

Research Article

Understanding MOOC Reviews: Text Mining using Structural Topic Model

Xieling Chen¹, Gary Cheng^{1,*}, Haoran Xie², Guanliang Chen³, Di Zou⁴

¹Department of Mathematics and Information Technology, The Education University of Hong Kong, New Territories, Hong Kong SAR, China

²Department of Computing and Decision Sciences, Lingnan University, New Territories, Hong Kong SAR, China

³Faculty of Information Technology, Monash University, Clayton, Australia

⁴Department of English Language Education, The Education University of Hong Kong, New Territories, Hong Kong SAR, China

ARTICLE INFO

Article History

Received 25 June 2021

Accepted 17 November 2021

Keywords

MOOC course reviews
programming courses
learner dissatisfaction
structural topic model
text mining

ABSTRACT

Understanding the reasons for Massive Open Online Course (MOOC) learners' complaints is essential for MOOC providers to facilitate service quality and promote learner satisfaction. The current research uses structural topic modeling to analyze 21,692 programming MOOC course reviews in Class Central, leading to enhanced inference on learner (dis)satisfaction. Four topics appear more commonly in negative reviews as compared to positive ones. Additionally, variations in learner complaints across MOOC course grades are explored, indicating that learners' main complaints about high-graded MOOCs include *problem-solving*, *practices*, and *programming textbooks*, whereas learners of low-graded courses are frequently annoyed by *grading* and *course quality* problems. Our study contributes to the MOOC literature by facilitating a better understanding of MOOC learner (dis)satisfaction using rigorous statistical techniques. Although this study uses programming MOOCs as a case study, the analytical methodologies are independent and adapt to MOOC reviews of varied subjects.

© 2021 The Authors. Publishing services by Atlantis Press International B.V.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

1. INTRODUCTION

Learner satisfaction is increasingly concerned in distance education as a measure of learners' perception of Massive Open Online Courses (MOOC). Online reviews are commonly considered important to influence learners' decisions about whether or not to take an online course [1]. Particularly, negative reviews are more credible, altruistic, and influential as compared to positive ones [2]. Therefore, to improve MOOC learner satisfaction and reduce complaints, executives of MOOCs have become interested in studying topics within course reviews on MOOC platforms to investigate important MOOC features to learners.

Researchers have explored MOOC learner (dis)satisfaction determinants by analyzing online course reviews using manual coding and computer-aided approaches. Conclusions of manual coding approaches depend seriously upon coder expertise [3]. Such subjective biases lead to difficulties in result replication, and the limited sample sizes cause results to be less comprehensive. As such, computer-based text mining approaches like probabilistic topic models for textual course review analysis in MOOC learner (dis)satisfaction research are becoming increasingly important (e.g., [4,5]). However, the traditional way of applying topic models focuses primarily on negative reviews, leading to the potential danger of mixing the positive sides since reviewers usually do not complain from the beginning to the end in writing a negative

review. Therefore, not all topics mentioned in negative reviews through topic models are true negative. To effectively understand causes for MOOC learner (dis)satisfaction, we ought to detect topics significantly appearing only or more in negative reviews as compared to positive ones. As such, the Structural Topic Model (STM) [6] is theoretically suitable since it allows researchers to include covariates that can affect topic distributions [7].

To that end, this study utilizes STM for MOOC review analysis. To demonstrate the methodological effect, we analyze 21,692 programming course reviews to determine the genuine determinants of programming MOOC learners' (dis)satisfaction and reveal the variations in learner complaints across courses with varied grades.

2. LITERATURE REVIEW

A number of studies (Table 1) explore MOOC learners' concerning issues based on manual content analysis or using computer-aided approaches (e.g., frequency analysis and topic modeling) or questionnaire surveys.

Manual content analysis is a common approach for identifying what learners concern about via textual course review analysis (e.g., [1,8]). For example, based on three top-ranked MOOCs, Hew et al. [8] identified factors, including (1) problem-driven learning, (2) lecture accessibility and enthusiasm, (3) active learning, (4) learner-learner interaction, and (5) resources and materials, that accounted for MOOC popularities. Based on manual coding of a sample of 8274 randomly selected MOOC

*Corresponding author. Email: chengks@eduhk.hk

Peer review under responsibility of KEO (Henan) Education Technology Co. Ltd

Table 1 | Research on the exploration of influential factors for MOOC learner satisfaction

Study	Data source	Main findings	Methods
Liu et al. [5]	12,524 reviews from 50 courses in Guokr Web	Topics with the positive (negative) sentiment: course content (course logistics and video production)	Topic modeling
Hew et al. [1]	A random sample of 8274 reviews in 400 MOOCs	Significant factors affecting satisfaction: course instructor, content, assessment, and schedule	Manual coding methods
Peng and Xu [13]	775 posts and 165 responses concerning a course named “Financial Analysis and Decision-making”	Completers show positive sentiment toward course-related content, and non-completers show negative sentiment toward platform construction	Topic modeling
Hew [8]	965 reviews of three highly-rated MOOCs	Five influential factors: (1) problem-driven learning with adequate explanations, (2) lecturer accessibility and enthusiasm, (3) active learning, (4) learner–learner interaction, and (5) resources and materials	Manual coding methods
Ezen-Can et al. [14]	550 posts from 155 learners for an 8-week MOOC	Seven clusters: (1) resources and learner achievement, (2) personal ideas, (3) factual moves, resource access, (4) classrooms, tools, and programs, (5) pedagogical plans, (6) student learning, and (7) support and equity.	Topic modeling
Peng et al. [15]	6163 reviews of 50 highly-rated courses in mooc.guokr.com	Five significant topics: (1) learners’ feelings and experiences about MOOCs, (2) literature course named “A Dream in Red Mansions,” (3) discussion of architecture, (4) basic knowledge of the introduction and planning of marketing, and (5) historical course called “Records of the Grand Historian.”	Topic modeling
Gameel [10]	1786 questionnaires of MOOCs learners	Important factors affecting satisfaction: learner perceived usefulness, instruction and learning aspects of MOOCs, and learner-content interaction	Questionnaire survey and statistical analysis

course reviews, Hew et al. [1] identified course structure, videos, instructors, content and resources, interaction, and assessment as important factors affecting learner satisfaction. Although manual coding approaches have strong interpretation power guaranteed by well-designed frameworks, the reliability of coding results depends seriously upon analyst knowledge, and the coding process is laborious, leading to difficulties in result replication with limited samples [7,9].

