Developing a Multidimensional Framework for Analyzing Student Comments in Wikis

Xiao Hu[1*], Christy Weng-Lam Cheong[2] and Samuel Kai-Wah Chu[1]
[1]Faculty of Education, University of Hong Kong, Hong Kong // [2]Macao Polytechnic Institute, Macao //
xiaoxhu@hku.hk // wlcheong@ipm.edu.mo // samchu@hku.hk
[*]Corresponding author

(Submitted June 13, 2016; Revised January 27, 2017; Accepted June 23, 2017)

ABSTRACT

This study develops a framework for analyzing student comments in Wikis of group writing to inform learning assessment. It first drew on the literature to develop a framework consisting of three modules measuring student interaction, meaning construction and thinking development in the writing process. In-service teachers were interviewed to ensure framework practicality and inform subsequent refinement. A sample of 1,482 Wiki page comments was collected from 48 groups of secondary school students in Hong Kong to test the developed framework. Statistical analyses and association rule mining were conducted to the coded data to explore the relations among coding categories. This study aims to raise the attention on page comments in the analysis of student activities in Wiki and provided empirical evidence on category relations, which will be instructive for further research and practice in Wiki-supported learning.

Keywords

Wiki, Student comments, Social interaction, Meaning construction, Thinking development

Introduction

Wiki, as a social media application, allows users to develop contents collaboratively. It is regarded as a useful tool to facilitate project-based learning activities (Li, Chu, & Ki, 2014; Lo, 2013; Wang, 2014). An increasing number of studies are devoted to explore its usage and affordance, demonstrating its values in strengthening student collaboration and facilitating knowledge acquisition (e.g., Aydin & Yildiz, 2014; Cullen, Kullman, & Wild, 2013).

A Wiki is made up of pages contributed by users. Each page consists of content and comments. The content part is where group writing is developed and presented, and for which revision history is tracked. The comment part is where individual users may leave short messages for their collaborators. According to Du, Chu and Chan (2016), comments on Wiki pages are closely related to various activities students perform in Wikis. Some comments facilitate communications by criticizing ideas; some point out writing issues, and others facilitate group coordination and collaborations (Du et al., 2016). When responding to each other’s comments, students may possibly engage in further discussions or page revisions, leading to an impact on the effectiveness or quality of collaborative writing (Judd, Kennedy, & Cropper, 2010; Woo, Chu, & Li, 2013). Figure 1 illustrates an example Wiki page with comments.

Wiki data are potentially useful for analyzing and monitoring students’ engagement and writing progress. The question is then how these data can be effectively analyzed to inform learning assessment. Existing studies on Wikis in education mainly focus on page content and edits (e.g., Macfadyen & Dawson, 2010; Romero-Zaldivar,Pardo, Burgos, & Kloos, 2012; Romero, López, Luna, & Ventura, 2013) while page comments are largely ignored. Analyses on page content and edits aimed to assist teachers in formative and summative assessment (e.g., McKenzie et al., 2013; Williams, 2014) and in identifying and monitoring students’ learning behaviors (e.g., Berland, Martin, Benton, Smith, & Davis, 2013; Brooks, Erickson, Greer, & Gutwin, 2014; Tobarra, Robles-Gómez, Ros, Hernández, & Caminero, 2014). However, few studies have investigated students’ comments made on Wiki pages. Students’ page comments in Wiki can provide additional evidences of student interactions and contributions. With proper processing, they can be made use by teachers in understanding, assessing and monitoring students’ learning. This study attempts to fill this research gap by developing a framework for analyzing students’ page comments in Wiki that can inform learning assessment. Specifically, the following research questions will be answered:
RQ1: How can student comments be categorized for facilitating teachers in learning assessment?
RQ2: Are there any relations among the categories of student comments?

