
An Analysis of Cognitive Learning Context in MOOC Forum Messages

Jian-Syuan Wong

College of Information Sciences
and Technology
The Pennsylvania State
University
University Park, PA 16802 USA
jxw477@psu.edu

Anna Divinsky

School of Visual Arts
The Pennsylvania State
University
University Park, PA 16802 USA
axd289@psu.edu

Bart Pursel

College of Information Sciences
and Technology
The Pennsylvania State
University
University Park, PA 16802 USA
bkp10@psu.edu

Bernard J. Jansen

Social Computing Group
Qatar Computing Research
Institute, HBKU
Doha, Qatar
jjansen@acm.org

Abstract

In this research, we analyze a large number of discussions of forum messages from three MOOC courses using a keyword taxonomy approach to identify learning processes occurring among the students. We conduct analysis on more than 100,000 forum messages from 14,647 forum threads from three MOOCs, with a combined 200,000+ enrollment. We map messages to six levels of Bloom's Taxonomy for cognitive learning. The results of this research indicate that learning processes of particular cognitive learning levels could be observed within discussions on MOOC forums. Results imply that different types of forum communications have features associated to particular learning levels, and the volume of higher levels of cognitive learning domains increase as the course progresses.

Author Keywords

MOOC; e-learning; cognitive learning

ACM Classification Keywords

K.3.2 [Computer and Information Science
Education][Miscellaneous]

Introduction

Massive open online courses (MOOCs) for distant learning offer several benefits to online learning, such as receiving quality education with fewer limitations relative to traditional course offerings. Most MOOCs are

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s).
CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
<http://dx.doi.org/10.1145/2851581.2892324>

| | |
|---------------------------------|--|
| Cognitive Learning Level | One of the six classification of cognitive learning in Bloom's Taxonomy |
| Online Discussion Forum | Feature of a site where people can engage in conversations by posting messages |
| Message | A thread, post, or comment posted to an online discussion forum |
| Keyterm | A term used in a message to identify a cognitive learning level |

Table 1: Definition of Key Constructs

free and open to public, providing learners the accessibility to domain experts of various subjects. Due to the large amount of enrolled learners and video-based lectures, communications between student-instructor and student-student within MOOCs become problematic. Therefore, technology-based communications, such as online discussion forums and blogs, have been adopted by MOOC providers to enhance interactions among learners and instructors [6]. The online discussion forum has become indispensable for students and instructors of MOOCs. Peer discussions in courses have been studied in prior work, showing that peer deliberations enhance students learning performance [7]. Moreover, discussion forums might also facilitate learning communities built by and among students that would improve student learning processes [9].

In this research, we study the use of cognitive learning theory for understanding both information seeking and learning processes among the discussions with distinct conversational forum message types, such as posts and comments. In addition, we seek to develop an inferential model for identifying features of the information seeking exchanges based on the contents of the discussions. In our analysis, Bloom's taxonomy [2] is utilized to examine MOOC forum discussions in order to identify the learning processes occurring within these forum conversations.

Bloom's Taxonomy proposes a series of learning domains that classified questions by level of abstraction and has become one of the most widely accepted cognitive learning frameworks. One of the governing principles of the taxonomy is its descriptive outline in

which each type of learning objective can be characterized in a context free manner [1].

Related Work

Because of the emergence of MOOCs, researchers have studied various aspects of asynchronous discussion forums. Dawson studied the ego-network of each student who participated in discussions on a forum using social network analysis [4] finding that students' interactions and participations can be considered a potential factor for predicting the grade of a student. In addition, discussions in online forums are also essential for instructors to monitor and identify problems relating to the assignments, grades, and class surveys, along with monitoring learners progress [8]. Wong et al. [10] studied how active users behaved on a MOOC forum, and the result implied that those enthusiastic users generally make a good contribution to discussions.

However, prior work on learning discussion forums has not proposed a model for characteristics of messages associated to a learning taxonomy. One potential methodology to attain this objective, which we exploit in this research, is the interpretation of forum messages are part of a learning process (e.g., [2]). Therefore, we posit that communications within a discussion forum are information seeking processes [5] that can be analyzed with a cognitive learning approach.

