Unpacking the Relationship between Discussion Forum Participation and Learning in MOOCs: Content is Key

Alyssa Friend Wise
Learning Analytics Research Network
New York University
239 Greene Street, 2nd Floor
New York, NY 10003 USA
alyssa.wise@nyu.edu

Yi Cui
Learning Analytics Research Network
New York University
239 Greene Street, 2nd Floor
New York, NY 10003 USA
yc65@nyu.edu

ABSTRACT

This study examined the relationship between discussion forum contributions and course assessment results in a statistics MOOC. An important feature of the study is that it distinguished between discussions that were related to the learning of course material ("content-related") and those which were not ("non-content"). Another contribution is that the study evaluated the additional usefulness of social centrality measures in predicting course grade after the quantity of forum contributions has been accounted for. Results showed that, overall, 15% of course learners contributed to the forums and these learners had a significantly higher rate of successfully passing the course than non-contributors (64% vs 32% passing). Learners who made posts to both content-related and non-content threads had a higher passing rate than those who only contributed to one type or the other. Among learners who successfully passed the course, there were no differences in course grade when comparing discussion contributors and non-contributors overall; however those who contributed to content-related threads performed slightly better than those who did not (course grade of 87% vs 85%). A predictive model based on the number of posts made to content-related threads explained a small proportion of variance in course grades; addition of social centrality measures did not significantly improve the variance explained by the model.

CCS CONCEPTS

• **Applied computing~ E-learning**

KEYWORDS

Massive open online courses; discussion forum; learning outcomes; social network analysis

LAK'18, *March 7-9, 2018, Sydney, NSW, Australia*

ACM ISBN 978-1-4503-6400-3/18/03…$15.00
https://doi.org/10.1145/3170358.3170403

ACM Reference Format:
A. F. Wise and Y. Cui. 2018. Unpacking the Relationship between Discussion Forum Participation and Learning in MOOCs: Content is Key. In *LAK'18: International Conference on Learning Analytics and Knowledge, March 7–9, 2018, Sydney, NSW, Australia. ACM, New York, NY, USA*, 10 pages. https://doi.org/10.1145/3170358.3170403.

1 INTRODUCTION

Lack of social interaction and support for emergent needs are often cited as factors that compromise learning experiences and outcomes in Massive Open Online Courses (MOOCs) [14, 15]. In order to help address these concerns, discussion forums are widely provided as venues for interpersonal interaction and are used for various purposes, such as Q&A about assignments and quizzes, discussions of topics related to the course subject, help seeking / giving concerning technical and logistic issues, as well as socializing [26, 32].

Research efforts have investigated connections between forum participation and learning in varied ways, such as examining the correlations between participation quantity and course grade as well as predicting pass / fail based on learners' social centrality in the discussion forums [e.g., 12, 16, 18, 28]. However, findings from these studies do not show a consistent relationship between forum participation and course performance. For example, [17] found a significant correlation between measures of forum social centrality and final grade in an algebra MOOC, but no relationship between the same variables in a finance course. In a different study on a biology MOOC, [18] found that social centrality was a significant predictor to distinguish students who earned a certificate of completion from those getting distinction, but not for identifying those who earned no certificate at all. Finally, [16] found that social centrality was not a significant predictor of final grade for any of the three courses they examined, after the number of threads contributed to was taken into account. A possible explanation for the lack of agreement in the existing literature is that prior studies rarely distinguished forum activities that are related to learning of course content from those that are not. MOOC literature shows that whether or not forum discussions are related to the learning of course content can have important implications for learners' participation and interaction patterns [10, 11, 32, 33]. For this reason, distinguishing forum

participation based on the relatedness to course content may help uncover relationships with course performance that would otherwise be concealed.

Understanding the relationship between discussion forum activities and learning outcomes in MOOCs is important both theoretically and practically. Theoretically, it can provide a foundation to better investigate and articulate the mechanism(s) by which participation contributes to and/or displays learning. This is important both for courses which employ discourse as a central element of the pedagogical design and those in which it serves a supplemental function. Practically, such understanding can inform the design and facilitation of forum activities to maximize learning benefits. It might additionally provide grounds for the inclusion of forum activity as an integral (rather than supplemental) element of MOOC design. In addition, it can be used to help learners make efficient use of their time in the discussions by focusing on activities that most benefit their learning. This could impact not just learning, but also course completion, as lack of time is a major factor responsible for attrition among MOOC learners [13].

In this study, we investigate the relationship between forum participation and learning while differentiating discussions related and unrelated to the content of the course. In the following sections, we first review prior research efforts and justify the importance of considering discussion content when exploring this problem space. Then we describe the methods we used to investigate the intended relations. Finally, we report and discuss the findings and propose areas for future research.

2 LITERATURE REVIEW

The relationship between discussion forum participation and learning in MOOCs has been investigated in multiple ways. All work in this area must make two key decisions at the start of a study: (1) how will learning be conceptualized and operationalized; and (2) how will discussion forum participation be conceptualized and operationalized. The relationship between them has generally been posited in a consistent way: more of whatever metrics are chosen to measure valued discussion participation should predict higher values for whatever metrics are used to measure learning.

2.1 Measuring Learning

The conceptualization and operationalization of learning in MOOCs has been relatively similar (and narrow) thus far, focusing on course performance measures such as pass / fail and grades. For example, scores on assignments [12] and overall course grades [17] have been examined for correlation with forum participation variables. Passing / failing and course grades have also been predicted using forum participation variables [e.g., 16, 17, 28]. While on one hand it is logical to draw on course- / instructor-defined criteria for success and the corresponding assessment instruments, this also constructs "learning" as a black-box metric, where the actual contents of the metric may differ dramatically from one course to the next. As previous educational research has shown, which instructional techniques and pedagogical activities are most beneficial for learning depends on the kinds of outcomes intended [34]. This is problematic for previous MOOC studies based on course grade which are forced to claim that discussion forum activity (or any other activity they study) is more or less associated with "learning" writ large without being clear about what that learning is. This may lead to both false positives and false negatives in examining factors that predict learning due to a lack of specificity and differentiation in the outcome(s) of interest [9].