To answer RQ1, a categorization framework was developed from the literature. In-service teachers were interviewed to inform modifications. The refined framework was then tried out on a sample of secondary school students' comments on their group project Wikis to inform further framework refinement. To answer RQ2, statistical analyses and association rule mining were conducted to the coded data, enabling discussions on the relationships between coding categories. This study helps raising the attention on page comments in the analysis of student activities in Wiki and providing empirical evidence on the theoretical and practical values of the proposed framework. This will be instructive for research and practice in Wiki-supported learning.

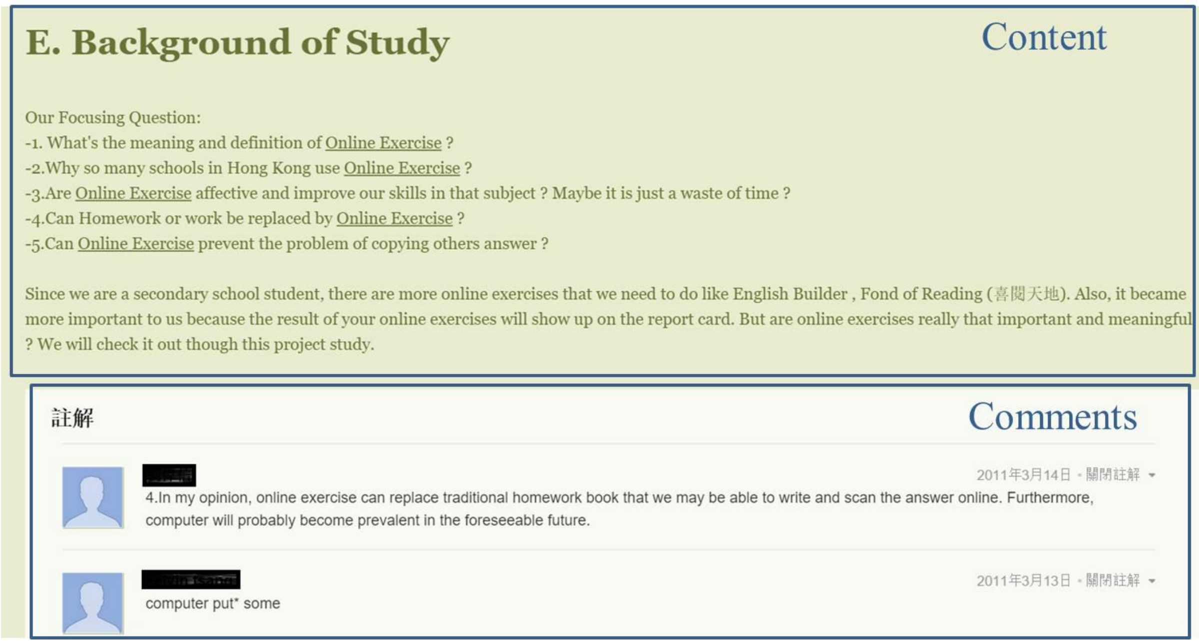

Figure 1. An example Wiki page with comments

Literature review

Wiki as a tool for collaborative learning

Collaboration, communication, critical thinking and problem solving are among the identified 21st century skills (Bruett, 2006; Partnership for 21st Century Skills, 2009). Wiki has been found to be a powerful platform that facilitates training and promoting these skills, for it can support an autonomous, collaborative and inquiry-oriented learning environment (Hakkarainen & Sintonen, 2002). In Wiki, students collaboratively write on the same pages, through communication and negotiation, among themselves and with teachers. This process leads to the co-construction of knowledge, as grounded in the social constructivist theory (O'loughlin, 1992). By allowing students to comment and modify each other's writing in a cumulative fashion, Wiki motivates and helps students practice critical thinking (Cabiness, Donovan, & Green, 2013; Wake & Modla, 2012). In particular, through commenting on Wiki pages, students were found to have enhanced their social competences and metacognitive skills (Notari & Doebeli, 2012).

