribution of 65.6% male and 34.4% female, and 77.8% having as field of expertise *Natural Sciences* (e.g Physics, Chemistry, Biology, Astronomy, Earth Sciences, Space Sciences), or *Formal Sciences* (eg, Mathematics, Logic, Computer Sciences). On the remaining options, appart from the field of expertise *Psychology and Education* represented by two surveyees, the remaining academic fields of expertise have just one individual. The details displayed in Table 1 and Table 2 present the specifics regarding age and nationality of the surveyees. In the latter, both the current professional nationality and the country of origin were considered.

Table 1. Personal Details of Surveyees: Age

Age/years	%
20-29	6
30-39	3.1
40-49	43.8
50-59	40.6
60-69	6

Table 2. Personal Details of Surveyees: nationality

Countries	Country of origin	Current citizenship
Portugal	78.1%	81.3%
Spain	15.6%	18.7%
Senegal	3.1%	0%
Venezuela	3.1%	0%

2.2. Sample: Professional Details

With exception of one case, all surveyees are active teacher/professor in higher education institutions, with 90.6% affiliated to a public University and two in a private Higher Education institution. Regarding professional experience, 50% have more than twenty years of teaching experience in academic institutions, and an equal percentage of 18.8% have between 6 to 10 or 11 to 20 years of teaching experience on the same context. During those teaching years, 56.2% taught/lecture in only one city in the same country, while the remaining 40.6% indicated the option of having taught on two to four different cities in the same country. Still regarding teaching cultural environment, 78.1% have taught/lecture in only one country, 12.5% in at least two countries and 9.4% in more than two countries. Currently, 6 of the surveyees teach in Spain while the remaining exert the profession in Portugal.

3. Results and Discussion

In this section, the results of the second part of the survey, *MOOCs grading and evaluation*, are presented and discussed. The survey's questions of this part could be divided into three groups. The first group consisted in two questions (*Suppose one of your courses or*

seminars would be transformed in a MOOC. Choose what means of evaluation you consider suitable and If you answer that none of the methods listed in the previous question are reliable, please indicate ONE method you think it is reliable for grading a MOOC) intending to examine which, of an array of usual MOOC grading methodologies, were preferred by the surveyees. The answer admitted more than an option. From the 32 answers, 56.2% chose *Peer-review tasks and multiple choice quizzes*, while 34.4% elected *Peer-review essays and peer-review tasks*. No surveyee chose the option 'none of the listed methods' and therefore the second question was locked for the surveyees.

The second group comprised three questions aiming to determine the 'comfort zone' of the surveyees regarding the inclusion of some MOOC grading methods in orthodox evaluation. The first question, *Inclusion of MOOC grading I: If you had to include online peer-review essays as part of the grading on your formal course, indicate the option with the prospective maximum percentage of the final grade*, 40.6% chose the option corresponding to 30% of the final grade and almost a third of surveyees chose the option corresponding to 25% of the final grade. Only 9.4% would opt for a maximum of 10% of the final grade while the remaining surveyees were comfortable with the inclusion of this grading methodology up to 20% of the final grade.

The second question, *Inclusion of MOOC grading II: If you had to include online multiple option quizzes with limited time as part of the grading on your formal course, indicate the option with the prospective maximum percentage of the final grade*, saw yet again a scenario where the options corresponding respectively to 30% and 25% of the final grade not only have the higher percentage of answers but that value is equal, 31.3%. From the remaining answers, 21.9% chose up to 20% of the final grade and just five would only accept up to 10% of this grading method in the final grade.

The third and final question of this group, *Inclusion of MOOC grading III: If you had to include online peer-review tasks as part of the grading on your formal course, indicate the option with the prospective maximum percentage of the final grade*, had half of the surveyees choosing the option corresponding to 30% of the final grade, and the second choice, corresponding to option of up to 25% of the final grade, had 31.3% of positive answers. Only two would just include up to 10%, with the remaining choosing the option of up to 20%.