2.2 Measuring Discussion Forum Participation

2.2.1 Amount of (Some Kind of) Contribution. In contrast to learning, there has been much greater variation in how discussion forum participation is conceptualized and measured. The most basic approach is to measure discussion forum participation by the raw quantity of contributions learners make to discussion forums, such as the number of threads contributed to or number of posts made. For instance, [23] examined a digital story-telling MOOC and found that final course grade had a significant but small ($r < .15$) correlation with the extent to which students posted to the discussion forum and responded to others' posts. Notably, in this study discussion participation was collected through self-report in a post-course survey, rather than directly from the system. In another study on a business MOOC, [10] compared the final score of forum contributors with non-contributors (who watched or downloaded at least one lecture) and found that contributors' final score was significantly higher than the non-contributors' with a large effect size ($d = 4.0$).

In addition to making use of contribution quantity variables, some studies have taken into account other characteristics of forum contribution, such as the type of posts (e.g., thread starting post vs reply post, see [1]) and the content of discussions that learners contributed to [e.g., 28]. In this study [28] examined the association between several characteristics of learner's forum contributions (quantity, topic of post, and discourse behavior reflected in post) and final course scores in a psychology MOOC. Using a bag-of-words model developed based on 2,000 hand-coded posts, they first categorized forum posts into on-topic and off-topic discussions; then categorized on-topic posts as active, constructive, or interactive behaviors based on Chi's Interactive, Constructive, Active, and Passive (ICAP) framework [6]. By building several regression models, it was found that final course score could be predicted by whether a learner had contributed to discussion forum (binary yes/no), the number of posts contributed, the percentage of on-topic posts in a learner's total contributions, and the presence of active and constructive behaviors in their contributions. Statistics for evaluating model fit and predictive power were not provided.

2.2.2 Position in the Social Network of Contributions. Going beyond simple counts of activity, some studies have proposed that learner's social network properties can be meaningful as indicators of participation and relationships formed during forum interaction. For example, [17] examined correlations between a learner's social centrality in forums and final course grade in algebra and finance MOOCs. They built undirected social networks for the forums based on copresence of learners in a

common thread and found that degree (the number of direct connections a node has) and betweenness (the number of times a node is part of the shortest path between any two other nodes) had a significant but very small positive correlation ($r < .05$) with final score in the algebra MOOC. However, no significant correlations were found between any centrality measures and final score for the finance MOOC. [19] examined associations between social centrality and the types of certificate learners obtained (no certificate, pass, and distinction). They built directed weighted social networks based on direct replies for two offerings of a programming MOOC taught in different languages and computed weighted degree, betweenness, and closeness (average of the shortest path lengths from a node to all other nodes in the network). For each offering, three multinomial regression models were built to predict the type of certificate obtained, each based on one of the three centrality measures. It was found that for one offering, weighted degree was significantly associated with the likelihood to obtain a certificate and a significant predictor of distinction. For the other offering, all three measures were significantly associated with the likelihood to obtain a certificate and significant predictors of both passing and distinction. No information about model fit or variance explained was provided.

2.2.3 Combining Quantity and Network Measures. Some studies have combined both quantity and network measures of discussion forum participation to predict course performance. For instance, [16] built linear models to predict final course grades in a MOOC on innovation (two offerings) and a MOOC on programming. The prediction was based on the number of threads contributed to, degree, betweenness, closeness, and Bonacich power (a learner's connectedness to influential learners in the network). For each course, three social networks based on copresence of learners in a common thread were built for three weekly sub-forums selected from the beginning, middle, and end of the course. It was found that the predictors only correlated significantly with final course grade in some of the networks, with number of threads being the feature that most frequently correlated with course grade (6 out of 9 networks), followed by degree, betweenness, closeness, and Bonacich Power; the correlations were all relatively weak ($r < .20$). Moreover, adding the social centrality measures to linear regression models based on the number of threads contributed to did not explain significantly more variance in course grade. The amount of variance explained by number of threads was not provided.

In addition to combining quantity and network measures, some studies have also incorporated non-discussion indices of learning activity. For instance, [12] used the quantity of discussion forum contribution together with learner's use of course website and lecture videos to model pass / fail for each unit (60% of homework problems correct) in a MOOC on computer networks. The quantity of forum contribution was measured by the number of forum posts and total number of words posted by a learner; other predictors included number of sessions a student logged into the course platform, page-load requests, total video playing time, and numbers of videos played, rewinds, pauses, fast forwards, and slow plays. They found that homework grade was significantly correlated with both forum variables: number of posts ($r = .20$) and total number of words ($r = .32$). Moreover, total number of words was found useful for predicting unit pass / fail status along with several of the non-forum variables (best model produced AUC = .783, accuracy = 76%). In another example, [18] used forum and non-forum related indicators in week 1 of an introductory biology MOOC to predict the type of certificate learners obtained at the end of the course. The predictors included learner's degree in a social network based on direct-reply relationship, average quiz score, number of peer assessments completed, and whether or not the learner was an incoming student of the university that offered the MOOC (who received incentives to participate). Two logistic regression models were built for predicting distinction vs. normal certificate and normal vs. no certificate. For the distinction vs. normal model (AUC = .947, accuracy = .926), degree was found to be a significant predictor together with number of peer assessments completed and being an incoming student of the university. Degree was not a significant predictor in the normal vs. no certificate model.

2.2.4 Beyond Contribution Based Measures. While posting has commonly been considered as the major form of learning-related participation in discussion forums, forum participants can also learn through reading posts. In fact, reading others' posts represents the reception of ideas (rather than the expression), which is critical to most models of learning through discussion and makes up the majority of time users spend in online forums (e.g., 75% in [35]). Furthermore, when non-posting behaviors are considered, the proportion of students in a MOOC who can be considered to have participated in (and potentially learned from) the forums grow dramatically. [21] found that only 1/3 of visitors to the forum of a MOOC on computer science and programming posted or voted; the other 2/3 of forum visitors engaged in less visible ways, such as reading posts and following threads (to be notified of new contributions).

The potential usefulness of non-posting activities for understanding learning has been explored in a small number of MOOC studies. [7] compared the final score of learners who visited the discussion forum at least once and non-forum-users who viewed at least one lecture, homework, or quiz in a software engineering MOOC and found that forum users' scores were significantly higher that the non-users' (median score of 22% vs. 0%). [1] used number of forum views as one of nine indicators to predict pass/ fail and final course grade in an undergraduate humanities course (for both For-Credit and Non-Credit learners). Forum views was found useful as a predictor in both pass / fail models (For-Credit model AUC = .685, kappa = .470; Non-Credit model AUC = .674, kappa = .288). In addition it was significant in the course grade model (along with four other variables) for Non-Credit students and its inclusion caused two other variables (both measuring the number of replies) to flip to negative signs. This suggests that findings of a positive association between replies and grades in models which do not include a viewing variable may not represent the real dynamics at play.