There are also scholars exploring factors affecting MOOC learner satisfaction using questionnaire-based surveys alongside statistical methodologies like regression analysis. For example, by analyzing 1786 MOOC learners’ questionnaires, Gameel [10] indicated the significant role of learner perceived usefulness and learner-content interaction in affecting MOOC learner satisfaction. Although findings from such questionnaire-based survey studies help understand MOOC learners’ preferences and facilitate new MOOC design and development, only a limited amount of information is provided without explanation. Also, interviewers commonly have different interpretations of the same questions; therefore, subjectivities cannot be avoided. Additionally, potential researcher imposition may affect result efficacy since researchers usually make assumptions when developing questionnaires [11,12].

In addition to manual coding methodologies and questionnaire-based surveys, a mainstream of automatic text mining approaches for MOOC learner satisfaction research is topic modeling [4,5, 13–15]), which, with the basis of term co-occurrence, have topics (i.e., a collection of terms with high co-occurrence probabilities) as outputs [16,17]. A topic modeling analysis of 50 MOOCs [5] indicates that “learners focus more on course-related content with positive sentiment and course logistics and video production with negative sentiment (p. 670).” Although such analyses enhance our

understanding of important topics that are frequently mentioned by learners, no study focuses on identifying topics appearing significantly more frequently in negative reviews as compared to positive ones.

Results regarding influential factors of learner satisfaction in the existing studies varied. For example, according to Liu et al. [5], learners complained more about course logistics and video production. Hew et al. [1] indicated the significant effects of the course instructor, content, assessment, and schedule on learner satisfaction. Peng and Xu [13] concluded that non-completers were likely to show negative sentiment toward platform construction. Considering the inconsistent findings and the research gap that no study compares positive and negative reviews by adopting topic models, we contribute to MOOC learner satisfaction literature by (1) uncovering topics that frequently appear in negative reviews, and (2) exploring differences in learner complaints among courses with varied grades.

3. DATA AND MODEL SETUP

3.1. Data Collection and Preprocessing

The programming MOOC reviews were obtained from Class Central¹ on 5th June 2020. By further eliminating duplicated MOOCs and those without reviews, 934 MOOCs remained. The course metadata and review data of these MOOCs were automatically crawled. A total of 25,694 review comments were collected. After excluding reviews without providing useful information concerning specific aspects of MOOC courses (e.g., “A completed this

¹<https://www.class-central.com/>.

course” and “Very good”) and non-English ones, 25,135 reviews remained.

Among the 25,135 reviews, 5-, 4-, 3-, 2-, and 1-point ratings were 88.42%, 12.61%, 1.66%, 0.70%, and 0.76%, respectively (Table 2). To deal with the unbalanced rating distribution, we followed the strategy in Hu et al. [7] to define negative reviews as those with 1- or 2-point ratings, and 5-point rating reviews were treated as positive ones. Lastly, 21,692 reviews for 304 courses were used for topic modeling and analysis.

For the 21,692 reviews, terms were extracted using Natural Language Tool Kit (NLTK),² with which word normalization and exclusion of default stop-words (e.g., “the” and “is”), punctuations, and numbers were conducted. Second, terms that were written in British and American English (e.g., “analyze” and “analyse,” “beginner” and “beginner”) were manually unified. Third, user-defined stop-words such as “everything,” “everyone,” and “everybody” were excluded. Fourth, manual stemming, that is, converting terms to root forms (e.g., “automate,” “automatic,” and “automation” were converted to “automat”), was conducted. Also, misspellings and non-words were excluded (e.g., “x2,” “x27,” “Xavier,” “xc,” “xd,” “xi-xii,” “x-mas,” “yh,” “yo,” and “y-o”). Furthermore, terms that represent the names of a person or course were excluded to reduce noise. In addition, we filtered unimportant terms with a term frequency-inverse document frequencies value of not more than 0.03.

3.2. STM Model Setup

In STM, each topic and document is assumed to be a distribution of terms and a mixture of topics, respectively. With a significant improvement over Latent Dirichlet Allocation, STM allows examining how document-level structure information affects topic prevalence and content. Figure 1 presents the plate diagram of STM, in which nodes are variables with labels indicating their roles in the modeling process. The unshaded (shaded) nodes represent hidden (observed) variables. Rectangles indicate replication: $n \in \{1, 2, \dots, N\}$ indexes terms in each document; $k \in \{1, 2, \dots, K\}$ indexes each topic supposing K topics; and $d \in \{1, 2, \dots, D\}$ indexes documents. STM aims to understand the unseen

topics based on the observable terms, W , and generates two matrices, that is, document-topic distributions θ and topic-term distributions β .

Taking advantage of STM’s ability to add topic prevalence parameters to influence a document’s distribution over topics, we modeled the impact of review extremity (positivity and negativity), course grade, and their interactions on topic prevalence μ , which further determines document-topic distribution θ . To explore how document-level covariates influenced topic prevalence, *Negative* and *CourseGrade* were defined as covariates to indicate review orientation and course ratings separately. *Negative* equals 1 for negative reviews with 1- and 2-point ratings, and 0 for positive reviews with 5-point ratings. *CourseGrade* ranges from 1 to 5. The relationship between topic prevalence and covariates is indicated by $Prevalence = g(Negative, CourseGrade, Negative \times CourseGrade)$, where $g(\bullet)$ represents a generalized linear function.

A key step in STM is to choose the number of topics K . Such a task was performed by using both quantitative and qualitative techniques. Specifically, from the quantitative perspective, we proceeded by estimating various models with varied numbers of topics (i.e., from five to 30). We compared the estimated models using their scores on average semantic coherence and exclusivity [7,18] (Figure 2), two popular indexes in topic modeling for an initial selection of the input number of topics. Three models ($K = 14, 15$, and 16) outperformed others and were selected for

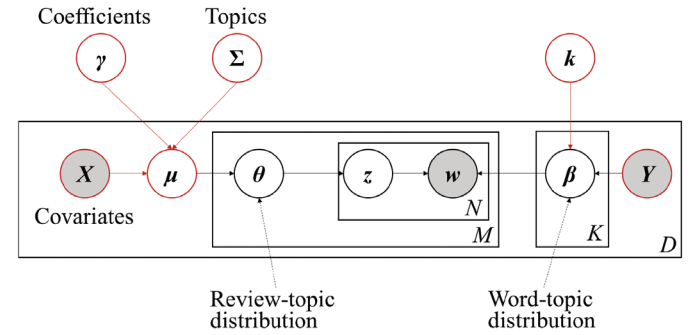


Figure 1 | Plate diagram of STM.