Research Objective

From prior studies, we understand that online discussion forums are viewed as valuable and crucial for distant learning. If the online discussion forums are really a valuable learning resource, there should be indications of such learning in the discussion forum

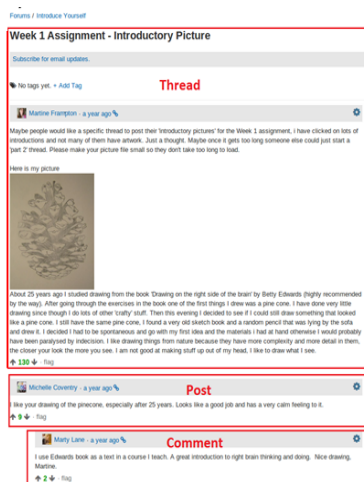


Figure 1. Screenshot of forum discussions

Note:

Messages can be of three types, which are:

- A *thread* contains a title and context of the thread, which allow users to start a new discussion topic on a forum.
- A *post* is the message applied to respond to a thread.
- A *comment* is created to reply to a post.

messages. Therefore, we have the following research objective, which is to **Identify forum messages exchanges occurring within a learning context**. The keywords from Bloom's taxonomy [11] are employed in this research to analyze forum messages to investigate the research objective. In addition, we examine how keywords are used in discussions.

Data

The forum data is collected from three online courses offered on Coursera, which are, (a) Introduction to Art: Concepts & Techniques (ART). (b) Creativity, Innovation, and Change section 1 (CIC01). (c) Creativity, Innovation, and Change section 2 (CIC02). These three MOOCs ran from six to eight weeks. The data consists of three hierarchical parts, including threads, posts, and comments. A *thread* contains a title and context of the thread, which allow users to start a new discussion topic on a forum. A *post* is the message applied to respond to a thread. A *comment* is created to reply to a post. The numbers of forum messages

created in each week of three courses are presented in table 1. In these courses, the ART course (7 weeks) has the most forum messages (more than 50,000) created in the course period. Conversely, approximately 12,500 messages are made in CIC02 course (6 weeks).

As can be seen in table 2, most of forum messages including threads, posts, and comments are created in the first week of three courses comparing to all other weeks. However, a decreasing trend can then be observed from week 2, which could result from the issue of high attrition and dropout rate of MOOCs [3]. Among three types of forum messages, posts and comments were commonly exploited by users to participate in discussions. This observation implies that once a new discussion topic was initiated in a thread, other users can then use posts to reply to the thread. If further communication is necessary, comments also can be utilized to respond to a specific post (see figure 1 as an example).

| | Art | | | CIC01 | | | CIC02 | | |
|-------|--------|--------|---------|--------|--------|---------|--------|-------|---------|
| Week | Thread | Post | Comment | Thread | Post | Comment | Thread | Post | Comment |
| 1 | 5,193 | 10,308 | 7,559 | 2,047 | 9,942 | 7,236 | 721 | 2,590 | 1,897 |
| 2 | 926 | 3,430 | 3,698 | 1,273 | 4,188 | 3,227 | 270 | 1,381 | 962 |
| 3 | 539 | 2,635 | 3,343 | 749 | 2,230 | 1,847 | 163 | 844 | 744 |
| 4 | 358 | 1,669 | 2,502 | 370 | 1,497 | 1,108 | 104 | 604 | 637 |
| 5 | 262 | 1,171 | 2,007 | 221 | 1,061 | 962 | 71 | 339 | 458 |
| 6 | 228 | 976 | 2,026 | 164 | 653 | 627 | 65 | 321 | 420 |
| 7 | 663 | 1,245 | 1,031 | 144 | 593 | 419 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 116 | 483 | 311 | 0 | 0 | 0 |
| Total | 8,169 | 21,434 | 22,166 | 5,084 | 20,647 | 15,737 | 1,394 | 6,079 | 5,118 |

Table 2. Numbers of threads, posts, and comments created in three courses.