Despite their conceptual importance in discussion forum learning and potential usefulness in predicting outcomes, non-

posting activities are not always recorded, preventing inclusion in studies. This was the case for the present research. Future MOOC design should make the capture of this data a priority. Additionally, the (re)design of MOOC forum interfaces to allow for read data at the post-level would enable the creation of novel indices for prediction.

3 STUDY FRAMING

The review of the literature reveals three critical issues that merit attention. First, although forum contribution quantity and network variables are often found to be positively associated with course performance, correlations are relatively low and results from prediction studies are mixed as to which of these variables (if any) are consistently useful predictors of course performance. Conceptually, contribution quantity and network centrality variables measure distinct aspects of forum participation, therefore examining and comparing their usefulness for predicting course performance can contribute to our understanding of MOOC forum learning. Specifically, contribution quantity directly tracks the amount of participation while network centrality variables denote social relationships [31]. Moreover, network variables require substantially more effort to compute and are less straightforward to interpret; their use must then be justified by additional predictive value beyond that offered by the simpler quantity measures. To address these issues, the current study examines the relative usefulness of forum contribution quantity and network centrality variables for predicting course performance.

Second, with the exception of [28], little prior work has taken into consideration the content of discussions when examining the association between forum participation and course performance. Discussions in MOOC forums often involve a broad range of topics that may or may not relate to learning of course content [26, 28] and thus would be more or less expected to predict course performance. For instance, [32] found that only 30-55% of discussion threads in five MOOCs on statistics, psychology and physiology were content related. Furthermore, differences have been found in the discourse characteristics and social network properties of content and non-content discussions [11, 25, 31] which are often participated in by largely distinct learners [31]. Thus failing to distinguish between these two kinds of discussions presents a threat to validity and creates a confound that may obscure the relationships that either one has with course performance. Therefore this study specifically investigates quantity and network variables for content, non-content and all discussions to identify if they are differently associated with performance.

Finally, the details needed to fully understand the data analysis conducted in these studies and to interpret the results are not always provided in the literature. For example, forum contribution quantity and network variables often follow a power-law distribution [2, 3] and therefore should undergo transformation before inclusion in models based on assumptions of normality (e.g., linear regression). However such details about data distribution and processing are not commonly reported (and perhaps not even performed). This absence hinders interpretation and comparison of the findings, and creates challenges for replication studies. In addition, not all studies report information about the quality of the models produced, either in terms of predictive power or goodness-of-fit. Regardless of achieving significance, if the relationships between discussion forum participation and course performance are weak, we need to adapt, expand or rethink our approach.

With these three challenges in mind, the current study aims to investigate the associations between learning performance and forum participation by addressing two research questions:

RQ1: Are there differences in MOOC completion and final course grade for learners who did or did not contribute to (content and non-content) discussions?

RQ2: Is forum contribution (measured by quantity and network measures for content and non-content discussions) useful for predicting MOOC course grades?

4 METHODS AND RESULTS

4.1 Data Source

This study used data from StatMed'14, a completed MOOC offered in 2014 on the Stanford open-source platform Lagunita. The course is an introductory course on probability and statistics with a special focus on statistics in medical studies. The course provided a discussion forum as a supplementary venue for optional participation in nine topic areas, including General, Video, Homework, Course Material Feedback, External Resources, Tech Support, Introductions, Study Group, and Platform Feedback. The forum was participated by two members of the instruction team. Learners were invited to post questions and comments about the course in the forums for response by other learners, the TA and the instructor. The forum used a standard linear interface with threads listed on the left and the expanded messages shown on the right. Data available included demographics, final course grade, and discussion forum logs.

4.2 Data

4.2.1 Demographics. There were 21,316 entries in the demographic data table (after removal of the two instructor ids). Each entry included a user id, gender, year of birth, level of education, and country. A total of 15,073 distinct user ids were identified. Of these, 11,029 appeared only once; 4,044 ids appeared 2 to 18 times to produce the remaining 10,287 entries. Multiple entries of the same id contained identical information for gender, year of birth, and level of education, but differed in country. The multiple country values for the same id were produced if a learner accessed the platform from different IP addresses which each yielded its own row in the table. These duplicated rows were removed and yielded 15,073 unique ids.

Of the 15,073 unique learners, 39% reported as female, 54% reported as male and 7% did not provide gender information. Age (calculated from self-reported year of birth) ranged from 0 to 130. It appears some learners have provided false data but the results

cannot be verified. Age distribution is given here to provide a sense of the learner population's age (see Table 1).

Similar to what has been reported in other MOOC studies [4, 5], this learner population is well educated with 83% having a bachelor's or more advanced degree (see Table 2).

These 15,073 learners were from 163 countries. Due to the multiple country values for some user ids, the country data was not broken down to describe the geographical distribution of the learner population.

4.2.2 Grade. The grade book contained 11,665 entries, each for one learner id. Grades of these learners ranged from 0 to 1, M = .116, SD = .265 (see Fig. 1).

4.2.3 Forum. Forum information provided in the dataset includes the following: thread id; post id; user id; post position in thread (starting post, reply post, or reply to reply post); parent post; post text; post creation date and time; and number of votes post received. Thread titles were not included in the data set. There were 566 learner ids in the discussion forum data. Together with the instructors, they made a total of 817 thread starting posts and 2,312 replies in the discussion forum. Five posts that only contained punctuation or non-English language were removed. In this process, one learner who only made one punctuation-only post was removed from the dataset.

Table 1. Age of Learners in Demographic Table (N = 15,073)

Age Group	#	%
0-9	29	0
10-19	367	2
20-29	5,772	38
30-39	4,641	31
40-49	1,728	11
50-59	753	5
60-69	274	2
70-79	49	0
80-89	2	0
>= 90	7	0
/N*	1,451	10
Total	15,073	100

*/N = data not provided.