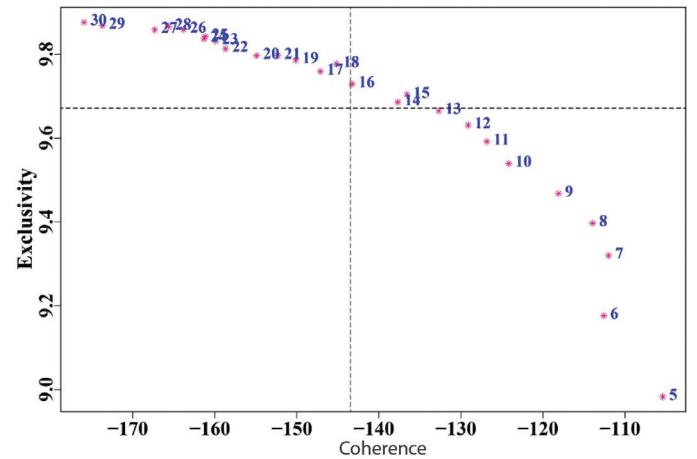


Figure 2 | Model diagnosis using exclusivity and coherence.

Table 2 | Statistics of the included reviews

Items	Frequency	%
Distribution of course grade		
Grade 1	8	2.63
Grade 2	14	4.61
Grade 3	35	11.51
Grade 4	89	29.28
Grade 5	158	51.97
Distribution of rating		
Grade 1	183	0.76
Grade 2	168	0.70
Grade 3	400	1.66
Grade 4	3,043	12.61
Grade 5	21,341	88.42

²<https://www.nltk.org/>.

Figure 3 Frequently used terms in (a) positive reviews (term frequency ≥ 200) (download the dynamic map via https://drive.google.com/file/d/1CdD7otd4zT5UKUTzmCKR6ay_sjtqwvYI/view?usp=sharing) and (b) negative reviews (term frequency ≥ 10) (download the dynamic map via https://drive.google.com/file/d/1E3A7my0QLcBZ54av92RABvavrnlBV8_v/view?usp=sharing).

Table 3 | Frequently used terms in positive reviews

Word	F	%	Word	F	%	Word	F	%
python	10,120	47.42	follow	1131	5.30	database	547	2.56
learn	8206	38.45	structure	1122	5.26	design	523	2.45
programming	7680	35.99	people	1072	5.02	present	515	2.41
easy	4007	18.78	provide	1040	4.87	organize	511	2.39
recommend	3688	17.28	content	1016	4.76	hope	508	2.38
beginner	3675	17.22	week	990	4.64	university	500	2.34
understand	3641	17.06	need	982	4.60	book	491	2.30
teach	3516	16.48	like	969	4.54	previous	488	2.29
class	3498	16.39	quiz	969	4.54	overall	473	2.22
start	3370	15.79	online	965	4.52	thought	470	2.20
basic	3239	15.18	challenge	916	4.29	lesson	467	2.19
interest	2728	12.78	see	911	4.27	understandable	451	2.11
instructor	2541	11.91	compute	903	4.23	day	445	2.09
explain	2537	11.89	explanation	893	4.18	fundament	444	2.08
experience	2472	11.58	web	858	4.02	tool	442	2.07
help	2411	11.30	skill	848	3.97	manner	436	2.04
code	2383	11.17	able	838	3.93	hard	435	2.04
assignment	2375	11.13	specialization	830	3.89	put	432	2.02
professor	2354	11.03	finish	829	3.88	appreciate	431	2.02
language	2324	10.89	build	828	3.88	different	430	2.01
knowledge	2080	9.75	keep	827	3.88	idea	430	2.01
teacher	2046	9.59	programmer	799	3.74	read	429	2.01
fun	2030	9.51	topic	771	3.61	informative	428	2.01
concept	1977	9.26	job	766	3.59	science	416	1.95
material	1954	9.16	write	759	3.56	forum	411	1.93
time	1925	9.02	background	734	3.44	introduce	407	1.91
complete	1850	8.67	cover	731	3.43	right	398	1.86
video	1801	8.44	pace	701	3.28	real	396	1.86
enjoy	1795	8.41	continue	687	3.22	funny	395	1.85
work	1739	8.15	project	683	3.20	methodology	393	1.84
introduction	1695	7.94	step	674	3.16	create	391	1.83
give	1687	7.90	game	673	3.15	wish	389	1.82
know	1679	7.87	level	670	3.14	subject	387	1.81
lecture	1598	7.49	say	670	3.14	interactive	386	1.81
clear	1531	7.17	difficult	651	3.05	come	385	1.80
look	1469	6.88	practice	650	3.05	prior	382	1.79
helpful	1438	6.74	begin	640	3.00	review	381	1.79
love	1397	6.55	problem	616	2.89	confidence	378	1.77
coursera	1368	6.41	perfect	607	2.84	bonus	376	1.76
useful	1270	5.95	world	605	2.83	style	369	1.73
student	1256	5.89	engage	602	2.82	advance	364	1.71
feel	1203	5.64	detail	601	2.82	series	364	1.71
simple	1203	5.64	fascinate	589	2.76	future	362	1.70
use	1181	5.53	try	588	2.76	effort	361	1.69
exercise	1159	5.43	point	577	2.70	general	360	1.69
data	1154	5.41	information	563	2.64	chapter	359	1.68
example	1136	5.32	study	556	2.61	practical	357	1.67

F, frequency.

“python,” “material,” “quality,” “explain,” “exercise,” and “time,” are mentioned more frequently.