Current research on student assessment in the Wiki context seems to fall into two main streams. The first stream tends to suggest ways of integrating student peer assessment into the assessment framework (e.g., Šerbec, Strnad, & Rugelj, 2010; DeWever, van Keer, Schellens, & Valcke, 2011). The aim is more to ensure a realistic and fair assessment rather than to facilitate a comprehensive formative assessment that can trigger timely and appropriate teacher intervention during the learning process. The second stream focuses on developing systems exploiting Wiki data to inform continual assessment (e.g., Kubincová, Homola, & Janajev, 2012; Palomo-Duarte, Dodero, Medina-Bulo, Rodríguez-Posada, & Ruiz-Rube, 2014). These studies tend to be platform-specific and quantitative in nature, with a scope largely confined to the page edits. While useful, they fall short in providing comprehensive information about the quality of student learning.

It is important to be aware that student collaboration resides in the composite of activities enabled in Wiki. Studies found that page commenting may help promoting group collaboration and revealing group dynamics (Judd et al., 2010). Woo et al. (2013) further indicated that page comments were sometimes "revision-oriented" triggering page edits that led to enhancement of writing quality. It is therefore necessary to extend the focus from page contents and edits to page comments for a more all-round review of student activities in Wiki.

Student online discussions

Student comments in Wiki are similar to discussions on online forums to some extent, especially those forums designed for inquiry-based collaborative learning (Tirado, Aguaded, & Hernando, 2011). While the former is rarely studied, numerous studies have been conducted to investigate the dynamics of student online discussions. Many of them analyzed the content of student postings by categorizing them according to a range of dimensions. Pena-Shaff and Nicholls (2004) developed a framework of interactive learning to encode student postings in an online discussion board into "meaning construction" categories such as clarification, elaboration and interpretation. They also considered the interactivity among students and differentiated individual reflections from conversational interactions. Tirado et al. (2011) analyzed the activity records of an online discussion forum from the perspectives of "psycho-social relations," "positive interdependence" and "construction of meaning." Similarly, Xie and Ke (2011) conducted a content analysis of student online discussions in terms of "social interaction," "knowledge construction" and "regulation of learning" (i.e., the coordination and management of the collaborative learning process). They observed that students tended to be involved in lower order cognitive activities when interacting with others but higher order cognitive activities when working on their own. In So's study (2009), student online discussions during group projects were analyzed with regard to collaborative learning behaviors and social presence behaviors. They found that most discussions were about group work facilitation while relatively few activities were challenging and explaining/elaborating.

These studies considered both social and cognitive dimensions of student interactions which also reside in Wiki commenting. These dimensions serve well as the theoretical basis for this study. However, unlike online forums where the discussion itself is the expected artefact, Wiki commenting is to facilitate the co-construction of the Wiki page contents. Therefore, besides idea exchange and negotiation, it is essential for Wiki comments to be (1) able to help sustain group interaction and collaboration (Judd et al., 2010), (2) relevant and contributive to the overall discussion about the project in question, and (3) able to reflect the quality of thinking of the commenter (Woo et al., 2013). In addition, the commenting area of Wiki platforms are often designed much simpler than fully-fledged discussion forums (e.g., without subject lines). With these differences on purposes, functionalities and interfaces, whether and to what extent the existing frameworks developed for online forum discussions can be used to effectively analyze Wiki comments remain open questions. In this study, therefore, we draw from these studies to develop a multi-focal framework for page comment analyses adequate and adaptable at different project phases.

Methods

This study adopts an iterative approach using a combination of literature review, stakeholder interview, experimental coding, statistical analysis and association rule mining, a data mining approach to find out relationships between items or categories in a dataset. An initial framework was first developed from the literature (FW V0) and then modified based on interviews with in-service teachers (FW V1). A sample of students' page comments on group project Wikis were collected and coded with the refined framework (FW V1). The distribution of the coded data informed yet another round of framework revision (FW V2). Association rule mining was then applied to the data coded with FW V2 to find out relationships among the categories. The results informed the final round of framework revision (FW V3).