Table 2. Education Level of Learners in Demographic Table (N = 15,073)

Level of Education	#	%
None	19	0
Elementary	14	0
Junior secondary	96	1
Secondary	1,147	8
Associates	209	1
Bachelors	3,956	26
Masters/PD*	5,813	39
Doctorate	2,673	18
Other	178	1
/N**	968	6
Total	15,073	100

*PD = professional degree

**/N = data not provided.

After data cleaning, a total corpus of 817 threads with 2,307 replies, and 565 learner ids remained in forum data; 11,664 learners remained in course grade data. A total of 555 learners had both forum data and course grade data.

In previous work, the 817 threads were classified as either being content-related or non-content using a unigram and bigram based-model (accuracy for thread starting and reply posts > .81) in conjunction with the DIPTiC method which categorizes threads by comparing the model classification of thread starting post and distribution of replies (estimated categorization accuracy = .88) [8]. The quantity of learner's overall forum contribution was measured by the total number of threads they contributed to and the total number of posts they made. The quantity of learners' contribution to different discussion contexts was measured by number of posts in content-related and non-content threads.

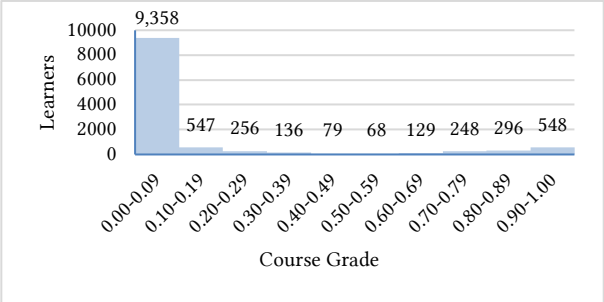

Figure 1: Final course grade of learners in grade book (N = 11,665).

In previous work, three social networks were constructed for the overall discussion forum, the content discussions, and the non-content discussions in this course [31]. The networks were constructed based on limited copresence: all users in small threads (< 5 replies) are connected to each other; in larger threads users are connected to all other users in their sub-thread and the thread starter only. This tie definition produces a comparable network to other conventional choices, such as Direct Reply and Star [31]. In this study, for each forum participant, several social nework centrality measures were computed for the overall, content, and non-content networks, including degree, weighted degree, betweenness, closeness, and eigencentrality (influence of a node based on its connectedness to other influential nodes in the network).

4.3 Forum Contribution and Pass / Fail

4.3.1 Forum contribution and pass / fail. First, the relationship between successfully passing the course (course grade >= 0.6) and forum contribution (contributors / non-contributors) was examined by performing a Chi-Square test of independence on all learners with course grade data available. A significant interaction was found, X^2 (1, N = 11,664) = 1,332.20, p < .001 showing that forum contributors were more likely to pass the course (57%) than non-contributors (8%) (see Table 3).

Table 3. Distribution of Passing X Contribution for All Learners with a Course Grade (N = 11,664)

	Fail (0.00-0.59)	Pass (0.60-1.00)
Forum Contributors (%)	240 (43%)	315 (57%)
Non-Contributors (%)	10,203 (92%)	906 (8%)

A preliminary inspection of learners' course grades revealed that of the 11,664 learners with grades, 8,314 (71%) received a final score of 0.01 or less. It is likely that these learners never engaged seriously in the course and the assignments. To focus more precisely on learners who seriously engaged with the course,

these 8,314 learners (63 forum contributors, 8,251 non-contributors) were excluded from the data for subsequent analyses. Among the 3,350 learners who remained, 492 (15%) were contributors and 2,858 (85%) were non-contributors. The previous analysis was re-run for the set of learners earning a final score greater than 0.01, again producing a statistically significant result X^2 (1, N = 3,350) = 189.33, p < .001. Again, the comparison showed a higher percentage of passing among forum contributors (see Table 4).

Table 4. Distribution of Passing X Contribution for Learners Who Received > 0.01 for Course Grade (N = 3,350)

	Fail (0.02-0.59)	Pass (0.60-1.00)
Forum Contributors (%)	177 (36%)	315 (64%)
Non-Contributors (%)	1,952 (68%)	906 (32%)

4.3.2 Contribution type and pass / fail. Forum contribution is not monolithic. Based on the characterization of discussions according to content-relatedness as described in Section 4.2.3, it was found that of the 492 forum contributors earning grades over 0.01, 165 contributed only to content discussions, 178 contributed only to non-content discussions, and 149 contributed to both kinds.

To take into account differences in the topics that learners engaged with, a Chi-Square test of independence was performed on these 492 learners to examine the differences in passing the course across learners of different forum contribution types: NC (non-content discussion contribution only), C (content-related discussion contribution only) and NC+C (contribution to both kinds of discussions). The comparison indicated a statistically significant interaction between the type of forum contribution and passing the course, X^2 (2, N = 492) = 14.74, p = .001. The largest residuals, indicating deviation from what would be expected due to chance, were found for NC+C (see Table 5) indicating that those who contributed to both kinds of discussions were more likely to pass the course (77% vs 57% / 60%). This could be because contribution diversity is associated with a higher passing rate. It could also be that contribution diversity does not matter and learners who contributed to both kinds of discussions simply had a higher overall quantity of contribution.

Table 5. Distribution of Passing X Contribution Types for Forum Contributors (N = 492)

	Fail (0.02-0.59)	Pass (0.60-1.00)
NC (%), AR*	76 (43%), 2.3	102 (57%), -2.3
C (%), AR*	66 (40%), 1.3	99 (60%), -1.3
NC+C (%), AR*	35 (23%), -3.8	114 (77%), 3.8

*AR = adjusted residue

Follow-up analyses were conducted to disentangle the effects of contribution diversity and quantity. First, the quantity of forum contribution across contribution type was compared as measured by: (a) total number of threads contributed to; (b) total number of posts made; (c) number of posts in content threads; and (d) number of posts in non-content threads. Because the data for these variables violated the assumption of normality, non-parametric tests were used.