4.2. Topic Summary and Validation

Table 5 displays the topic modeling results. The first column shows the labels that briefly summarize the meaning of each topic. The second column presents the top terms having the highest frequencies and exclusivities of each topic. The third column indicates topic prevalence. The most concerning topic is *learner friendliness*, followed by *solution explanation*, *learning perception*, *learning attitude*, and *course interactivity and difficulty*.

4.3. Identification of Negative Topics

Figure 4 visualizes the relative topic prevalence in positive and negative reviews, where dots are “the mean values of the estimated differences and the bars are 95% confidence intervals for the difference (p. 422) [7].” The results evidently show four negative topics, including *grading*, *problem solving and practices*, *course quality*, and *programming textbooks*. We also provide examples of representative reviews for each negative topic (Table 6).

The top negative topic is *grading* with representative words such as “grade,” “question,” “wrong,” “distract,” “autograde,” “mark,” “answer,” and “check.” The examination of representative reviews indicates that the topic is commonly associated with technical

Table 4 | Frequently used terms in negative reviews

Word	F	%	Word	F	%	Word	F	%
learn	89	25.36	provide	31	8.83	second	19	5.41
code	83	23.65	waste	31	8.83	watch	19	5.41
time	78	22.22	introduction	29	8.26	add	18	5.13
teach	70	19.94	people	29	8.26	discussion	18	5.13
give	69	19.66	difficult	28	7.98	grade	18	5.13
video	69	19.66	lesson	28	7.98	last	18	5.13
assignment	64	18.23	forum	27	7.69	least	18	5.13
beginner	64	18.23	point	27	7.69	online	18	5.13
understand	64	18.23	say	27	7.69	require	18	5.13
python	61	17.38	follow	26	7.41	able	17	4.84
programming	60	17.09	poor	26	7.41	android	17	4.84
student	58	16.52	quality	26	7.41	data	17	4.84
complete	57	16.24	read	26	7.41	expect	17	4.84
instructor	57	16.24	spend	26	7.41	google	17	4.84
use	56	15.95	write	26	7.41	hard	17	4.84
class	54	15.38	easy	25	7.12	helpful	17	4.84
need	53	15.10	experience	25	7.12	lack	17	4.84
see	53	15.10	topic	25	7.12	simple	17	4.84
work	53	15.10	web	25	7.12	specialization	17	4.84
feel	47	13.39	bad	24	6.84	change	16	4.56
lecture	47	13.39	different	24	6.84	clear	16	4.56
material	46	13.11	disappoint	24	6.84	detail	16	4.56
example	44	12.54	frustrate	24	6.84	error	16	4.56
help	44	12.54	project	24	6.84	fail	16	4.56
know	44	12.54	ask	23	6.55	language	16	4.56
recommend	43	12.25	level	23	6.55	lecturer	16	4.56
content	42	11.97	professor	23	6.55	money	16	4.56
week	41	11.68	cover	22	6.27	outdated	16	4.56
explain	40	11.40	free	22	6.27	pay	16	4.56
look	39	11.11	confuse	21	5.98	present	16	4.56
quiz	38	10.83	explanation	21	5.98	talk	16	4.56
concept	37	10.54	answer	20	5.70	wrong	16	4.56
knowledge	35	9.97	book	20	5.70	c	15	4.27
information	34	9.69	c++	20	5.70	create	15	4.27
problem	34	9.69	hour	20	5.70	design	15	4.27
start	34	9.69	instruct	20	5.70	finish	15	4.27
basic	33	9.40	interest	20	5.70	idea	15	4.27
question	33	9.40	reason	20	5.70	structure	15	4.27
try	33	9.40	show	20	5.70	update	15	4.27
exercise	32	9.12	star	20	5.70	useful	15	4.27
coursera	31	8.83	udacity	20	5.70	advance	14	3.99
issue	19	5.41	practice	19	5.41	begin	14	3.99

F, frequency.

Table 5 | Topic summary

Labels	Representative terms	%
<i>learner friendliness</i>	understandable, basic, prior, friendly, beginner, suitable, logic	17.23
<i>solution explanation</i>	easy, concept, explain, clear, exercise, digest, precise	12.49
<i>learning perception</i>	interest, cool, python, master, useful, well-structured, pleasure	10.44
<i>learning attitude</i>	enjoy, feel, complete, pace, comfortable, passionate, fun	8.17
<i>course interactivity and difficulty</i>	simple, confidence, interactive, help, lock-down, scratch, join	7.73
<i>learning motivation and engagement</i>	game, humor, willing, engage, bore, motivate, informative	6.51
<i>learning experience</i>	continue, excite, slow, worry, experience, fascinate, refresh	5.96
<i>programming style and skill</i>	session, style, brilliance, idle, software, breakpoints, cpu	5.20
<i>IT infrastructure for MOOCs</i>	database, sql, db, module, flow, control, dictionary	5.09
<i>grading</i>	grade, question, wrong, autograde, mark, answer, check	5.07
<i>programming project</i>	xml, json, api, html, parse, project, network, mini	4.87
<i>problem solving and practices</i>	solve, balance, homework, encounter, practical, real, mistake	4.05
<i>programming textbooks</i>	compute, describe, section, science, textbook, home	3.67
<i>course quality</i>	outstanding, key, limit, perspective, quality, link, rich	3.52

issues concerning submission, file uploading, website execution, and grader settings, as well as issues concerning unfair or error in grading and not available for trying again.

