Computers & Education 165 (2021) 104133

Contents lists available at ScienceDirect

Computers & Education

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

MOOC discussion forums: The interplay of the cognitive and the social

Irena Galikyan [a,b,*], Wilfried Admiraal [a], Liesbeth Kester [c]

[a] Leiden University Graduate School of Teaching, Leiden University, the Netherlands
[b] American University of Armenia, Yerevan, Armenia
[c] Department of Education, Faculty of Social Sciences, Utrecht University, the Netherlands

ARTICLE INFO

Keywords:
MOOC discussion forums
Content analysis
Interplay of cognitive and social aspects
Course performance

ABSTRACT

Massive Open Online Course (MOOC) platforms capture the digital traces of millions of learners and generate an avalanche of "numbers" on learner behavior in MOOCs. Yet little is known about the dynamics through which MOOCs can support individual learning as the cognitive and social constituents of this complex process and their interplay within this process do not clearly surface in this large mass of "numbers". This study analyzed the content generated by learners in a MOOC discussion forum with a particular focus on the still under-explored cognitive dimension of learning in MOOCs and demonstrated how certain levels of cognitive engagement relate to learning. It further examined the interplay between the cognitive and social aspects, revealing the moderating role of the social aspect in the association between the lowest level of cognitive engagement and learning in a MOOC environment. The study concludes with discussing the theoretical and practical implications of the findings and with highlighting the need to consider the interdependencies between the cognitive and social variables and learning when designing and evaluating MOOCs.

1. Introduction

Coming to life in 2008, Massive Online Open Courses or MOOCs have since been transforming higher education by spreading it out and making it massively accessible. Today, after a decade, more than 900 universities from around the globe offer over 12 thousand courses reaching out to around 100 million learners through MOOC platforms (Shah, 2019). These virtually unlimited learner numbers come to stand in sharp contrast to the limited instructional guidance available in MOOCs, the massiveness of which makes it practically impossible to monitor and facilitate individual engagement in the learning process. Therefore, to reach their potential and have tangible educational impact, MOOCs might deploy alternative learner support mechanisms, such as establishing self-sustaining collaborative learning communities in which learners interact with each other to construct knowledge (Gillani & Eynon, 2014; Kop, Fournier, & Mak, 2011; Ramesh, Goldwasser, Huang, Daume, & Getoor, 2013). This capacity of MOOCs, in its turn, is conditioned by "the active engagement of several hundred to several thousand 'students' who self-organize their participation according to learning goals, prior knowledge and skills, and common interests" (McAuley, Stewart, Siemens, & Dave Cormier, 2010, p. 4).

In a typical MOOC, the primary space for collaborative knowledge construction is the MOOC discussion forum. As such, it is the only space for learners to engage in textual dialogue and in itself presents unique data on the content generated by learners within a

* Corresponding author. Leiden University Graduate School of Teaching, Leiden University, the Netherlands.
E-mail addresses: i.galikyan@iclon.leidenuniv.nl, iren.galikyan@gmail.com (I. Galikyan).

https://doi.org/10.1016/j.compedu.2021.104133
Received 31 March 2020; Received in revised form 11 January 2021; Accepted 14 January 2021
Available online 19 January 2021
0360-1315/

MOOC (Ezen-Can, Boyer, Kellogg, & Booth, 2015). Consequently, there is a growing body of learning analytics research seeking to decode learner engagement in the vast amount of learner trace data stocked in MOOC discussion forums. Some studies (e.g., Coetzee, Fox, Hearst, & Hartmann, 2014) have interpreted learner engagement as learner participation indices (e.g., the number of posts viewed and posted in discussion forums) and have revealed correlations between forum participation on the one hand and higher grades and retention on the other. Others have looked into the qualitative constituent of learner posts and, examining the content of learner contributions, have found significant correlations between learners' level of forum activity and learning gains (e.g., Wang, Wen, & Rosé, 2016; Wang, Yang, Wen, Koedinger, & Rosé, 2015).

These positive correlations as such do not go beyond establishing a mere connection between individual effort and achievement and do not provide practical implications for course design "beyond exhorting students to be more active" (Reich, 2015, p.34). At the same time, they come to be at odds with the radical decline documented in MOOC discussion participation over time (Brinton et al., 2014). This documented decline, in its turn, is at odds with findings of Seaton, Bergner, Chuang, Mitros, and Pritchard (2013) that identify MOOC discussion forums as the most popular resource referred to by learners during homework completion. The authors suggest that the significant amount of time spent by some students in the discussion forum could be explained by either its pedagogical or social utility or both.

Such disagreement found in the literature in relation to discussion forums' significance to MOOC learning (Wise & Cui, 2018a) suggests that the question should not be about *whether*, but *how* individual cognitive engagement in MOOC discussion forums relates to learning in MOOCs. However, narrowing the analysis down to the level of individual cognition (Jones, 2015) leaving out the social aspect of collaborative knowledge construction is not going to help in finding answers. The individual learner is "an evolving actor, who changes through interaction with others and with new learning experiences" (Stahl, Law, Cress, & Ludvigsen, 2014, p. 366), and understanding the dynamics through which MOOCs make or fail to make learning happen is not possible without insights into the mechanism through which the cognitive and social aspects 'work together'. This study set out to explore the nature of the relationship between these two and to examine how individual learner interactions with other learners might affect the extent to which individual cognitive engagement predicts learners' performance.

1.1. *Cognitive and social aspects of learner interaction in MOOC discussion forums*

Discussion forums can be seen as "the only channel for support and for information exchange between peers" in MOOCs (Boroujeni, Hecking, Hoppe, & Dillenbourg, 2017, p. 128). Despite this, there is still lack of clarity regarding both what discussion forums are achieving as such (Onah, Sinclair, & Boyatt, 2014) and the way and the purpose for which MOOC discussion forums have been studied so far (Almatrafi, 2018). Existing studies on MOOC discussion forums have investigated learner-generated content within the forums from various perspectives, such as analysis of learner sentiments to predict learner dropout (Wen, Yang, & Rose, 2014b), classification of speech acts to predict instructor intervention (Arguello & Shaffer, 2015), identification of linguistic features of content-based and non-content-based starting posts (Wise, Cui, & Vytasek, 2016). Several studies have analyzed learner-generated content with the purpose of detecting and evaluating learner cognitive engagement while others have focused on learner interactions within the forums in order to shed light on the social aspect of learning in MOOCs. As the current study focuses on both the cognitive and social aspects, we will next present a detailed overview of the findings of existing research on these two aspects of learning in a MOOC.