For total number of *threads* contributed to, a Kruskal Wallis test showed a statistically significant difference among the NC, C, and NC+C contribution types, X^2 (2) = 297.641, p < .001. Post-hoc Mann-Whitney tests with a Bonferroni correction showed that learners making posts in both NC+C threads contributed to more threads overall than those contributing only to C threads (p < .001, r = .755), and that those contributing only to C threads contributed to more threads overall than those contributing only to NC threads (p = .001, r = .150). For total number of *posts* contributed, a Kruskal Wallis Test showed a statistically significant difference among the NC, C, and NC+C contribution types, X^2 (2) = 244.492, p < .001. Post-hoc Mann-Whitney tests with a Bonferroni correction showed that learners making posts in both NC+C threads contributed more posts overall than those contributing only to C threads (p < .001, r = .682), and that those contributing only to C threads contributed more posts overall than those contributing only to NC threads (p = .001, r = .183). For number of *posts in C threads*, a Mann-Whitney test showed that learners making posts in both NC+C threads made more posts in C threads than those who only contributed to C threads, (p < .001, r = .327). Finally for number of *posts in NC threads*, a Mann-Whitney test showed that learners making posts in both NC+C threads made more posts in NC threads than those who only contributed to NC threads (p < .001, r = .334). These results indicate that learners posting to both C and NC threads participated more than the other two groups when considered both overall and for each kind of post individually (see Table 6).

Table 6. Comparison of Median Contribution Quantity across Different Participation Types (N = 492)

Types of Threads Contributed to	# Threads (Total)	# Posts (Total)	# Posts (C)	# Posts (NC)
NC (N = 178)	1	1	0	1
C (N = 165)	1	1	1	0
NC+C (N = 149)	4	5	3	2

Even with differences in the quantity of contribution, it is possible that diversity of contribution could also explain differences in pass / fail. To separate the effects of participation diversity and quantity on passing rate, logistic regression was used to model passing the course as a function of number of threads, number of posts, and participation diversity (one or both types of discussion). Log transformation was used to make the contribution quantity data suitable for modelling: number of threads (initial skewness / kurtosis = 5.75 / 46.25; final skewness / kurtosis = 1.95 / 3.93); number of posts (initial skewness / kurtosis = 8.58 / 95.68; final skewness / kurtosis = 1.87 / 4.07). The resulting model did not substantially explain variance in passing the course.

4.4 Forum Contribution and Course Grade

The course grade distribution of learners receiving > 0.01 in the course was bimodal with the lowest grade frequencies near the pass/fail threshold (see Fig. 2). This implies that passers and non-passers are best considered as two distinct populations. In this paper, further investigation focused on the relationship between discussion contribution and course grade for those who passed the course (N = 1,221). Non-passers will be investigated in the future.

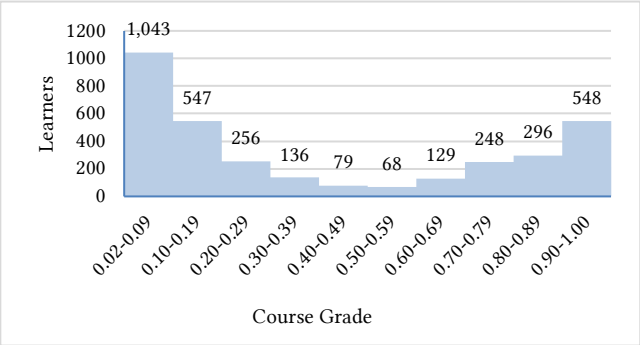

Figure 2: Course grade for learners who received > 0.01 course grade (N = 3,350).

4.4.1 Comparing course grade of passers with different contribution types. There was no overall difference in course grade between those who contributed to the discussion forums (in any way) (N = 315) and those who did not (N = 906), $t_{(1,219)}$ = 1.610, p = .107. However, a comparison of course grades across the three different contribution types (NC, C, NC+C) and no contribution did indicate a statistically significant difference, F (3, 1,217) = 3.295, p = .020. A planned comparison contrasting learners who contributed to content discussions (C and NC+C) with those who did not (NC + Non-contributors) revealed higher course grades for those contributing to content discussions t (1,217) = 3.056, p = .002 (see Table 7).

Table 7. Mean Grade and Standard Deviation for Passers in Different Participation Groups (N = 1,221)

	Non-Contributors (N = 906)	NC (N = 102)	C (N = 99)	NC+C (N = 114)
Mean	0.847	0.835	0.874	0.864
(SD)	(0.108)	(0.109)	(0.099)	(0.105)
Mean	0.845		0.869	
(SD)	(0.108)		(0.102)	

4.4.2 Predicting course grade using forum contribution variables. In order to examine the usefulness of forum contribution variables for predicting learners' course grades, a linear regression model was built. Two types of participation variables (quantity and network) were used as predictors (see Table 8). C closeness, NC closeness, and NC eigencentrality met criteria for normality and were included without transformation. All other predictors had extreme skewness and kurtosis values and appeared to follow a power law distribution; they were therefore log transformed. One variable (overall closeness) remained strongly kurtotic (11.24) even after transformation and therefore was excluded from the model. All variables included in the model were standardized before being input (see Table 8). The distribution for the outcome variable (grades of those who passed the course) did not violate normality.

Table 8. Forum Participation Predictor Variables

	Predictors	Original Skewness / Kurtosis	Final Skewness / Kurtosis
Quantity Variables	Threads	4.95 / 33.87	1.62 / 2.34
	Total posts	7.21 / 65.78	1.65 / 2.94
	C posts	8.01 / 79.71	1.24 / 1.92
	NC posts	3.89 / 18.73	1.01 / 1.06
Network Variables	Degree	3.56 / 17.07	0.67 / 0.26
	Weighted degree	9.44 / 105.15	1.11 / 1.95
	Betweenness	4.13 / 20.63	0.88 / -0.81
	Eigencentrailty	1.67 / 4.73	1.45 / 3.54
	C degree	4.49 / 28.72	0.57 / -0.31
	C weighted degree	9.51 / 107.90	0.93 / 1.02
	C betweenness	6.18 / 46.64	1.45 / 0.60
	C closeness	-0.14 / -0.43	NA
	C eigencentrailty	2.08 / 6.93	1.75 / 4.68
	NC degree	2.59 / 8.66	0.63 / -0.50
	NC weighted degree	4.57 / 26.80	0.84 / 0.10
	NC betweenness	3.77 / 15.26	1.82 / 1.71
	NC closeness	0.10 / 0.15	NA
	NC eigencentrality	0.95 / 0.24	NA