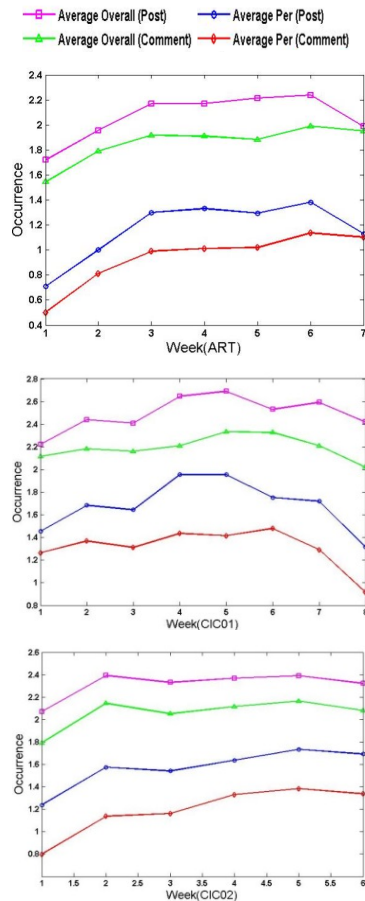


Figure 2. Average and Overall use of cognitive learning terms in posts & comment over three MOOC courses.

| Classification | Definition | Classifying forum message |
|----------------------|---|--|
| Remembering | Retrieving, recognizing, and recalling relevant knowledge | Message must describe, list, or name factual information |
| Understanding | Constructing meaning from oral, written, and graphic messages through interpreting, exemplifying, classifying, summarizing, inferring, comparing, explaining | Message must translate, construe, interpret, or extrapolate information |
| Applying | Carrying out or using a procedure through executing, or implementing | Message must exploit information and put the resulting knowledge into action |
| Analyzing | Breaking material into constituent parts, determining how the parts relate to one another and to an overall structure or purpose through differentiating, organizing, attributing | Message must deduce, scrutinize, or survey information |
| Evaluating | Making judgments based on criteria and standards through critiquing | Message must appraise or relate information to the real world |
| Creating | Putting elements together to form a coherent whole | Message must formulate, generate, restructure, or combine information |

Table 3. Anderson and Krathwohl's taxonomy with definitions (conceptual level) and Classification characteristics for discussion forum messages (operational level).

Method

The analysis and identification of possible learning contexts within forum messages is conducted using keyterms derived from the concept of revised Bloom's Taxonomy, an updated version of Bloom's Taxonomy proposed by Anderson and Krathwohl [1]. The conceptual definition of each cognitive domains of updated Bloom's Taxonomy is shown in table 3. Additionally, how forum messages can be categorized with those domains is also represented in table 3.

As presented in table 3, each classification level denotes a specific domain of a cognitive learning objective, starting from the simplest (*Remembering*) to the most complex (*Creating*). Take *Remembering* as an example; the learning goal of a learner is that s/he is able to recall the methods, basic concepts, or processes of learned materials. In order to identify forum messages that can be categorized into *Remembering* level, we would have to locate terms that describe, list, or name factual information. Therefore, we apply cognitive learning keywords associated to each domain

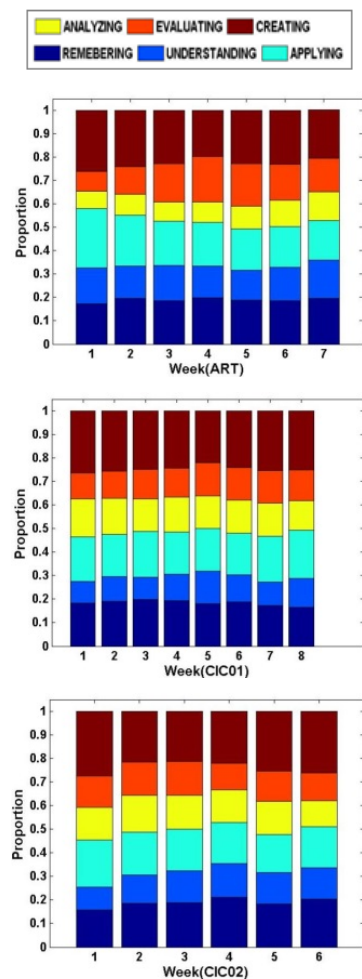


Figure 3. Use of cognitive learning terms in posts in three courses.