To start with studies that focused on the cognitive aspects of MOOC discussion forums, Wen, Yang, and Rosé (2014a) utilized linguistic markers, i.e., the level of learner language abstraction, to measure cognitive engagement and to predict learner dropout based on their level of cognitive engagement. It was found that the more learners engaged in personal interpretation in their discussion posts, the lower the learner dropout rate from the discussion forums. Wong, Pursel, Divinsky, and Jansen (2015) employed Bloom's revised taxonomy (Anderson & Krathwohl, 2001) to identify key terms for each of the six cognitive domain categories and classified forum messages using an automated algorithm. Although it was revealed that the use of cognitive learning terms in the discussion forum increased over the duration of the course, the researchers had to admit that the use of the taxonomy for classification was "not straightforward" (p. 5). Furthermore, recognizing the limitations of automated measures in decoding cognitive engagement, Wang et al. (2016) employed manual coding of discussion content based on the ICAP (Interactive-Constructive-Active-Passive) framework (Chi & Wylie, 2014) which proposes that learning gains increase as the learner progresses in cognitive engagement level. Higher-order thinking behavior was revealed to result in more learning than paying general or focused attention to course materials. However, the researchers found interactive behaviors to be rare in the analyzed discussion forum and had to group the interactive and constructive levels together into one higher-order thinking level and, therefore, were not able to estimate the distinct effects of the two levels on student learning gains.

In addition, studies examining the social aspect of MOOCs, i.e., learner interactions with other learners, have applied SNA to identify learner interaction patterns. For example, a study by Jiang, Fitzhugh, and Warschauer (2014), examining the relationship between learner's centrality in discussion forums and their performance in two MOOCs, reported mixed results (a significant but small correlation in one MOOC but no relationship in the other). In another study (Houston, Brady, Narasimham, & Fisher, 2017), direct learner interactions—the number of threads a learner contributed to and the number of peers a learner interacted with—were reported to have stronger correlations with MOOC final grade than indirect measures of a learner's social network positioning. Moreover, recognizing the crucial importance of the content of discussions, Wise and Cui (2018b) distinguished between content and non-content discussions in their analysis of the explanatory power of quantity of contributions and measures of social centrality for course performance, thus tapping into the interrelationship between cognitive and social aspects of learning in MOOCs. They showed that the number of contributions to content threads accounted for 3% of variance in course performance, whereas the addition of social centrality metrics and other measures did not significantly improve the explanatory power of the model. It should be noted that overall the study documented a weak relationship between forum contributions and course performance. At the same time, another study examining the interdependencies of temporal patterns, contributed content, and structural roles in MOOC discussion forums (Boroujeni et al., 2017) revealed that even peripheral learners, in this case one time help seekers, play a significant role in triggering discussions on the course content by posting, for example, content-related information requests.

The above-mentioned contradictory findings on the relationship between the cognitive and the social constituents of discussion participation and learning in MOOCs call for a more elaborated understanding. They come to back up the argument that the question of *how* a learner evolves within a MOOC cannot be answered if the analysis is narrowed down either to individual level cognitive processes or to learner interactions with other learners (Eynon, Hjoth, Yasseri, & Gillani, 2016; Jones, 2015). This is further supported by our previous research (Galikyan & Admiraal, 2019) which revealed the moderating role of learner interactions in the association between cognitive presence and academic performance, suggesting that the cognitive and social aspects are contingent on each other in their influence on learning. Therefore, understanding the relationship of individual cognitive engagement in MOOC discussion forums to learning in MOOCs necessitates insights into the interrelation of the cognitive and social aspects in their association to the change process called individual learning (Suthers, 2006). The current study set out to unravel the complex dynamics of this interrelationship. Based on our previous research and literature review, it first examined the levels of individual learner cognitive engagement, through the analysis of MOOC forum content, and then tested how levels of cognitive engagement and individual learner interactions with other learners—the total number of different threads a learner contributes to—affect each other in predicting learner MOOC performance.

Thus, the research questions that the study aimed to address were the following:

- RQ1 What levels of cognitive engagement characterize co-construction of knowledge in MOOC discussion forum?
- RQ2 How do learner interactions and different levels of individual cognitive engagement interact in their influence on learner MOOC performance?

2. Method

2.1. Data

The study was carried out on the data from the *Miracles of Human Language: An Introduction to Linguistics* MOOC offered by Leiden University. This 6-week course gives an introduction into the study of linguistics, and the data came from the three offerings of the course in the beginning of 2018. Course materials include lecture videos and readings. Assessment consists of six weekly quizzes (10% each) and a final exam (40%). The MOOC provides a discussion board with separate forums for each week for learners to interact and seek help. The course provides a discussion board for learners to interact and seek help, however, participation in the discussion forums is optional. There are five forum categories: (a) General discussion, (b) Share resources and join study groups, (c) Questions & answers, including feedback, (d) Meet and greet, and (e) Weekly discussions, each corresponding to a particular module (week) of the course. Each weekly forum in its turn consists of subforums/threads, through which the instructor assigns discussion tasks on weekly topics. A sample task reads as follows: "Given the fact that there are different languages and that these languages may contain untranslatable concepts, would you say that people can fundamentally differ in their ways of thinking? Do you think that people agree on the meaning of concepts if they share a native tongue? Try to think of an untranslatable matter in your language and explain its meaning to others, including those who speak your language." Each task asks learners to write a response to the task and respond to at least to two other students. The forum is supervised by volunteer moderators who receive instruction and coaching by [institution removed for peer review], in addition to support from Coursera. The moderators are instructed to monitor behavior, not the content of the course. Students are given community guidelines beforehand, detailing what (positive) behavior was expected. New threads with the same discussion task on the same topic are started for each course offering by the automated system. Learners in their turn can choose to contribute to an existing thread or initiate a new thread. This results in multiple threads on the same topic.