The second most negative topic is *problem solving and practices*. The associated complaints mainly center around classroom time arrangement, ambiguous wording of problems, and inappropriate difficulty level of exercises (either too difficult or pointless). Also, there are many learners complaining that they have to upgrade (to paid accounts) to complete quizzes. Furthermore, some learners

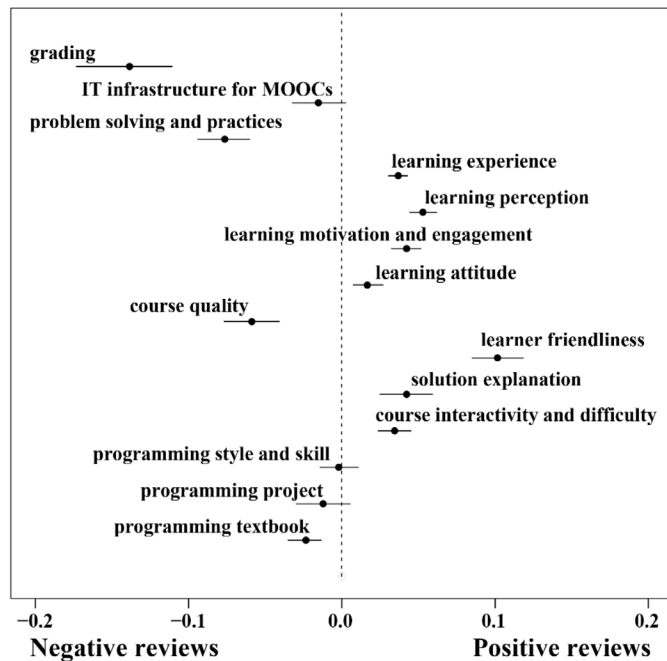


Figure 4 | Differences in the topic prevalence (negative reviews versus positive reviews).

point out that instructors' tips about practices are often too general and not helpful. Additionally, many learners complain that there are ambiguous quiz questions without explicit explanation.

The third topic is *course quality* with representative reviews commonly linked with an amateurish presentation, sloppy codes, poor and limited code examples, misleading and wrong information, unnecessary repetition, as well as incorrect or insufficient explanations of important concepts and functions, which fails to convey them in a concise and easy-to-understand format for beginners. In addition, there are learners complaining about broken links, especially "more info" links that just go to visual studio home pages where there are more advertisements than information.

The fourth topic is *programming textbooks*, where learners mainly complain that (1) chapters in textbooks are too brief, (2) textbooks are less useful while doing homework, (3) textbooks are extremely theoretical with no real-world information, and (4) much of the content is over-simplified to the point of ridiculous.

Results of the negative topic identification offer insights into factors accounting for MOOC learner satisfaction toward programming courses. Learners can be easily impressed by courses that are learner-friendly, especially for beginners. Furthermore, most learners are satisfied with the content and materials with moderate difficulty. In addition, instructors' clear explanation of solutions and their interaction with learners can motive learner engagement.

4.4. Topic Correlation Analysis

The topic correlation [19] in Figure 5 intuitively indicates the essence of MOOC learner (dis)satisfaction. In the figure, topics are represented as nodes, and two nodes with high co-occurrence probabilities are linked. A shorter link between two nodes indicates a closer correlation. The sizes of nodes indicate their prevalences in the corpus. The colors of nodes indicate topic extremities, where

Table 6 | Examples of representative reviews for negative topics

Negative topics	Examples of reviews
<i>grading</i>	<p>"You can't submit quizzes to see if you got them right, and all your coding problems are marked wrong. You can still do the coding, and the website will error check and all, but you won't get the pass/fail grade to be sure you got it right."</p> <p>"If using any new version of ruby or rails, you will have to modify multiple items in your files/system to get the grader to work."</p>
<i>problem solving and practices</i>	<p>"I would have given 5 stars but my assignment has not got a fair marks. And I am not passed due to that grading."</p> <p>"The time stated to complete the course was 4–6 h per week. With the lectures at 1–2 h per week, that left 3–4 h to complete the homeworks.....not a chance."</p> <p>"The wording of the problems too ambiguous; maybe that is the CS style? Some of the exercise seemed quite pointless. Inconsistent level of difficulty. The last few problems required one to look for general patterns to code."</p> <p>"You have to upgrade (to paid account) in order to complete quizzes. It's very unfortunate."</p> <p>"Course is practically useless without exercises. Verified certificates cost money and it does mean students who not pay can not practice."</p>
<i>course quality</i>	<p>"This course covers many important aspects of the JavaScript syntax, but ultimately fails to convey them in a concise and easy-to-understand format that should be the basis of a beginner's class."</p> <p>"I am an experienced programmer and teacher of programming. I started this course because I thought - reasonably, given the title- that the course would get right to C++. Instead, there's a great deal of C and basic algorithms. This is a waste of my time!"</p>
<i>programming textbook</i>	<p>"In general, the course is all like reading out loud the textbook."</p> <p>"The textbook was not that helpful when it came to the homework problems."</p> <p>"The chapters in the Think Python textbook are too brief."</p> <p>"I am at home i have a kid so i was unable to study and doing the assignments in time."</p>

the bluer (redder) colors indicate higher negativities (positivities). Topics of the same color tend to link closely, while topics of varied colors are commonly disconnected. For example, *grading* is closely associated with *problem solving and practices*, *programming projects*, *programming style and skill*, *course quality*, and *programming textbooks*, indicating that learners who write about grading issues tend to complain as well about programming issues, practices, and course quality. Such graphical topic correlation analysis enhances our understanding of possible causalities for MOOC learner dissatisfaction.

4.5. The Moderating Effect of Course Grades

Given the interaction term of *Negative* with *CourseGrade* in our model, we examined the variations in negative topics' proportions across course grades (Figure 6), with the *x*- and *y*-axes indicating course grades and expected topic proportions, respectively. The blue and red lines represent negative and positive reviews, separately, with dashed lines depicting the 95% confidence interval.

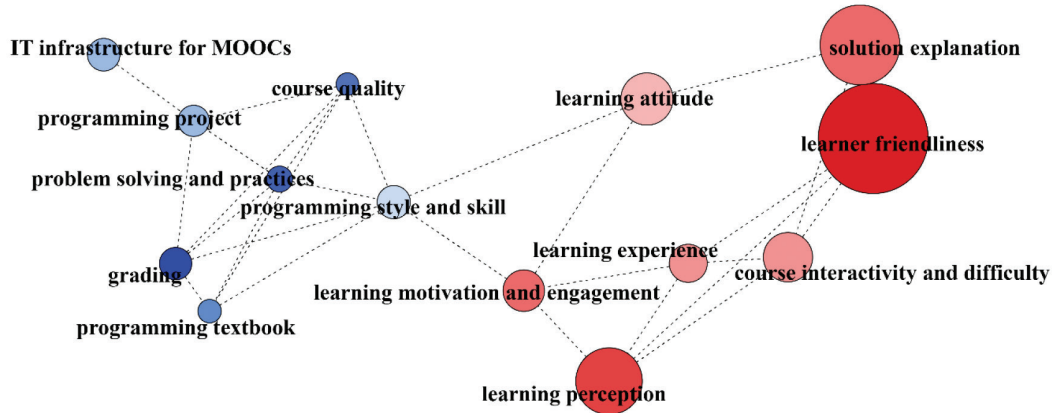


Figure 5 | Graphical topic correlations.