[11] to analyze forum messages created on three MOOCs.

In this research, we begin with identifying Bloom's Taxonomy keyterms in each forum message, and then calculate the average occurrences of keyterms in a message. With the analysis, we are able to acquire a better understanding of the use of the Taxonomy's keywords by students on the forum. In addition, we examine the weekly usages in all courses, which provide the insight about the trend of cognitive keywords usage throughout the course as an indication of learning levels by week. We then accumulate these terms to evaluate the distribution of keyterms in each level. Based on this result, the course instructor and designer could have an understanding whether students' learning processes are aligned to the learning objective of each MOOC. For example, *Creating* is especially crucial for ART and CIC courses (e.g., the ultimate goal of ART MOOCs is that learners can use what they have learned from these courses to create their own art pieces.) Therefore, if a higher proportion of keywords of the *Creating* level can be observed, it could imply that students' learning progresses match the course objectives.

Result

In order to identify the learning contexts among discussions on MOOCs forums, we investigate both the occurrences and usage distribution of Bloom's Taxonomy keyterms. Figure 2 presents the average occurrences of cognitive keyterms in forum messages. Average Overall Post/Comment denotes the total occurrences of keyterms divided by the number of all posts/comments. Instead, Average Per Post/Comment indicates the sum of keyterms only divided by the

posts/comments that contained at least one key term. Therefore, we know the values of Average Per Post/Comment should be greater than or equal to the values of Average Overall Post/Comment. The result shows average occurrences of keyterms considering all messages (occurrences/# all messages) and only the messages containing keywords (occurrences / #messages containing keywords).

As can be seen in figure 2, the occurrence of cognitive keyterms is also notably higher in posts relative to comments, indicating that posts may be the most impactful learning events. The pattern of keyterms occurrences are varied in the three courses. The amount increases through the course period in ART course; however, a downward curve is observed in CIC01. Upon investigation, many forum messages of CIC01 in the last weeks are created for having casual conversations rather than discussing about the course content. Those conversations do not include any learning activities, leading to the decreased trend of keyterms usage.

We then analyze the occurrences of keyterms at each level for both post and comments, as shown in figures 3 and 4. In figure 3, we map the weekly occurrences of Bloom's Taxonomy keyterms in forum posts based on the cognitive level. We see some general takeaways. First, there is a great variance in the distribution of cognitive term types at the start of the MOOCs. The reason for the variations could be that different discussion topics may be dominant in certain weeks, which lead to higher proportions of some cognitive levels. However, by the end of the MOOCs, a relatively tightening distribution is observed, resulting in a reduced variance of cognitive term distribution by level

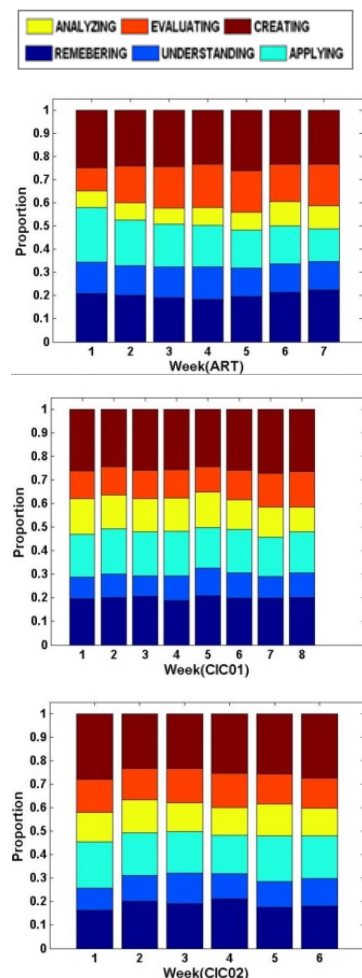


Figure 4. Use of cognitive learning terms in comments in three courses.

in the messages. The possible reason leading to the difference could be that students begin to share their opinions and feedbacks about the course, as well as their appreciations. Thus, discussion topics are very diverse, reducing the variances of different cognitive levels.