The third group consisted of two questions, *Inclusion of MOOC grading IV: Choose the reason(s) for choosing the percentages from the list below* and *Inclusion of MOOC grading V: Would you include MOOC grading methods on your non online courses?*, and aimed at a first understanding the reasons for the choices of the previous groups and if those choices were acceptable in orthodox academic courses. An overwhelming percentage of 68.8% considered that their choices made on the former maximum percentages of inclusion of

MOOC grading methods maintained the overall quality of the final grade and/or minimized eventual fraud from the students. The minimization of fraud as solo reason for choice was chosen by 15.7% of the surveyees, while four of the professors considered these methods did not evaluated adequately the intended academic skills. Finally, 81.3% of the surveyees would include MOOCs grading methods on their non-online courses. From these results and taking in consideration also the data of Section 2, it is clear that for a large percentage of experienced academics the inclusion of some MOOC grading methods in non online courses presents a reasonable possibility, as long as its influence in the final grade is between 20 to 30%. As such, that implies those methods are seen as effective evaluation tools within certain constraints.

One cannot escape comparing these results with those discussed by Pessoa *et al.* on the case of computer programming MOOCs assessment, as almost 80% of the surveyees have as field of expertise Natural and Formal Sciences (Pessoa, 2015). In this study, after analyzing twenty three grading rubrics gathered from several academic institutions around the world (which included countries as culturally diverse as Brasil, China, Canada, or Australia), the three most relevant criteria not only were the most frequent and rated criteria but were also used in peer-assessment as reliable on grading programming skills. It was inferred that online assessment could be perceived as trustworthy when the correcting rubrics were appropriately developed.

In fact, peer assessment, viewed as a quality mark by some authors (Mora, 2012; Kulkarni, 2013), was present in the majority of choices as a well founded grading method for the sample discussed in this work. Further, from the five academics that chose all the options for grading, four considered that their choices maintaining the overall quality of grading without also choosing the option regarding the minimization of fraud. More interestingly, however, is that though only six of the surveyees would not include these methods in their non online courses, just one consistently choose the option of up to 10% of final grade in the several questions of the second group. Four of them had evidence to be comfortable with a percentage equal or higher than 25% of the final grade on at least two of the MOOC grading methods. There would seem to be a contradiction between the confidence of the reliability of the grading methods in theory and when the personal professional practice is applied. This certainly would be of interest to explore in a future stage of research.

Finally, we considered to be pertinent, given the large percentage of surveyees with Natural or Formal Sciences as field of expertise, to compare the perspective solely of these academics per country, *i.e.*, contrasting Portuguese *versus* Spanish academic point of view regarding the inclusion and maximum percentages of MOOCs methodologies on the final grade. This would also allow to cross-reference with the conclusions of Pessoa *et al.*, on the MOOC assessment of programming courses (Pessoa, 2015).

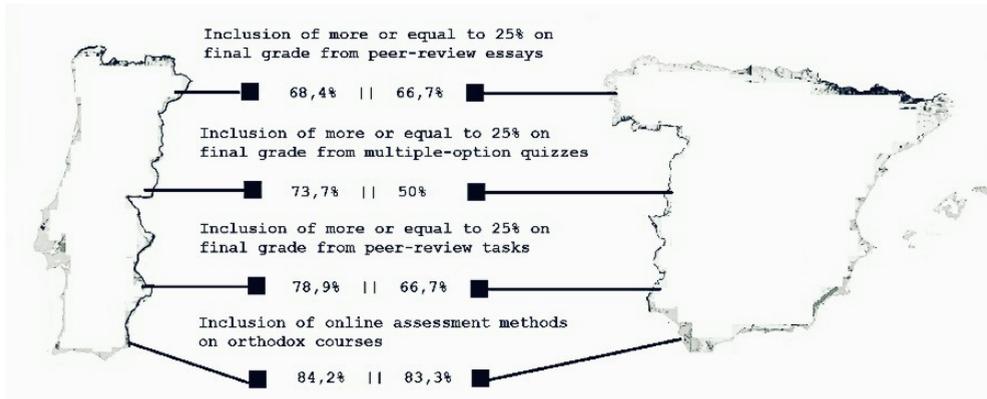


Figure 1. Results regarding only surveyees with Formal or Natural Sciences as field of expertise, for the second and third group of questions. Portugal is represented left, and Spain on the right.