Initial framework development

According to the three requirements mentioned in the previous section, related frameworks and/or taxonomies on the following three aspects in the literature of student online asynchronous discussions were adopted: Social Interaction (Bales, 1950; Tirado et al., 2011), Meaning Construction (Pena-Shaff & Nicholls, 2004), and Thinking Development (Krathwohl, 2002). A composite framework (FW V0) consisting of three modules (i.e., Social Interaction, Meaning Construction and Thinking Development) was consequently developed and presented in the Appendix.

Interviews with in-service teachers

To examine the practicality of this initial framework and to further refine it, the opinions of seven in-service secondary school teachers were sought via five semi-structured interviews. Convenience sampling was adopted to invite the interviewees. Among them, four were invited to inform the refinement (i.e., "exploration group") and three to comment on the refined scheme to ensure its feasibility and practicality (i.e., "evaluation group"). Both groups were asked (1) what they would need to know about student comments in Wikis; (2) how they thought of the framework in relation to their practical needs; and (3) how the framework can be modified. The evaluation group was also asked to comment on the refined framework developed during the exploration stage. Because of the difference in purpose, interviews were conducted in pairs with the exploration group to allow collective brainstorming, discussion and a certain extent of consensus building. Interviews with the evaluation group were conducted individually to ensure independent judgment. Table 1 summarizes the demographic characteristics of the interviewed teachers. The refined framework after the interviews is referred to as FW V1 and is presented in the Appendix.

Table 1. Demographic information of the interviewees

	Exploration group	Evaluation group
Gender		
Male	2	1
Female	2	2
Teaching experience		
1-3 years	1	1
8-10 years	1	1
Over 10 years	2	1
Subject fields(1)		
Humanities (Chinese, English, Design)	2	2
Social Science (History)	1	0
Science (Chemistry, Computer, Geography)	2	1

Note. (1) One interviewee teaches more than one subject.

Experimental coding

To test the applicability of FW V1, we applied it to a sample of student comments collected from the Wikis developed by 48 student groups from a junior secondary school in Hong Kong. Of the 48 groups, 30 were from Form 1 (equivalent to Grade 6) and 18 were from Form 2 (Grade 7). The total number of students involved was 238. Their age ranged from 12 to 14 during the study period. Each group consisted of about five students who were required to collaborate on an inquiry-based project for the Liberal Studies course on a five-month period. A Google Site was created exclusively for each group to facilitate collaboration. The groups were required to write their project reports on their Google Sites using Wiki pages to differentiate sections (e.g., introduction, methodology). During the project period, the students could post comments on each page and use this feature to communicate and discuss with one another. For this study, the comments attached to each page, which are predominantly written in English, were collected. Among the 48 Wikis, 7 contained no comments (6 from Form 1 and 1 from Form 2) and 1 had access control in place. They were thus removed from the sample.

The unit of analysis varied in existing studies involving content analysis of student online discussions, including unit of meaning (Bales, 1950), sentence (Pena-Shaff & Nicholls, 2004) and postings (Weltzer-Ward, 2011). To sustain the comparability across modules, sentence was used in this study as the unit of analysis. In case of grammatical irregularities, an operational definition of sentence, a set of words resembling a simple and coherent utterance, was adopted. Table 2 provides a descriptive summary of the sample that consists of 1,482 units.

Table 2. A descriptive summary about the dataset

	Overall	Form 1	Form 2
Total number of groups	48	30	18
Number of groups removed before analysis	8	7	1
Number of groups to be studied	40	23	17
Total number of units	1,482	1,056	426

All units identified were coded manually according to FW V1 and the results were used to inform further refinement of the framework to FW V2. To ensure coding quality, comments from one fourth of student groups randomly selected in both Forms were double-coded by a second independent coder. Cohen's Kappa was calculated to measure the level of interrater reliability. The distribution of the coded data across different categories were then analyzed and compared to findings in related literature.