The MOOC data included demographic information, such as age, gender, and understanding of the course subject matter before taking the course, discussion forum logs, and MOOC final grade. In total, there were 6265 messages posted to 127 discussion threads, from 633 learners, with average age of 34.7 years, 66% self-reported as female, 33% as male and 1% preferred not to provide gender information. Overall, the sample was well educated with 64% holding a bachelor's or higher degrees. MOOC final grades of these learners ranged from 0.83 to 100% ($M = 31.76$, $SD = 31.16$).

2.2. Measures

The content of the MOOC discussion forums was analyzed using the coding instrument developed by Veerman and Veldhuis-Diermanse (2001) and validated and extended by Schellens and Valcke (2005). The instrument makes a distinction between non-task-related and task-related contributions. Non-task-related contributions are categorized as *Planning, Technical, Social,* and *Nonsense*. Task-related contributions, representing the cognitive dimension, are categorized as i) *New Information* (facts, experience/opinions, and new theoretical ideas), ii) *Explicitation*, and iii) *Evaluation*. Thus, *New Information* contributions present relevant content that is new in the context of the discussion and constitute the basic cognitive level; contributions categorized as *Explicitation* refine and elaborate already stated information and represent the intermediate cognitive level; and *Evaluation* contributions that critically discuss earlier contributions "on strength and relevance in the light of the task" (Veerman & Veldhuis-Diermanse, 2001, p. 626) capture the advanced cognitive level. Thus, "the consecutive types of communication represent higher levels of knowledge construction" (Schellens & Valcke, 2005, p. 961). Considering that a single message posted to a discussion forum may contain more than one theme or idea, following Henri (1992), the *unit of meaning* was chosen as the unit of analysis. Two researchers coded 70 messages with an inter-rater agreement of Cohen's kappa of 0.89, 95% CI [0.835, 0.953], which indicated excellent agreement beyond chance (De Wever, Schellens, Valcke, & Van Keer, 2006).

To understand the dynamics of learner cognitive engagement and interaction in MOOC discussion forum, it was necessary to complement content analysis with the analysis of the extent to which a learner interacts with other learners and thus becomes exposed to their ideas. In general, as a MOOC progresses, the discussion forum becomes flooded by multiple threads on the same topics. Following Houston et al. (2017) and Jiang et al., 2014, the number of *different* threads a learner contributed to was taken as a measure of interactions with other learners. The analysis included only threads to which learners contributed task-related contributions.

2.3. Analyses

For the first research question, descriptive statistics were used for the cognitive levels manifested in the MOOC discussion forum. To answer the second research question, a multiple regression analysis was performed with the MOOC final grade as the dependent variable, learner gender, age, and understanding of the course subject matter before taking the course (i.e., prior knowledge) as covariates, the levels of cognitive engagement and the number of different threads contributed to as independent variables. The analysis of the interaction effect of the levels of cognitive engagement and the number of different threads contributed to on the MOOC final grade was conducted using the PROCESS macro (version 3) developed by Hayes (2018).

3. Results

3.1. Levels of cognitive engagement

To address Research Question 1, we conducted content analysis of a total of 6265 messages. The analysis revealed 7295 units of meaning, out of which 1060 (14.53%) were non-task-related and 6235 (85.47%) were task-related contributions suggesting that the communication in the MOOC was predominantly task-related. Sample non-task-related contributions read as follows: "Lovely that we share our ideas;" "Hello, I am not sure about how it works on a MacBook Air." Sample task-related contributions read as: "Whether Esperanto can 'withstand' variation due to cultural influences of its speakers … time will tell and it may well depend whether it becomes more widely spoken as a first language in centralised locations;" "It's difficult to say what is natural or what is unnatural, but a Language lives if there are changes inside it, just like innovations bring forward the discoveries in sciences." The subsequent analyses were based on task-related contributions only. Table 1 presents the distribution of the categories of task-related units—*New Information, Explicitation,* and *Evaluation*—in the MOOC discussion forum. As demonstrated, a significant proportion—72.75% (4536 units)—was related to *New Information* whereas *Explicitation* and *Evaluation* constituted 20.1% (1253 units) and 7.15% (446 units), respectively.

In addition, Fig. 1 depicts the distribution of the categories of task-related units over the six weeks of the MOOC. In general, all three categories showed a decrease in their numbers from Week 1 to Week 6, the steepest being from Week 1 to Week 2. However, starting from Week 2, all three demonstrated diverse patterns. In Week 1, the number of *New Information* units (2857) was almost six times higher than the number of *Explicitation* units (490) and ten times higher than that of *Evaluation* units (288). Already in Week 2, the numbers of *New Information* and *Explicitation* category units decreased five-fold whereas the number of *Evaluation* units decreased more than ten-fold. Despite this, Week 6 shows quite similar numbers in the three categories, specifically 129 *New Information* units, 103 *Explicitation*, and 81 *Evaluation* units.

This suggests that the category that suffered the most drastic decline in the number of units when comparing the first and last weeks of the course is that of *New Information* with a 22-fold decrease in contrast to 4.7-fold and 3.5-fold decrease registered in *Explicitation* and *Evaluation* units. Fig. 2 displays the breakdown of the category units based on the types of instructor-created discussion tasks for each week. As the numbers presented in Fig. 2 suggested a possibility of a relationship between discussion tasks and cognitive levels, additional qualitative analysis of discussion tasks was performed. Although it was revealed that all the tasks asked learners to not only provide their own response to the question posed but also react to the posts of at least two other learners, i.e., build on the contributions of others, some tasks explicitly required refinement and elaboration of one's response while others called for critical analysis of

Table 1
Descriptive statistics.

	$N_{learners} = 608$		
	N	M	SD
1. Age		34.67	14.75
2. New Information	4536	7.46	6.39
3. Explicitation	1253	2.06	3.45
4. Evaluation	446	0.73	1.26
5. Threads Contributed		4.31	3.21
6. Final grade		31.76	31.16

Note. N = Number of units coded at cognitive engagement levels.