C = content-related, NC = non-content

Variables listed in Table 8 were all candidates for inclusion in the model. Forward selection was used to build the model with a .05 threshold for inclusion. In order to examine the usefulness of network measures for prediction after contribution quantity has been taken into account, variables were added in the order of quantity, then network. As there is no agreement in the prior literature on the usefulness of different quantity and network variables, within each category variables were entered into the regression model based on their correlation with course grade (see Table 9). The quantity variable "posts in content-related threads" had the highest correlation with course grade. After it was included in the model, no subsequent variable significantly increased the variance explained. To test whether other variables offered additional predictive power collectively (even if not individually), a separate model was built in which the variables were added in blocks (all quantity variables together, then all network variables). The network block did not explain significant additional variance (ΔR^2_{ADJ} = .017, F (8, 303) = 1.691, p = .100) on

top of that explained by the quantity block (R^2_{ADJ}= .029, F (3, 311) = 4.136, p = .007) and "posts in content threads" was the only individually significant predictor. Thus the final model for course grade was based on the single predictor "posts in content threads" (*beta* = .188, p = .001) and explained a significant but small proportion of variance in course grade (R^2_{ADJ} = .032, F (1, 313) = 11.427, p = .001). Note that adjusted R^2 in the final single variable model is larger than for the quantity block model because the inclusion of the additional variables did not improve the model more than would be expected by chance.

Table 9. Correlations between Course Grade and Participation Variables

	Predictors	*r*	*p* (2-tailed)
Quantity Variables	Threads†	.136*	.016
	Total posts†	.146**	.009
	C posts†	.188**	.001
	NC posts†	.009	.878
Network Variables	Degree†	.124*	.028
	Weighted degree†	.135*	.017
	Betweenness†	.071	.212
	Eigencentrality†	.150**	.008
	C degree†	.151**	.007
	C weighted degree†	.148**	.009
	C betweenness†	.104	.066
	C closeness	.118*	.037
	C eigencentrality†	.160**	.004
	NC degree†	.042	.459
	NC weighted degree†	.056	.323
	NC betweenness†	.119*	.035
	NC closeness	-.053	.348
	NC eigencentrality	.074	.193

C = content-related, NC = non-content, † indicates log transformed

*p < 0.05, **p < 0.01

5 DISCUSSION AND IMPLICATIONS

5.1 Making (Any Kind of) Forum Contribution is Associated with Passing the Course

Forum contributors had a higher rate of successfully passing the course (57% vs 8% passing) among all 11,664 learners who received a final course grade. This remained true when only the 3,350 learners who received a course grade greater than 0.01 were considered (64% vs 32% passing). Among forum contributors, those who made posts to both content and non-content threads had a high rate of passing than those who only contributed to one type of discussion or the other (77% vs 60% / 57%). These learners also made a greater number of contributions to both kinds of threads, thus it is not possible to disambiguate the effects of contribution quantity and diversity. Other studies have also reported a positive association between forum participation and passing a MOOC [1, 12]. However, for learners who passed the course, there was no difference in course grades between those who did and did not contribute to the discussions. While causality cannot be inferred from any of these studies, these contrasting findings suggest that the role of (undifferentiated) forum contribution may relate more to engagement, effort and persistence than directly to learning of course content. Prior studies have shown that many learners drop out of MOOCs due to a perceived lack of support and social connections [14, 15, 36]. Thus, the positive association between discussion contribution and the likelihood of passing suggests that learners' involvement in the discussion forum may have either fostered, or occurred as a result of, a sense of engagement in the course and commitment to passing it. A similar explanation was given for an earlier finding that experimentally manipulated instructor social presence affected learners' interactions but not course performance in a non-MOOC environment: at a basic level learners want to feel there are other people involved and invested in a course with them, but as long as this requirement is fulfilled, there is no further relationship between the "social" environment and course performance [30].

5.2 Contributing to Content-Related Discussions Predicts Course Grade

Among those who passed the course, there was no overall difference in course grades between those who contributed to the forums and those who did not; however, considering the type of contribution, those who contributed to content discussions performed slightly better than those who did not (final grade of 87% vs 85%). Additionally, five of the six variables measuring content contributions were significantly correlated with course grade, in contrast to only one of the six variables measuring non-content contributions. A significant but small proportion of the variance in course grades (3.2%) was predicted by the final model which included a positive effect of the number of posts in content-related threads. These results provide strong support for the importance of taking the content of forum interactions into account when examining the MOOC discussions with respect to learning of course materials. This aligns with the findings of [28] showing that course performance were predicted by a higher percentage of on-topic posts among learners' total contributions.

These findings have important implications for MOOC research. First, they highlight the importance of differentiating discussion forum data based on content-relatedness, especially when the research purpose is related to understanding learning of course content. Content and non-content discussions have been previously shown to display distinct characteristics in terms of language, interaction patterns, and learners involved [31, 32]. This study extends those findings to show a relationship with course performance as well. Thus prior studies finding a small relationship between discussion contribution quantity and course performance [1, 12], might find a stronger relationship if content-differentiation is incorporated. Second, while a *relationship* has been established, the *causal mechanisms* are not yet clear. Existing work does not speak to whether learners who participated in (content) discussions improved their understanding through their participation and thus got better grades or if learners who got higher grades were more competent and confident in engaging

with content discussions / answering content-related questions and thus participated more (or if some other variable such as engagement or interest affected both grades and discussion contributions). Research can investigate these questions by (a) collecting initial and final measures of interest / engagement; (b) examining the differential effects of different kinds of discussion participation (i.e. not only contributions); (c) controlling for entry levels of ability; (d) modeling changes in course performance and discussion participation over time; and, when possible (e) experimentally manipulating type and level of discussion participation.

Finally, it is important to highlight that the addition of social centrality measures did not significantly improve the variance in course grades that could be explained. [16] found similar results (SNA measures did not provide additional explanatory power over the quantity of discussion contribution) in predicting course grades in two MOOCs on innovation and one on programming. This is a strong sign that while SNA methods may be useful for understanding social interaction and relations in MOOC forums, they may not be the most appropriate measure for predicting learning outcomes measured by final course grades, especially when straightforward contribution quantity variables can be used.