Considering Figure 1, which displays the results of this analysis, it is evident that for the survey sample that is convergence in most of the two nations' perspectives. Considering the inclusion of online peer-review essays assessment equal or higher than 25% of the final grade and expressing favorability to inclusion of online assessment methods on their courses, the percentages are almost the same and expressively represent the opinion of majority of the academic professors. Though the majority is still case for the other two situations (inclusion of more or equal to 25% on the final grade from online assessment through multiple-option quizzes with limited time/peer-review tasks) there is a clear gap between the percentages. On both cases, Portuguese academics present percentages higher than 70% while such mark is never reached in the analogous Spanish sample. There would seem to be certain methodologies perceived as less reliable for the latter, though on the overall the inclusion of MOOCs assessment methods would present no disruption of grading confidence, within the given constraints. This agrees with the conclusions of Pessoa *et al.*, which also noted that the objective nature of the field of expertise could affect the perception of online assessment (Pessoa, 2015).

4. Conclusions

The Internet brought new avenues to the human experience spectra and Education is no exception. Studies, such a quantum-based approach model for the dynamics of higher education/societal transformation interface, are concerned, and rightly so, with the expectations of the professionals, students and general individuals regarding the novel and prospective educational paradigms (Marquez-Ramos, 2017). However, approaches to mass education, such as the ones through television or radio broadcasting, have risen and fall on

the 20th Century. One key factor lies on the reliability of assessment of such courses. Academia stands as the corner-stone for the social confidence on grading and its perspective on the reliability of such assessment methods will affect the success or downfall of online courses. This work aimed at being a preliminary study for understanding the perspective academics have on MOOC grading methods, and from the results it is evidenced that the methods used in MOOC grading are generally seen as acceptable and possible to be included on orthodox courses, as long as they do not represent more than a quarter of the final grade. Future research will aim at exploring the reasons for this constraint and which other online grading methodologies would be acceptable.

Acknowledgments

This work had the financial support of Fundação para a Ciência e a Tecnologia (FCT/MEC) through national funds and co-financed by FEDER, under the Partnership Agreement PT2020 (Projects UID/QUI/50006/2013 and POCI/01/0145/FEDER/007265). To these financing sources the authors are greatly indebted.

References

- Bliven, B. (1924). *The Century Illustrated Monthly Magazine*, 108, pp. 148.
- Kulkarni C., Wei, K. P., Le, H. (2013) Peer and self-assessment in massive online classes [J]. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20 (6), 33
- Márquez-Ramos, L., Mourelle, E. (2018). On the relationship between society and higher education: what path should we take? *Distance Education*, 39 (1) 19-36
- Mora, M. C., Sancho-Bru, J. L. , Iserte, J. L., Sánchez, F. T. (2012) An e-assessment approach for evaluation in engineering overcrowded groups. *Computers & Education*, 59 (2), 732-740
- Pessoa, A. M., Coelho, L., & Fernandes, R. (2015). Massive Open Online Course Management: Learning Science and Engineering through Peer-Reviewed Projects. In R. Queirós (Ed.), *Innovative Teaching Strategies and New Learning Paradigms in Computer Programming* (pp. 89-105). IGI Global
- Stray, C. (2001). The Shift from Oral to Written Examination: Cambridge and Oxford, *Assessment in Education: Principles, Policy & Practices*, 8, (1) 33-50.