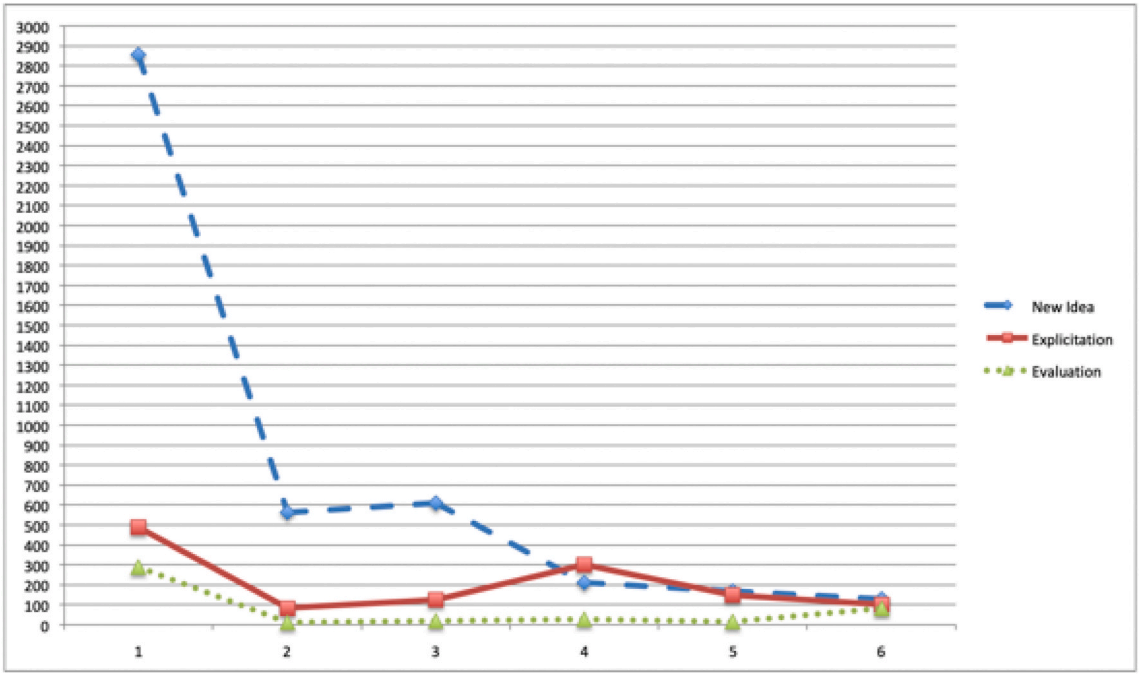

Fig. 1. Distribution of task-related communication by weeks.

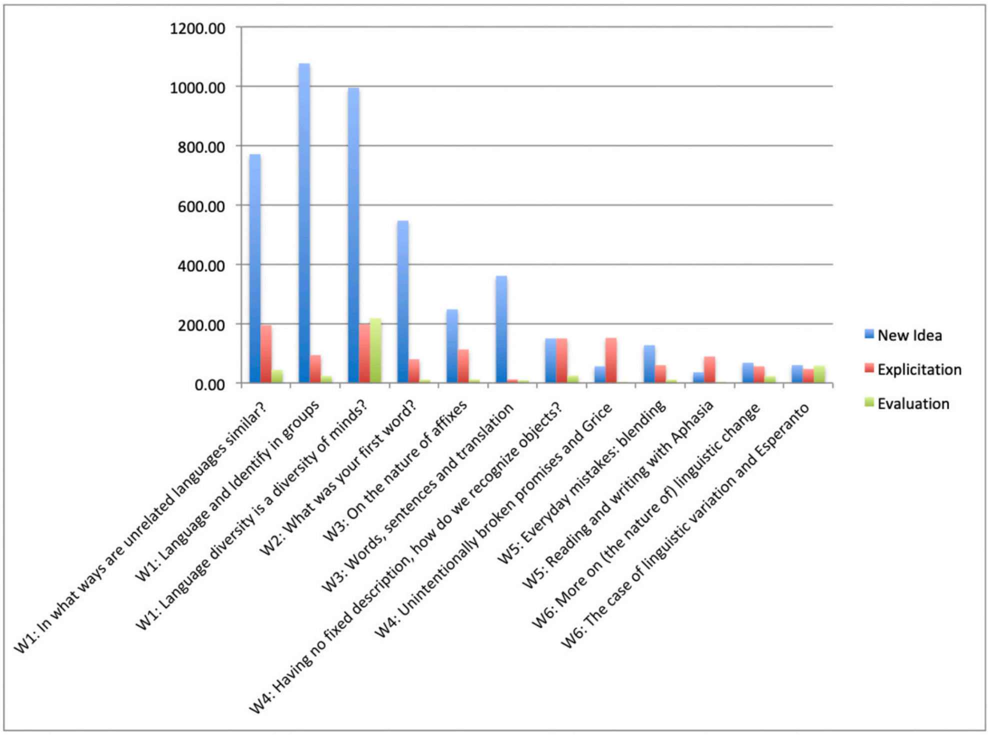

Fig. 2. Distribution of task-related units by discussion tasks.

argumentations. This seems to provide some preliminary insights into how cognitive engagement can vary based on discussion task type with some tasks evoking higher cognitive levels demonstrated by higher numbers in *Explicitation* and *Evaluation* category units.

3.2. Relationship between learner cognitive engagement, number of different threads contributed to, and MOOC performance

In response to the second research question, we first explored the possible correlations between the learners' levels of cognitive engagement and MOOC final grade. For this, the data on the three levels of cognitive engagement were aggregated at the learner level. Table 2 presents the bivariate correlations of each level of cognitive engagement with final grades. As demonstrated in Table 2, statistically significant relationships were found between the three levels of cognitive engagement—*New Information, Explicitation, Evaluation*—and course final grade ($r = 0.48$; $r = 0.59$; $r = 0.41$, with $p < .01$, respectively).