5.3 A Need to Reconceptualize Learning through MOOC Forum Participation

The above sections have explored the implications of specific results that were found in this study. However, taking a broader perspective, it is critical to note that this research does not document a strong relationship between MOOC discussion participation and course performance. Specifically, among passers, no difference was found between the grades of contributors and non-contributors; the difference in grades between learners who did and did not contribute to content discussions was only 2 percentage points out of 100 (SD = .265); and the final model of discussion contribution predicted less than 5% of the variance in course grades. How should this be interpreted? One possibility is that while MOOC discussions provide logistic and social functions, the current actual value for learning of course content is small. This is a reasonable argument given that little attention is given to discussions as an element of MOOC pedagogy and they are generally offered as an optional enhancement to the course, rather than an integrally designed element of it. However, it is also possible that there is valuable learning occurring in the discussions that is simply not captured by course grades. Learners come to MOOCs with a broad range of backgrounds, interests and goals. For example [22] found that some alumni learners participate in the same MOOCs repeatedly; instead of pursuing a high grade, their interest in participating may involve networking and assisting others. Their developing expertise and identity in the domain would be a form of learning [29], but not one captured by course grades. This could be examined by studying changes in learners' roles and social positions in the discussion over time. A different approach would be to look for evidence of learning external to the MOOC. For example [27] proposed investigating learners' participation in a community of practice related to the MOOC they had taken as an alternative way to understand and assess the success of a MOOC.

At the same time that we reconsider what kind(s) of learning we expect to occur through MOOC discussions, we also need to be more explicit in the models of activity thought to lead to these outcomes and align our methods to them. Current approaches have been limited to examining discussion participation as contribution; this both vastly underestimates the number of people involved [21] and confounds those asking questions with those providing answers. If the model for MOOC discussion learning is based on receiving helpful information and explanations about difficult concepts or procedures that will result in higher scores on coursework, then measures should exclude the answer-givers and include everyone who accessed the information. Previous work has shown that non-contribution activity (such as reading posts, following threads, giving "thumbs-ups") represents the vast majority of discussion participation [21, 35]. Some of these measures have also been shown useful in predicting MOOC course grades [1]. In contrast, if the model for MOOC discussion learning relates to growing personal networks and developing expertise, confidence and identity in an area, then a focus on the contributions made is appropriate. It would also, however, call for very different kinds of measures and analyses than simple counts of quantity. In this context SNA centrality might indeed be useful, not as a predictor of learning but as a direct proxy for it (though see caution in how ties are constructed and interpreted, [31]).

Finally, we note that the extent of learners' involvement in MOOC discussions has been consistently underestimated. Common estimates are based on those who contribute as a percentage of those who enroll and range from 3 to 10% [4, 24]. However both the numerator and denominator are inappropriate here. Those involved should be counted as those who derive value from the forums (whether they contribute or not) and they should be counted as a percentage of the total number seriously engaged in the course (excluding those who enroll to just "window shop", see [20]). In this study, after taking out learners with minimal engagement (course grade of 0.01 or less), 15% were found to contribute to the forums. Following [21]'s 2-to-1 guideline of readers-to-writers, this suggests that as many as 45% of learners participated in the forums in one way or another.

5.4 Limitations and Future Research

This study investigated relationship between forum participation and course performance in a MOOC on statistics in medical studies. Future research can examine to what extent these findings generalize to other learning contexts and pedagogical designs. For instance, discussion forums in this course were not designed as an integral part of learning process, but were provided for optional participation. Future efforts can examine the usefulness of forum participation for supporting learning of course content where their use is designed as a pedagogical element of the course. In addition, future work can attempt to disambiguate whether contribution diversity, quantity, or both are related to passing MOOC courses.

6 CONCLUSION

Discussion forums are widely used in MOOCs to offer social interaction and learning support. However, the connections between forum participation and learning have not yet been well established. This study showed an overall relationship between contributing to forum discussions and passing the course, as well as a specific relationship between the quantity of content-related contributions and course grade. These findings indicate the critical importance of considering the content of forum discussions when examining their relationship to course performance. The overall weak relationship between forum contribution and course grade suggests a need to reconceptualize conventional perspectives on what learning occurs in MOOC discussion forums and how it is assessed.

ACKNOWLEDGMENTS

We thank Stanford University and the MOOCPosts Dataset team for their assistance in accessing and working with the data.

REFERENCES

[1] M. Almeda, J. Zuech, R. Baker, C. Utz, G. Higgins, and R. Reynolds. In press. Predicting student achievement from online interaction in a mixed for-credit/open course. To appear in *Online Learning*.

[2] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. 2000. Classes of small-world networks. In *Proceedings of the National Academy of Sciences, 97*(21), 11149-11152. DOI: 10.1073/pnas.200327197

[3] A. L. Barabási and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286, 5439 (Oct. 1999), 509-512.

[4] Y. Belanger and J. Thornton. 2013. Bioelectricity: A quantitative approach to Duke University's first MOOC. Retrieved September 14, 2017 from http://hdl.handle.net/10161/6216.

[5] L. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, and D. T. Seaton. 2013. Studying learning in the worldwide classroom: Research into edX's first MOOC. *Research & Practice in Assessment* 8 (Summer 2013), 13-25.

[6] M. T .H. Chi, 2009. Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science* 1, 1(Jan. 2009), 73-105. DOI: 10.1111/j.1756-8765.2008.01005.x

[7] D. Coetzee, A. Fox, M. A. Hearst, and B. Hartmann. 2014. Should your MOOC forum use a reputation system?. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, New York, NY, USA, 1176-1187. DOI: 10.1145/2531602.2531657

[8] Y. Cui, W. Q. Jin, and A. F. Wise. 2017. Humans and machines together: Improving characterization of large scale online discussions through Dynamic Interrelated Post and Thread Categorization (DIPTiC). In *Proceedings of the 4th ACM Conference on Learning@ Scale*. ACM, New York, NY, USA, 217-219. DOI: 10.1145/3051457.3053989

[9] D. Gašević, S. Dawson, T. Rogers, and D. Gasevic. 2016. Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education* 28 (2016), 68-84.

[10] N. Gillani and R. Eynon. 2014. Communication patterns in massively open online courses. *The Internet and Higher Education* 23 (2014), 18-26.

[11] N. Gillani, T. Yasseri, R. Eynon, and I. Hjorth. 2014. Structural limitations of learning in a crowd: Communication vulnerability and information diffusion in MOOCs. *Nature Scientific Reports* 4 (2014). DOI: 10.1038/srep06447

[12] S. Guo and W. Wu. 2015. Modeling student learning outcomes in MOOCs. In *Proceedings of the 4th International Conference on Teaching, Assessment, and Learning for Engineering*. IEEE, 1305-1313.