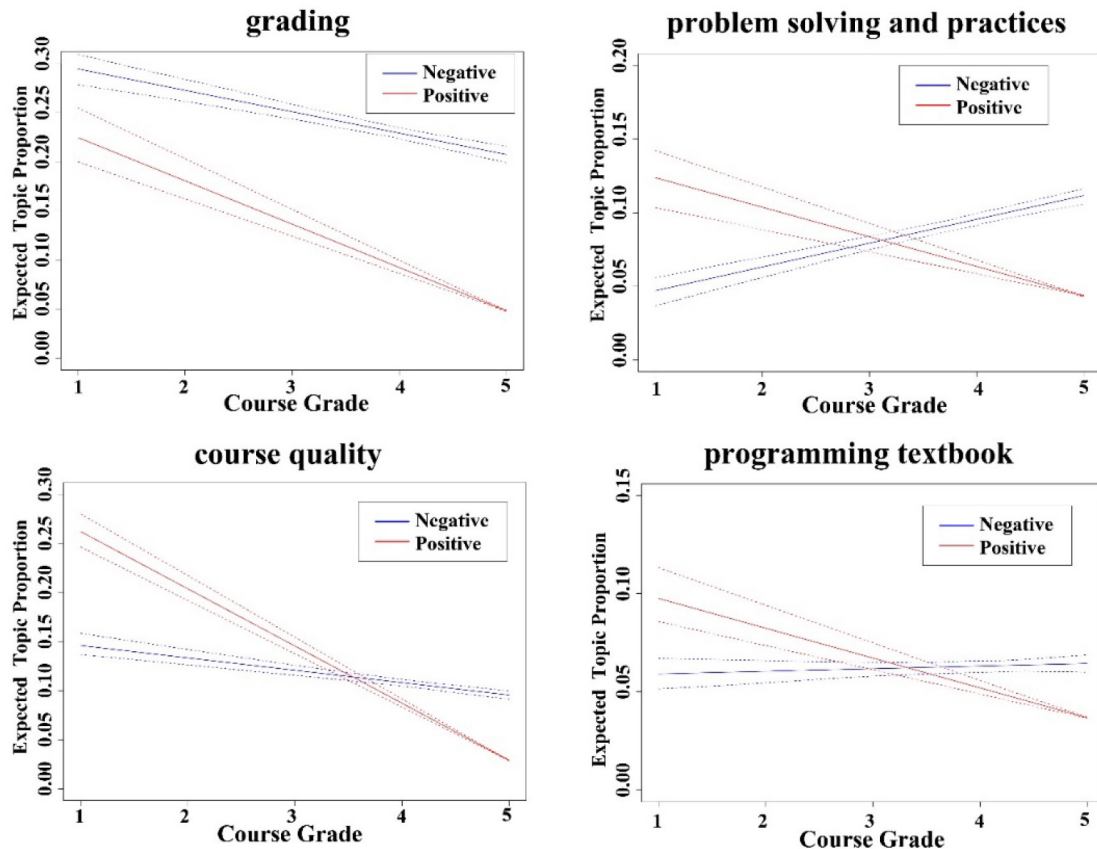


Figure 6 | Moderating effects of course grades.

Results indicate that the popularities of two negative topics (*problem solving and practices* and *programming textbooks*) in negative reviews increase alongside the increased course grades. For example, regarding *problem solving and practices*, its popularity in negative reviews increases significantly from about 5% (*CourseGrade* = one) to 10% (*CourseGrade* = five). On the contrary, its popularity in positive reviews experiences a significant decrease alongside the increased course grades. The popularities of the other two negative topics (i.e., *grading* and *course quality*) in the negative reviews decrease alongside the increased course grades. For example, the popularity of *course quality* in negative reviews decreases from about 25% (*CourseGrade* = one) to 5% (*CourseGrade* = five).

The above results indicate that the major sources of MOOC learners' dissatisfaction are issues concerning *grading*, *practices*, and *textbooks*. For high-graded courses, *problem-solving and practices* and *programming textbooks* are the main causes for learner dissatisfaction, while for low-graded courses, *grading* and *course quality* are the primary causes.

5. DISCUSSION AND CONCLUSION

This first in-depth study applies STM on MOOC learner reviews to explore the major causes for learner complaints. We identified two most frequently mentioned topics (i.e., learner friendliness and solution explanation) and four negative topics (i.e., grading, problem solving and practices, course quality, and programming textbook design) appearing significantly more in negative reviews as compared to positive ones in statistics. Such analysis contributes to MOOC learner dissatisfaction literature by allowing learners' true complaints to be heard. The followings are discussions of the above six topics, based on which we provide suggestions to MOOC providers and lecturers to promote learner satisfaction.

Learner friendliness, which may involve system/service use and content learning, is the top factor mentioned by learners. A possible reason is that most learners use MOOCs for self-learning of programing. Thus, courses that provide basic programing knowledge are more suitable and are perceived as satisfactory among learners. Another possible explanation is that learners might have inadequate experience in MOOC learning and are unfamiliar with MOOC environments. Thereby, not all learners are skillful or capable of fully utilizing MOOC functionalities and services, leading to fear in MOOC use [20] and thus lower satisfaction.

Grading is another essential factor that affects learners' satisfaction toward MOOC learning. Such a result is consistent with the finding of Ramesh et al. [21]. To improve learners' satisfaction toward grading, technicians should pay attention to submission, file uploading, website execution, and grader settings to provide pleasant experiences for learners in dealing with assignments. Second, automatic graders' precision can be improved by using advanced deep learning algorithms to avoid unfairness or errors in grading. Additionally, retrying opportunities can be provided to learners.