In figure 4, we map the occurrences of cognitive keyterms by level per week for comments. Interesting, we do not see the same tightening of cognitive term variances as the MOOCs progress, as we did for posts. Based on figures 3 and 4, we observe that the use of keyterms is varied in different weeks and courses. Nonetheless, keyterms usage for posts and comments within each course show a relative similar pattern.

The proportion of *Creating* is the highest for both posts and comments in all three courses, which range from 19.6% to 28% (posts in week 4 of ART - comments in week1 of CIC02). The use of *Analyzing* is the least common in all three courses, only 7.3% - 15.7% (posts in week 1 of ART - posts in week 2 of CIC02) keyterms belongs to this cognitive level.

MOOCs have become a significant resource for online learning. Due to the geo-location limitations, interactions and communications are problematic among learners. An online discussion forum provides an alternative approach to resolve the issue allowing learners and instructors to communicate with others. In order to evaluate forum users' learning progress through information exchanges during discussions, we apply a revised Bloom's taxonomy for the inquiry. Keyterms corresponding to six cognitive levels are derived to explore the learning progress in each week. We find that the highest proportions of forum messages

are mapped to creating domain in all courses, implying that those messages formulate, generate, restructure, or combine information. By mapping forum messages to each learning domain, course instructors could have a better understanding whether students' learning progresses meet the teaching objectives. In addition, the Bloom's Taxonomy keyterms can be used as an important feature to identify forum messages in different learning levels. A course instructor is then able to locate the discussions with a specific learning level.

Future Work

In this research, keywords corresponding to learning cognitive domains of Bloom's Taxonomy are employed to study the learning context in MOOC forum messages. This research provides preliminary results of identification and analysis about possible educational discussions. In order to have a better insight about how the learning process is performed in a MOOC forum, we will apply techniques of natural language processing, such as entity extraction, to learn what subject is occurring within a learning discussion. Therefore, if a learning key term is located in a forum message, we can then discover the possible discussion topic (e.g., assignment, quiz, etc.). Additionally, text summarization can also be conducted on large amount of learning context to summarize the discussions.

References

- [1] Anderson, L.W. and Krathwohl, D.R., 1999. A revision of Bloom's taxonomy of educational objectives. *Theory Into Practice* (Vol. Complete e, p. xxix, 352 p.)
- [2] Bloom, B.S., COLLEGE, C.O., and Krathwohl, D. R. Taxonomy of educational objectives: The

- classification of educational goals, by a committee of college and university examiners. *Handbook 1: Cognitive domain*. New York: Longmans (1956).
- [3] Clow, D., 2013. MOOCs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge ACM*, 185-189.
- [4] Dawson, S., 2010. 'Seeing'the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology* 41, 5, 736-752.
- [5] Jansen, B.J., Sobel, K., and Cook, G., 2011. Classifying ecommerce information sharing behaviour by youths on social networking sites. *Journal of Information Science* 37, 2, 120-136.
- [6] Mak, S., Williams, R., and Mackness, J., 2010. Blogs and Forums as Communication and Learning Tools in a MOOC.
- [7] Smith, M.K., Wood, , W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., and SU, T.T., 2009. Why peer discussion improves student performance on in-class concept questions. *Science* 323, 5910, 122-124.
- [8] Stephens-Martinez, K., Hearst, M.A., and Fox, A., 2014. Monitoring MOOCs: which information sources do instructors value? In *Proceedings of the first ACM conference on Learning@ scale conference ACM*, 79-88.
- [9] Swan, K., 2002. Building learning communities in online courses: The importance of interaction. *Education, Communication & Information* 2, 1, 23-49.
- [10] Wong, J.-S., Pursel, B., Divinsky, A., and Jansen, B.J., 2015. An Analysis of MOOC Discussion Forum Interactions from the Most Active Users. In *Social Computing, Behavioral-Cultural Modeling, and Prediction Springer*, 452-457.
- [11] Wong, J.-S., Pursel, B., Divinsky, A., and Jansen, B.J., 2015. Analyzing MOOC discussion forum messages to identify cognitive learning information exchanges. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community American Society for Information Science*, 23.