[13] C. Gütl, R. H. Rizzardini, V. Chang, and M. Morales. 2014. Attrition in MOOC: Lessons learned from drop-out students. *Learning Technology for Education in Cloud. MOOC and Big Data Communications in Computer and Information Sciences* 446 (2014). Springer, Cham, 37-48. DOI: 10.1007/978-3-319-10671-7_4

[14] K. F. Hew and W. S. Cheung. 2014. Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review* 12 (Jun. 2014), 45-58. DOI: 10.1016/j.edurev.2014.05.001

[15] K. S. Hone and G. R. El Said. 2016. Exploring the factors affecting MOOC retention: A survey study. *Computers & Education* 98 (2017), 157-168. DOI: 10.1016/j.compedu.2016.03.016

[16] S. L. Houston, K. Brady, G. Narasimham, and D. Fisher. 2017. Pass the idea please: The relationship between network position, direct engagement, and course performance in MOOCs. In *Proceedings of the 4th (2017) ACM Conference on Learning@ Scale*. ACM, New York, NY, USA, 295-298. DOI: 10.1145/3051457.3054008

[17] S. Jiang, S. M. Fitzhugh, and M. Warschauer. 2014. Social positioning and performance in MOOCs. In *Proceedings of Graph-Based Educational Data Mining Workshop at the 7th International Conference on Educational Data Mining*. CEUR-WS, 55-58.

[18] S. Jiang, A. Williams, K. Schenke, M. Warschauer, and D. O'Dowd. 2014. Predicting MOOC performance with week 1 behavior. In *Proceedings of the 7th International Conference on Educational Data Mining*. Educational Data Mining Society, USA, 273-275.

[19] S. Joksimović, A. Manataki, D. Gašević, S. Dawson, V. Kovanović, and I. F. De Kereki. 2016. Translating network position into performance: Importance of centrality in different network configurations. In *Proceedings of the 6th International Conference on Learning Analytics & Knowledge*. ACM, New York, NY, USA, 314-323. DOI: 10.1145/2883851.2883928

[20] B. Korsunsky and C. Li. 2017. By hook or by MOOC: Lessons learned and the road ahead. *The Physics Teacher* 55, 3(2017), 146-148.

[21] E. Mustafaraj and J. Bu. 2015. The visible and invisible in a MOOC discussion forum. In *Proceedings of the 2nd ACM Conference on Learning@ Scale*. ACM, New York, NY, USA, 351-354. DOI: 10.1145/2724660.2728691

[22] M. Nelimarkka and A. Vihavainen. 2015. Alumni & tenured participants in MOOCs: Analysis of two years of MOOC discussion channel activity. In *Proceedings of the 2nd ACM Conference on Learning@ Scale*. ACM, New York, NY, USA, 85-93. DOI: 10.1145/2724660.2724671

[23] T. Phan, S. G. McNeil, and B. R. Robin. 2016. Students' patterns of engagement and course performance in a Massive Open Online Course. *Computers & Education* 95 (Apr. 2016), 36-44. DOI: 10.1016/j.compedu.2015.11.015

[24] O. Poquet and S. Dawson. 2015. Analysis of MOOC Forum Participation. In *Proceedings of the 2015 Australian Society for Computers in Learning and Tertiary Education Conference*. ASCILITE, 224-234.

[25] L.A. Rossi and O. Gnawali. 2014. Language independent analysis and classification of discussion threads in Coursera MOOC forums. In *Proceedings of 2014 IEEE 15th International Conference on Information Reuse and Integration*. IEEE, 654-661. DOI: 10.1109/IRI.2014.7051952.

[26] G. S. Stump, J. DeBoer, J. Whittinghill, and L. Breslow. 2013. Development of a framework to classify MOOC discussion forum posts: Methodology and challenges. In *Proceedings of NIPS 2013 Workshop on Data Driven Education*. NIPS Foundation, 1-20.

[27] Y. Wang, L. Paquette, and R. Baker. 2014. A longitudinal study on learner career advancement in MOOCs. *Journal of Learning Analytics* 1, 3(2014), 203-206.

[28] X. Wang, D. Yang, M. Wen, K. Koedinger, and C. P. Rosé. 2015. Investigating how student's cognitive behavior in MOOC discussion forums affect learning gains. In *Proceedings of the 8th International Conference on Educational Data Mining*. Educational Data Mining Society, USA, 226-233.

[29] E. Wenger. 1998. *Communities of practice: Learning, meaning, and identity*. Cambridge University Press, Cambridge, UK.

[30] A. F. Wise, J. Chang, T. Duffy, and R. Del Valle. 2004. The effects of teacher social presence on student satisfaction, engagement, and learning. *Journal of Educational Computing Research* 31, 3(2004), 247-271.

[31] A. F. Wise, Y. Cui, and W. Q. Jin. 2017. Honing in on social learning networks in MOOC forums: Examining critical network definition decisions. In *Proceedings of the 7th International Learning Analytics & Knowledge Conference*. ACM, New York, NY, USA, 383-392. DOI: 10.1145/3027385.3027446

[32] A. F. Wise, Y. Cui, W. Q. Jin, and J. Vytasek. 2017. Mining for gold: Identifying content-related MOOC discussion threads across domains through linguistic modeling. *The Internet and Higher Education, 32*, 11-28.

[33] A. F. Wise, Y. Cui, and J. Vytasek. 2016. Bringing order to chaos in MOOC discussion forums with content-related thread identification. In *Proceedings of the 6th International Conference on Learning Analytics & Knowledge*. ACM, New York, NY, USA, 188-197. DOI= 10.1145/2883851.2883916

[34] A. F. Wise and K. O'Neill. 2009. Beyond more versus less: A reframing of the debate on instructional guidance. In S. Tobias and T. M. Duffy eds. *Constructivist instruction: Success or failure?*. Routledge/Taylor & Francis Group, New York, USA, 82-105.

[35] A. F. Wise, J. Speer, F. Marbouti, and Y. T. Hsiao. 2013. Broadening the notion of participation in online discussions: Examining patterns in learners' online listening behaviors. *Instructional Science* 41, 2(2013), 323-343.

[36] A. M. F. Yousef, M. A. Chatti, U. Schroeder, and M. Wosnitza. 2014. What drives a successful MOOC? An empirical examination of criteria to assure design quality of MOOCs. In *Proceedings of the 14th International Conference on Advanced Learning Technologies*. IEEE, 44-48.