Similar to Chen et al. [22] and Hew [23], we found that problem-solving is frequently mentioned in negative reviews. The following improvements can be considered. First, problems should be clearly

expressed using precise rather than ambiguous wording. Second, the difficulty level of problems and practices should be adjusted appropriately to suit the course level. Furthermore, enough support from instructors should be provided to learners in solving problems, where instructors ought to be always available in discussion forums to answer learners' questions and give specific rather than general tips. Also, questions should be switched at moderate speed to allow enough time for learners to think and answer. In addition, solution explanation, as a top concern among MOOC learners, should also be carefully considered during instructors' problem-solving for learners. This is in line with Hew [8] that highlights the significance of problem-centered learning with clear expositions for engagement promotion in online courses. In programing learning, practical coding tasks and activities are commonly included to promote learners' understanding of concepts and knowledge. A problem-centered instructional strategy thus focuses on teaching learners concepts and skills via solving real-world tasks and problems. This is usually realized by instructors through performing simple tasks and offering explicit explanations and step-by-step instructions of specific coding procedures and relevant examples and non-examples of concepts. For example, for building an interactive game, instructors should specifically organize courses to assist learners in learning required concepts. Such procedure essentially contributes to bringing tangible meanings to the concepts and principles learned and motivates learners' learning interests.

Course quality is also identified to have a significant effect on learner satisfaction. Such a finding is consistent with previous studies (e.g., [24–26]), all of which indicated course quality as a key factor contributing to MOOC's successful execution. For example, Albelbisi [24] found that course quality significantly influenced learners' satisfaction in MOOC learning. Yang et al. [25] identified course quality as an essential factor to gauge MOOC success. Such a result provides evidence that courses with higher quality usually contribute to higher MOOC learners' satisfaction. The quality of MOOC courses, for example, design, output appropriateness, content, and ease of course material understanding, influences learner satisfaction and MOOC success. To improve course quality, sufficiently high-quality code examples should be provided to enhance learners' understanding of particular knowledge points. Also, instructors should provide sufficient explanations of important concepts and functions but avoid unnecessary repetitions of simple knowledge points. Furthermore, course content should be designed to suit target learners. For example, for basic courses, the content should be conveyed in a concise and easy-to-understand format. Additionally, there is a need to avoid including too many broken and unnecessary links, especially those with many advertisements.

A teaching textbook is an important carrier of a comprehensive compilation of course content [27] and has been found to be frequently mentioned in negative reviews. As suggested by DeLone and McLean [28], information quality can positively affect learners' satisfaction and their intention to learn. The positive effect of textbooks, especially those good textbooks that are written by instructors, are considered to be valuable since textbook use benefits learners as references of standard sources [29]. To improve programming textbook design, the textbooks should contain corresponding knowledge points that are helpful and

have reference value for learners in solving homework problems. Second, sufficient information should be contained in textbooks to cover the target knowledge of courses. Additionally, real-world information should be included to avoid textbooks being over-theoretical to enhance practical value. A new form of textbooks, that is, interactive textbooks, refers to “material involving less text, and instead of having extensive learning-focused question sets, animations of key concepts, and interactive tools (p. 2) [30].” Research indicates that learners who spare more effort in activities using interactive textbooks tend to achieve better performance in examination in comparison to those using traditional print textbooks. Thus, MOOC providers and instructors can consider providing self-written textbooks in an interactive format.

Additionally, we reveal MOOC learner complaints across MOOCs with varied grades. Results indicate that for high-graded MOOCs, learner complaints relate primarily to problem-solving, practices, and programming textbooks, whereas learners of low-graded MOOCs are frequently annoyed by issues about grading and course quality. Accordingly, providers of high-graded MOOCs are suggested to promote problem-solving efficiency and quality of practices and programming textbooks; providers of low-graded MOOCs should pay more attention to course quality and grading improvement.

The correlations among topics indicate the nature of MOOC learner (dis)satisfaction. For example, grading, as the most common negative topic, relates closely with programming textbooks and problem-solving and practices. An examination of specific review comments indicates that many learners complain that the grading of assignments and practices by autograders is often inconsistent with the knowledge points described in textbooks, causing confusion and uncertainties. Therefore, learners who write about grading tend to complain as well about programming textbooks or problem solving and practices. The graphical topic correlation analysis offers corpus-level implications about organizational structures and facilitates the understanding of why learners complain.

In sum, our study helps understand what causes MOOC learners to be dissatisfied and provides actionable suggestions on MOOC management. To improve MOOC learner satisfactions, MOOCs and online review websites can exploit automatic text mining methods like STM. Considering the dynamic nature of learner demands, MOOC managers ought to be constantly aware of learners’ learning experiences and feedback derived from their reviews. By integrating time, MOOC properties, and learner characteristics into topic models, managers are able to trace learner (dis)satisfaction dynamics and understand their MOOCs’ strengths and weaknesses.

This study has limitations. Firstly, we used reviews of programming MOOCs from one platform. For the purpose of result generalization, there is a need to extensively use MOOCs of different subjects and from various platforms to enhance the model’s effectiveness and ability. Considering that courses of different subjects might have different characteristics, further comparisons of topics hidden in course reviews of different subjects should also be conducted. Secondly, we included review extremities and MOOC grades as STM covariates. Future work is suggested to consider more covariates, for example, learner characteristics, to allow more fine-tuned insights into MOOC development.

CONFLICTS OF INTEREST

The authors declare they have no conflicts of interest.

AUTHORS’ CONTRIBUTION

Xieling Chen contributed in the design and methodology of the study. Gary Cheng and Haoran Xie contributed in data analysis and interpretation. Guanliang Chen and Di Zou contributed in manuscript preparation.

ACKNOWLEDGMENTS

The research described in this article has been supported by the One-off Special Fund from Central and Faculty Fund in Support of Research from 2019/20 to 2021/22 (MIT02/19-20), the Interdisciplinary Research Scheme of the Dean’s Research Fund 2019-20 (FLASS/DRF/IDS-2) and the Research Cluster Fund (RG 78/2019-2020R) of The Education University of Hong Kong, General Research Fund (No. 18601118) of Research Grants Council of Hong Kong SAR, and Lam Woo Research Fund (LWI20011) of Lingnan University, Hong Kong.