To evaluate the significance of the given variables in predicting learner MOOC grade, we proceeded with multiple linear regression. The normality of residuals was examined through normal Q-Q Plots. The Durbin–Watson value of 1.95 indicated that there were no auto-correlation problems. The multicollinearity of predictors was checked by VIF (Variance Inflation Factor) values below 5. The results of multiple regression analysis (Table 3, Model 1 to 2) indicated that with respect to the cognitive engagement there was a significant negative association between the level of *New Information* and learner MOOC grade (B = - 1.03, $p < .001$), a significant positive relationship between the level of *Explicitation* and learner MOOC grade (B = 1.87, $p < .001$), and a non-significant positive relationship between the *Evaluation level* and learner MOOC grade (B = 1.18, $p = .210$), Thus, the higher the frequency of learners' contributions rated at the level of *New Information*, the lower the learners' MOOC grade, and the higher the frequency of learners' contributions rated at the levels of *Explication* and *Evaluation*, the higher their overall MOOC grade. The number of *Threads Contributed* was positively related to MOOC grade (B = 6.24, $p < .001$). Follow-up analysis was conducted to examine the moderating role of the total number of different *Threads Contributed* in the relationship between the predictor variables and the outcome variable. To probe for potential variable interactions, we used Hayes' (2018) PROCESS v3 Model 1 with 5000 bootstrapping iterations. The results are summarized in Table 3 (Model 2 to 3). The number of different *Threads Contributed* was identified as a significant moderator of the relationship between the frequency of *New Information* contributions and MOOC grades (B = 0.17, $p < .001$, CI [0.083, 0.256], Model 2 to 3). As demonstrated by simple slope analysis (Fig. 3), the negative effect of the frequency of *New Information* contributions on MOOC grades was stronger at lower numbers of *Threads Contributed* (−1.98, [SE = 0.35], $p < .001$, 95% CI [-2.675, −1.288]) and was less strong at higher numbers of *Threads Contributed* (−0.89, [SE = 0.25], $p < .001$, 95% CI [-1.392, −0.399). Thus, for learners with low number of threads, the higher the frequency of contributions rated at the level of *New Information*, the lower the learner grade, and for learners with high number of threads, the fewer the frequency of contributions rated at the level of *New Information*, the higher the learner grade.

Thus, it appears that the total number of different threads contributed moderates the negative relationship between the frequency of engagement in the basic cognitive level and MOOC performance in a way such that the negative effect is stronger for lower numbers of different threads contributed as compared to higher numbers of different threads contributed. The comparison of the correlations between the lowest level of cognitive engagement, i.e., *New Information*, and the two higher levels, i.e., *Explicitation* and *Evaluation,* for learners with low and high number of *Threads Contributed* revealed weaker correlations among learners with low number of threads ($r = .22$; $r = 0.21$, with $p < .01$) than among those with high number of threads ($r = 0.61$; $r = 0.44$, with $p < .01$).

4. Discussion

The current study used the coding instrument of Veerman and Veldhuis-Diermanse (2001) to examine the cognitive levels of learner contributions to MOOC discussion forums. It aimed to reveal the dynamics of the relationship between the cognitive levels and MOOC performance as well as evaluate the potential role of the number of different threads contributed to in the association between learner cognitive engagement and MOOC performance.

Table 2
Summary of intercorrelations.

	1	2	3	4	5	6	7
1. Age	–						
2. Gender[a]	-.05						
3. Prior knowledge	-.08	.04					
4. New Information	.11**	.05	-.02*				
5. Explicitation	.04	-.03	.02	.70**			
6. Evaluation	.11**	-.05	.00	.53**	.59**		
7. Threads contributed	.10*	.05	-.00	.81**	.79**	.57**	
8. Final grade	-.04	.06	.02	.48**	.59**	.41**	.66**

Note. N = 607.
*$p < .05$. **$p < .01$.
[a] Dummy cod[a]d: 0 = male, 1 = female.

Table 3
Multiple linear regression models.

	Model 1		Model 2		Model 3	
Constant	31.25	(3.77)	36.00	(2.80)	33.31	(2.86)
Gender[a]	3.97	(2.6[a])	3.05	(1.99)	2.96	(1.97)
Age	-.06	(.09)	-.18**	(.06)	-.18**	(.06)
Prior knowledge	.32	(1.24)	.07	(.91)	.00	(.89)
New Information			−1.03***	(.25)	−1.44***	(.27)
Explicitation			1.87***	(.47)	1.10*	(.51)
Evaluation			1.18	(.94)	.78	(.94)
Threads contributed			6.24***	(.59)	6.70***	(.59)
Threads contributed × New Idea					.17***	(.04)
R^2	.005		.47***		.48***	
ΔR^2			.10***		.01***	

Note. N = 603. Unstandardized regression coefficients are reported with standard errors in parentheses. All predictors were centered prior to analysis. *$p < .05$; **$p < .01$; ***$p < .001$.
[a] Dummy coded: 0 = mal[a], 1 = female.

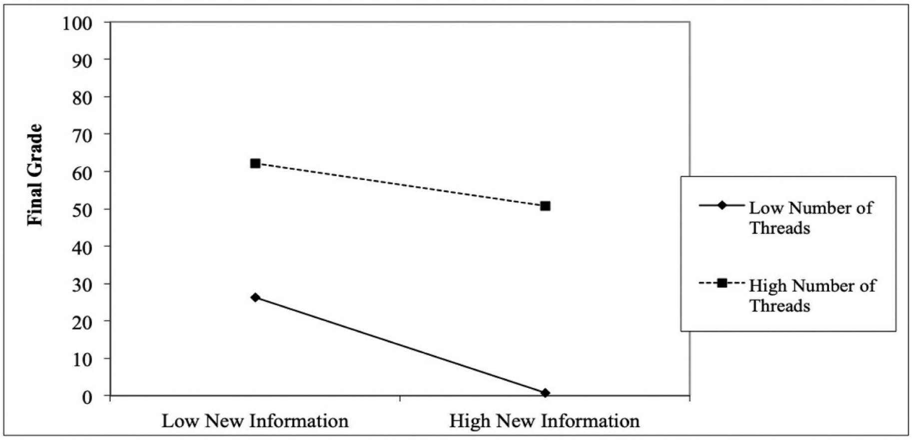

Fig. 3. Moderating Effect of the Number of Threads Contributed on the Relationship between New Idea and MOOC Grades. High and low levels of the Number of Threads Contributed one standard deviation above and below the mean.

4.1. *Cognitive engagement levels*

The results of the current study revealed that learner communication in the MOOC discussion forum was predominantly focused on the content of the course. The most common contributions were those reflecting the basic cognitive level, which were followed by contributions at the intermediate and advanced levels of cognitive engagement. Although this finding is similar to those of other studies on knowledge construction in MOOCs (Goggins, Galyen, Petakovic, & Laffey, 2016; Tawfik et al., 2017) which report mostly lower levels of knowledge construction, the results of the current study point to the possibility of a relationship between MOOC discussion tasks and specific levels of cognitive engagement. Such a relationship between task characteristics and phases of knowledge construction has been suggested to explain the prevalence of low-level knowledge construction found in other computer-supported collaborative learning contexts (e.g., Newman, Webb, & Cochrane, 1997; Schellens & Valcke, 2005; Veerman & Veldhuis-Diermanse, 2001). This finding indicates a need to acknowledge the possibility of such a relationship when designing discussion tasks in order to explicitly elicit contributions at higher cognitive levels, i.e., push learners to go beyond a provision of information or examples.

In addition, it emerges from the results that the number of learner contributions to MOOC discussion forum fall drastically as early as in Week 2. Though such decline in contribution numbers has been documented in other studies (e.g., Brinton et al., 2014; Tawfik, 2017), to our knowledge, this study is the first to demonstrate that it is the lowest cognitive level—the *New Information* category—that displays the most drastic reduction in the number of its units over the course and it is this level that shows a declining trend as the course progresses. At the same time, as indicated by our results, higher cognitive levels—*Explicitation* and *Evaluation*—demonstrate more 'stable behavior', their numbers displaying increase along with decrease throughout the duration of the MOOC. Moreover, towards the end of the MOOC, the numbers of contributions reflecting the three categories are almost equal suggesting a trade-off between quantity and quality of contributions to MOOC discussions as the course progresses. All of these further suggest that MOOC research should not get 'caught up' in contribution numbers. As such, these numbers cannot serve as a meaningful yardstick for understanding whether and how MOOCs support or fail to support learner engagement in knowledge construction. It is the levels of learner cognitive engagement that should be the focus when interpreting contribution numbers to MOOC discussion forums.

4.2. Cognitive engagement and MOOC performance

The above-mentioned propositions on the need to go beyond participation rates are further supported by the results of the analysis of the relationship between learner cognitive engagement and MOOC performance. As suggested by regression modeling, the lowest level of cognitive engagement, which is the one demonstrating the most drastic decrease over the duration of the course, is negatively associated with learner MOOC grades. This negative relationship suggests that learners who tend to reproduce isolated information and fail to comprehend the relations among the information are less successful in understanding the course content. This means that engagement in the lowest cognitive level, that is when a learner shares new facts, personal experiences/opinions, and/or theoretical ideas without any interpretation or without integrating it with the information presented earlier, could, to some extent, be indicative of surface processing of course content, which in its turn is associated with poor academic performance (Biggs, 1991). In contrast, engagement manifested at higher levels, at which learners build on each other's contributions by refining and evaluating their content, is associated with refined understanding of the content of the course. These findings support the hierarchical structure added to the typology by Schellens and Valcke (2005). The findings also support the assertion that discussions provide a rich medium through which insights into learner cognitive processes can be gained (Stump, DeBoer, Whittinghill, & Breslow, 2013), further emphasizing the need for reevaluating the potential of discussion forums for supporting knowledge construction in MOOCs (Shields, 2017).

Our study further examined the moderating role of the social dimension, i.e., the number of different threads a learner contributed to, in the association between cognitive engagement and MOOC performance. As suggested by the results, the total number of different threads a learner contributes moderates the negative effect of the lowest level of cognitive engagement on MOOC performance. The moderation is such that the negative effect is weaker for higher than for lower number of different threads contributed to. The results of the comparison of the correlations revealed weaker correlations between the lowest level of cognitive engagement and the two higher levels among learners with low number of threads than among those with high number of threads. All this might suggest that for learners with low number of threads, engagement in the lowest level of cognitive engagement has a more negative effect as these learners fail to sufficiently engage in the two higher cognitive levels. However, contributing to high number of different threads compensates for this failure to engage in critical discussion allowing learners exposure to a diversity of understandings and perspectives communicated by others and thus reducing the lack of understanding of the course content. The discussion forum provides the space for the emergence of various perspectives and understandings, and it is possible that the greater the number of threads a learner contributes to, the greater the 'volume' of the diversity of perspectives and understandings that the learner is exposed to and can learn from. As shown by research, students are capable of learning by reading the contributions of more expert participants (Cacciamani, Cesareni, Martini, Ferrini, & Fujita, 2012) and even 7-year-old students are able to recognize their own knowledge needs and identify and build upon the ideas of their peers in knowledge-building discussions (Resendes, Scardamalia, Bereiter, Chen, & Halewood, 2015). In fact, studies on learners' positions in CSCL communities and their cognitive outcomes suggest that the greater the number of contacts a learner has with other learners and the shorter the distance to other learners, the better is the learner's performance (e.g., Cadima, Ojeda, & Monguet, 2012; Cho, Gay, Davidson, & Ingraffea, 2007). Moreover, learners have a tendency to form ties with other learners of different performance levels within MOOC discussion networks (Jiang et al., 2014), and low-performing learners often seek and receive help from high-performing learners by initiating new threads in MOOC discussions (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2014). Thus, our finding confirms the crucial role of "dialogues and challenges brought about by differences in persons' perspectives" in knowledge construction (Pea, 1993) and could be indicative of how, due to the scale of MOOCs, peer-to-peer pedagogies are leveraged (Shields, 2017), and "many of the traditional roles and responsibilities of the teaching team are distributed among learners", such as when learners drive each other's understanding through discussion participation (Grover, Franz, Schneider, & Pea, 2013, p.43) or assess their peers' essay assignments using a rubric (Admiraal, Huisman, & Van de Ven, 2014). This proposition is further supported by research that highlights the shifting teacher role in online instructional contexts by showing that online discussions appear to operate as systems in which learners and teachers function as co-equal agents in knowledge construction (Park et al., 2015). Thus, it may be implied that MOOC discussion forum provides affordances and mechanisms for supporting learning by giving opportunity to every learner to benefit from the variety of perspectives presented in the discussion forum threads and thus to benefit from potential knowledge construction by engaging in discussions.

The findings of the current study seem to support the proposition that the avalanche of numbers on learners in MOOC discussion forums does not inherently lead to meaningful answers about learning in MOOCs (Reich, 2015) as MOOC discussion forums are much more intricate than numbers imply (Boroujeni et al., 2017). Finding answers to questions related to the extent to which MOOC discussion forums support learning necessitates going beyond numbers and rates. Our findings corroborate the premise that an individual learner evolves through their interaction with other learners within a MOOC and confirms the importance of both cognitive and social aspects for learning in a MOOC environment. The study emphasizes the need for analyzing the interrelationship between these two for an enhanced understanding of how MOOCs support learning (Boroujeni et al., 2017; Grover et al., 2013) as well as the need to consider the complex nature of these interdependencies in designing MOOC learning environments.