REFERENCES

- [1] K.F. Hew, X. Hu, C. Qiao, Y. Tang, What predicts student satisfaction with MOOCs: a gradient boosting trees supervised machine learning and sentiment analysis approach, *Comput. Educ.* 145 (2020), 103724.
- [2] S. Molinillo, A. Fernández-Morales, J.L. Ximénez-de-Sandoval, A. Coca-Stefaniak, Hotel assessment through social media—TripAdvisor as a case study, *Tour. Manag. Stud.* 12 (2016), 15–24.
- [3] X. Chen, D. Zou, G. Cheng, H. Xie, Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: a retrospective of all volumes of *Computers & Education*, *Comput. Educ.* 151 (2020), 103855.
- [4] F.F. Lubis, Y. Rosmansyah, S.H. Supangkat, Topic discovery of online course reviews using LDA with leveraging reviews helpfulness, *Int. J. Electr. Comput. Eng.* 9 (2019), 426–438.
- [5] S. Liu, X. Peng, H.N.H. Cheng, Z. Liu, J. Sun, C. Yang, Unfolding sentimental and behavioral tendencies of learners’ concerned topics from course reviews in a MOOC, *J. Educ. Comput. Res.* 57 (2019), 670–696.
- [6] M.E. Roberts, B.M. Stewart, D. Tingley, STM: R package for structural topic models, *J. Stat. Softw.* 10 (2014), 1–40.
- [7] N. Hu, T. Zhang, B. Gao, I. Bose, What do hotel customers complain about? Text analysis using structural topic model, *Tour. Manag.* 72 (2019), 417–426.
- [8] K.F. Hew, Promoting engagement in online courses: what strategies can we learn from three highly rated MOOCs, *Br. J. Educ. Technol.* 47 (2016), 320–341.
- [9] X. Chen, J. Chen, G. Cheng, T. Gong, Topics and trends in artificial intelligence assisted human brain research, *PLoS One* 15 (2020), e0231192.
- [10] B.G. Gameel, Learner satisfaction with massive open online courses, *Am. J. Distance Educ.* 31 (2017), 98–111.

- [11] K. Popper, *The logic of scientific discovery*, Routledge, London and New York, 2005.
- [12] S. Ackroyd, J.A. Hughes, *Data collection in context*, Longman Group UK Ltd., London, UK, 1992.
- [13] X. Peng, Q. Xu, Investigating learners' behaviors and discourse content in MOOC course reviews, *Comput. Educ.* 143 (2020), 103673.
- [14] A. Ezen-Can, K.E. Boyer, S. Kellogg, S. Booth, Unsupervised modeling for understanding MOOC discussion forums: a learning analytics approach, *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, Association for Computing Machinery, New York, USA, 2015, pp. 146–150.
- [15] X. Peng, S. Liu, Z. Liu, W. Gan, J. Sun, Mining learners' topic interests in course reviews based on like-LDA model, *Int. J. Innov. Comput. Inf. Control.* 12 (2016), 2099–2110.
- [16] D.M. Blei, Probabilistic topic models, *Commun. ACM* 55 (2012), 77–84.
- [17] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent dirichlet allocation, *J. Mach. Learn. Res.* 3 (2003), 993–1022.
- [18] M.E. Roberts, B.M. Stewart, D. Tingley, C. Lucas, J. Leder-Luis, S.K. Gadarian, et al., Structural topic models for open-ended survey responses, *Am. J. Pol. Sci.* 58 (2014), 1064–1082.
- [19] A. Garcia-Rudolph, S. Laxe, J. Saurí, M.B. Guitart, Stroke survivors on twitter: sentiment and topic analysis from a gender perspective, *J. Med. Internet Res.* 21 (2019), e14077.
- [20] R.I. Chang, Y.H. Hung, C.F. Lin, Survey of learning experiences and influence of learning style preferences on user intentions regarding MOOCs, *Br. J. Educ. Technol.* 46 (2015), 528–541.
- [21] A. Ramesh, S.H. Kumar, J. Foulds, L. Getoor, Weakly supervised models of aspect-sentiment for online course discussion forums, *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, vol. 1: Long Papers, Association for Computational Linguistics, Beijing, China, 2015, pp. 74–83.
- [22] X. Chen, D. Zou, H. Xie, G. Cheng, What Are MOOCs Learners' Concerns? Text Analysis of Reviews for Computer Science Courses, in: Y. Nah, C. Kim, S.Y. Kim, Y.S. Moon, S.E. Whang (Eds.), *Database Systems for Advanced Applications, DASFAA 2020 International Workshops, DASFAA 2020*, Springer, Cham, 2020, pp. 73–79.
- [23] K.F. Hew, Unpacking the strategies of ten highly rated MOOCs: implications for engaging students in large online courses, *Teach. Coll. Rec.* 120 (2018), 1–40.
- [24] N.A. Albelbisi, Development and validation of the MOOC success scale (MOOC-SS), *Educ. Inf. Technol.* 25 (2020), 4535–4555.
- [25] M. Yang, Z. Shao, Q. Liu, C. Liu, Understanding the quality factors that influence the continuance intention of students toward participation in MOOCs, *Education. Tech. Research Dev.* 65 (2017), 1195–1214.
- [26] J.M. Kevan, M.P. Menchaca, E.S. Hoffman, Designing MOOCs for success: a student motivation-oriented framework, *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, Association for Computing Machinery, Edinburgh, UK, 2016, pp. 274–278.
- [27] C. Zhang, S. Yi, B. Rao, An empirical study on the satisfaction of new form textbook based on ACSI, *Proceedings of the 5th International Conference on Distance Education and Learning*, Association for Computing Machinery, Beijing, China, 2020, pp. 128–132.
- [28] W.H. DeLone, E.R. McLean, The DeLone and McLean model of information systems success: a ten-year update, *J. Manag. Inf. Syst.* 19 (2003), 9–30.
- [29] P. Adamopoulos, What makes a great MOOC? An interdisciplinary analysis of student retention in online courses, *Proceedings of the 34th International Conference on Information Systems, ICIS 2013*, AIS, Milano, Italy, 2013.
- [30] A.D. Edgcomb, F. Vahid, R. Lysecky, A. Knoesen, R. Amirtharajah, M.L. Dorf, Student performance improvement using interactive textbooks: a three-university cross-semester analysis, *2015 ASEE Annual Conference & Exposition*, Seattle, Washington, 2015, pp. 26.1423.1–26.1423.17.