In a practical sense, the understanding of the levels of learner cognitive engagement and how these relate to learning is crucial for course instructors and designers as it allows to reconsider the yardstick for evaluating MOOC learning experiences in general and knowledge construction benefits of discussion forums in particular. In addition, the awareness of how the levels of learner cognitive engagement relate to learning in MOOCs could help MOOC instructors understand how to formulate discussion tasks in such a way as to promote higher levels of cognitive engagement. This would imply formulating discussion tasks that require building on other learners' contributions by refining, elaborating, and evaluating the information posted by others. This would also imply designing learning environments that promote interactions and help learners 'see' the knowledge construction potential of discussion forums by making them aware of the relationship between learning and active participation in MOOC discussion forums.

4.3. Limitations and future research

Considering that the data for the study come from a single course, the generalizability of our findings is limited. Future studies can evaluate the extent to which these findings can be generalized to other MOOCs as well as focus specifically on the relationship between the type of discussion task and level of cognitive engagement. It would also be interesting to examine learners' cognitive engagement levels and their relationship to learning in MOOCs in which participation in the discussion forum contributes to final grade.

5. Conclusion

The current study suggests that learner contributions to MOOC discussions reflect the level of learner cognitive engagement in MOOC content and have a complex relationship with learner performance in a MOOC. The findings on the interrelationship between the cognitive and social variables and learning in a MOOC environment offer a fresh perspective on the interdependencies of cognitive and social dimensions in shaping learning and as a result offer an enhanced understanding of knowledge construction in a MOOC context. The findings highlight the need for considering the interplay of the cognitive and the social when designing and evaluating learning environments as well as when implementing instructional strategies aimed at promoting optimal learning in MOOCs.

Credit author statement

Irena Galikyan: Conceptualization, design, analysis, writing; Wilfried Admiraal: editing/reviewing, supervision. Liesbeth Kester: editing/reviewing, supervision.

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*, 119–128.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: Arevision of Bloom's taxonomy of educational objectives*. New York: Longman.

Almatrafi, O. (2018). *Analyzing MOOC forums: Developing models to support instructors' monitoring of learners' posts* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 13421837).

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with massive online courses. In *Proceedings of the 23rd international conference on world wide web (WWW '14)* (pp. 687–698). New York: ACM Press. https://doi.org/10.1145/2566486.2568042.

Arguello, J., & Shaffer, K. (2015). Predicting speech acts in MOOC forum posts. In *Proceedings of the 9th international AAAI conference on web and social media, ICWSM'15, 2015*.

Biggs, J. B. (1991). Approaches to learning in secondary and tertiary students in Hong Kong: Some comparative studies. *Educational Research Journal, 6*, 27–39.

Boroujeni, M., Hecking, T., Hoppe, H. U., & Dillenbourg, P. (2017). Dynamics of MOOC discussion forums. In *Proceedings of the seventh international learning analytics and knowledge conference (LAK17)* (pp. 128–137). https://doi.org/10.1145/3027385.3027391

Brinton, C. G., Chiang, M., Jain, S., Lam, H. K., Liu, Z., & Wong, F. M. F. (2014). Learning about social learning in MOOCs: From statistical analysis to generative model. *IEEE Transactions on Learning Technologies, 7*(4), 346–359.

Cacciamani, C., Cesareni, D., Martini, F., Ferrini, T., & Fujita, N. (2012). Influence of participation, tutorship styles, and metacognitive reflection and knowledge building in online university courses. *Computers & Education, 58*, 874–884. https://doi.org/10.1016/j.physletb.2003.10.071

Cadima, R., Ojeda, J., & Monguet, J. M. (2012). Social networks and performance in distributed learning communities. *Educational Technology & Society, 15*(4), 296–304. Retrieved from http://www.jstor.org/stable/jeductechsoci.15.4.296.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243.

Cho, H., Gay, G., Davidson, B., & Ingraffea, A. (2007). Social networks, communication styles, and learning performance in a CSCL community. *Computers & Education, 49*(2), 309–329. https://doi.org/10.1016/j.compedu.2005.07.003

Coetzee, D., Fox, A., Hearst, M. A., & Hartmann, B. (2014). Should your MOOC forum use a reputation system?. In *Proceedings of CSCW 2014* (pp. 1176–1187). New York: ACM Press.

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 & Education, 46*(1), 6–28.

Eynon, R., Hjoth, I., Yasseri, T., & Gillani, N. (2016). Understanding communication patterns in MOOCs: Combining data mining and qualitative methods. In S. ElAtia, D. Ipperciel, & O. Zaïane (Eds.), *Data mining and learning analytics: Applications in educational research*. Hoboken, NJ: John Wiley & Sons.

Ezen-Can, A., Boyer, K. E., Kellogg, S., & Booth, S. (2015). Unsupervised modeling for understanding MOOC discussion forums: A learning analytics approach. In *Proceedings of the fifth international conference on learning analytics and knowledge (LAK17)* (pp. 146–150).

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

Galikyan, I., & Admiraal, W. (2019). Students' engagement in asynchronous online discussion: The relationship between cognitive presence, learner prominence, and academic performance. *The Internet and Higher Education, 43*, 100692.

Goggins, S. P., Galyen, K. D., Petakovic, E., & Laffey, J. M. (2016). Connecting performance to social structure and pedagogy as a pathway to scaling learning analytics in MOOCs: An exploratory study. *Journal of Computer Assisted Learning, 32*(3), 244–266.

Grover, S., Franz, P., Schneider, E., & Pea, R. (2013). The MOOC as distributed intelligence: Dimensions of a framework & evaluation of MOOCs. In *Proceedings of the 10th international conference on computer supported collaborative learning* (pp. 42–46). Retrieved from http://www.gerrystahl.net/proceedings/cscl2013/cscl2013proceedings2.pdf.

Hayes, A. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York, NY: The Guilford Press.

Henri, F. (1992). Computer conferencing and content analysis. In A. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden Papers* (pp. 117–136). London: Spinger Verlag.

Houston, S. L., Brady, K., Narasimham, G., & Fisher, D. (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* (pp. 295–298). New York, NY, USA: ACM. https://doi.org/10.1145/3051457.3054008.

Jiang, S., Fitzhugh, S. M., & Warschauer, M. (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* (pp. 55–58). CEUR-WS.

Jones, C. (2015). *Networked learning: An educational paradigm for the age of digital networks*. London: Springer.

Kop, R., Fournier, H., & Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *International Review of Research in Open and Distance Learning, 12*(7), 74–93.

McAuley, A., Stewart, B., Siemens, G., & Dave Cormier, D. (2010). Massive open online courses digital ways of knowing and learning. *The MOOC Model for Digital Practice*. Retrieved February 22, 2019 from http://www.davecormier.com/edblog/wp-content/uploads/MOOC_Final.pdf.

Newman, D. R., Webb, B., & Cochrane, C. (1997). Evaluating the quality of learning in computer supported co-operative learning. *Journal of the American Society for Information Science, 48*(6), 484–495.

Onah, D. F. O., Sinclair, J. E., & Boyatt, R. (2014). Exploring the use of MOOC discussion forums. In *Proceedings of london international conference on education* (pp. 1–4).

Park, J. H., Schallert, D. L., Sanders, A. J., Williams, K. M., Seo, E., Yu, L., et al. (2015). Does it matter if the teacher is there? A teacher's contribution to emerging patterns of interactions in online classroom discus-sions. *Computers & Education, 82*, 315–328.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions* (pp. 47–87). New York: Cambridge University Press.

Ramesh, A., Goldwasser, D., Huang, B., Daume, H., & Getoor, L. (2013). Modeling learner engagement in MOOCs using probabilistic soft logic. In *NIPS workshop on data driven education* (pp. 1–7).

Reich, J. (2015). Rebooting MOOC research. *Science, 347*(6217), 34–35. https://doi.org/10.1126/science.1261627

Resendes, M., Scardamalia, M., Bereiter, C., Chen, B., & Halewood, C. (2015). Group-level formative feedback and metadiscourse. *International Journal of Computer-Supported Collaborative Learning, 10*(3), 309–336.

Schellens, T., & Valcke, M. (2005). Collaborative learning in asynchronous discussion groups: What about the impact on cognitive processing? *Computers in Human Behavior, 21*(6), 957–975.

Seaton, D., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D. (2013). Who does what in a massive open online course? *International Journal of Human-Computer Studies, 68*, 223–241.

Shah, D. (2019). *Year of MOOC-based degrees: A review of MOOC stats and trends in 2018*. Retrieved from https://www.class-central.com/report/moocs-stats-and-trends-2018/.

Shields, K. D. (2017). *Crowdsourcing cognitive presence: A quantitative content Analysis of a K-12 educator MOOC discussion forum*. Doctoral dissertation. Retrieved from https://digitalcommons.kennesaw.edu.

Stahl, G., Law, N., Cress, U., & Ludvigsen, S. (2014). Analyzing roles of individuals in small-group collaboration processes. *International Journal of Computer-Supported Collaborative Learning, 9*, 365–370. https://doi.org/10.1007/s11412-014-9204-9

Stump, G., DeBoer, J., Whittinghill, J., & Breslow, L. (2013). Development of a framework to classify MOOC discussion forum Posts: Methodology and challenges. In *Proceedings of NIPS 2013 workshop on data driven education* (pp. 1–20).

Suthers, D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning, 1*(3), 315–337.

Tawfik, A. A., Reeves, T. D., Stich, A. E., Gill, A., Hong, C., Mcdade, J., et al. (2017). The nature and level of learner – learner interaction in a chemistry massive open online course. *Journal of Computing in Higher Education*, 1–21. https://doi.org/10.1007/s12528-017-9144-2

Veerman, A., & Veldhuis-Diermanse, E. (2001). Collaborative learning through computer-mediated communication in academic education. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *European perspectives on computer supported collaborative learning. Proceedings of the First European Conference on CSCL*. Maastricht: McLuhan Institute, University of Maastricht.

Wang, X., Wen, M., & Rosé, C. P. (2016). Towards triggering higher-order thinking behaviors in MOOCs. In *Proceedings of the sixth international conference on learning analytics & knowledge - LAK '16* (pp. 398–407). https://doi.org/10.1145/2883851.2883964

Wang, X., Yang, D., Wen, M., Koedinger, K., & Rosé, C. P. (2015). Investigating how student's cognitive behaviors in MOOC discussion forums affect learning gains. In *Proceedings of the 8th international conference on educational data mining* (pp. 226–233). Retrieved from http://www.educationaldatamining.org/EDM2015/proceedings/full226-233.pdf.

Wen, M., Yang, D., & Rosé, C. P. (2014a). Linguistic reflections of student engagement in massive open online courses. In *Proceedings of the 8th international conference on weblogs and social media ICWSM*.

Wen, M., Yang, D., & Rose, C. (2014b). Sentiment Analysis in MOOC DiscussionForums: What does it tell us?. In *Proceedings of the seventh international conference on educational data mining (EDM 2014)* (pp. 130–137). Massachusetts: International Educational Data Mining Society.

Wise, A. F., & Cui, Y. (2018a). Envisioning a learning analytics for the learning sciences. In *Paper presented at the 13th international conference of the learning sciences* (London, United Kingdom).

Wise, A., & Cui, Y. (2018b). Unpacking the relationship between discussion forum participation and learning in MOOCs: Content is key. In *Proceedings of LAK'18* (pp. 330–339). ACM.

Wise, A. F., Cui, Y., & Vytasek, J. (2016). Bringing order to chaos in MOOC discussion forums with content-related thread identification. In *Proceedings of the sixth international conference on learning analytics & knowledge - LAK '16 (New York, New York, USA, 2016)* (pp. 188–197).

Wong, J. S., Pursel, B., Divinsky, A., & Jansen, B. J. (2015). Analyzing MOOC discussion forum messages to identify cognitive learning information exchanges. In *Proceedings of the association for information science and technology* (Vol. 52, pp. 1–10). https://doi.org/10.1002/pra2.2015.145052010023