Relationship between comment categories

Association rule mining is a data mining technique used for identifying associations among frequently appeared patterns in a dataset (Han, 2012). It has been used in the education domain to find out relationships between variables, particularly in datasets with many variables (e.g., student emotion status, learning performances) (Baker, 2010). Unlike correlation analysis that is bivariate, association rule mining can discover relationships among multiple variables at the same time. Specifically, association rule mining aims to find "if-then" rules of the variables, in the form of "antecedent → consequence," where antecedent and consequence are conditions that some variable(s) has certain value(s). This study applied association rule mining to explore the associations among categories across the three modules. For example, a possible rule in this study might be "a comment is in SI-0 -> the comment is in MC-0." That is, if a comment is categorized into the SI-0 category, chances are it is also in the MC-0 category.

To identify interesting and significant rules, the FP-Growth algorithm (Han, 2012) was used with a minimum Support value of 0.50 and minimum Lift value of 2.0. Rules satisfying the criteria are defined as "interesting" ones (Han, 2012). In addition, the "Cosine" measures of interestingness proposed by Merceron and Yacef (2008) for association rules in educational data were adopted. The "Cosine" interestingness threshold is set on the level larger than 0.65. These threshold values were set with tradeoffs between frequency of occurrences and number of resultant rules.

Results

Refined framework FW V1

In view of FW V0, the interviewees in the exploration group agreed that the three modules in the framework were necessary. However, they were concerned of the categorization complexity, which may turn out to be impractical as it may take much time for teachers to understand the categories and interpret the results. Some categories were regarded rare among their students. An example was the category "reflection," which they thought students would seldom do upon commenting. The interviewees also agreed that it was necessary to combine categories similar in nature to make the framework more practically feasible. It was noted that more detailed categorization would be feasible or needed either when more resources (e.g., time and manpower) are made available to teachers or when student comments reflect a strong inclination toward a particular category.

To refine the scheme, possible code combinations were proposed to solicit discussion within the two exploration groups who also suggested combinations they found fit for their needs. Consensus was reached after each pair interview and the opinions from the interviews were consolidated to inform framework refinement. The refinement decisions made at this stage are summarized as below.

In Social Interaction, three sets of codes were combined due to their shared natures, namely, (1) SI-1 to SI-3 (Give Suggestions/Information/Opinions) being combined into SI-A (Giving Acts) as they represents different types of giving acts; (2) SI-4 to SI-7 (Ask for Suggestions/Information/Opinions/Help) being combined into SI-B (Make Requests) as they represents different types of requests; and (3) SI-11 (Encouragement) and SI-12 (Others) being combined into SI-E (Others) as they both concern students' ability in socializing. There were different opinions on whether SI-8 (Agree) and SI-9 (Disagree) should remain as individual categories. When commenting on the Meaning Construction module, one interviewee noted that student conflicts (i.e., MC-5) may need teachers' special intervention. In line with this opinion, these two categories were kept separate while SI-9 (Disagree) was combined with SI-10 (Show Antagonism) into SI-D (Disagree) as they are both acts of negating others.

In Meaning Construction, six codes, namely, MC-2 (Reply), MC-3 (Clarification), MC-4 (Interpretation), MC-6 (Assertion), MC-8 (Judgment), and MC-10 (Support), which involve acts of responding (to previous questions, statements or perspectives) or providing feedback were combined into MC-X (Responses). MC-7 (Consensus Building) and MC-9 (Reflection) were combined into MC-D (Result) as they indicate some kind of results (of a discussion and/or a learning process).

In Thinking Development, the interviewees were asked whether they preferred two categories (i.e., either there is cognitive development occurred or not) or three categories (i.e., no, low or high level of cognitive development). They tended to find dualism arbitrary and were inclined to have three categories. So TD-1 to TD-3 (Remember, Understand, and Apply) were combined into TD-L representing low level of cognitive development, and TD-4 to TD-6 (Analyze, Evaluate, and Create) into TD-H representing high level of cognitive development.

The feasibility and practicality of the consequent refined framework (FW V1) were confirmed in the evaluation group interviews. The interviewees generally expressed that they would be more willing to use the refined framework and knowing the overall picture would be sufficient in view of their heavy workload. One interviewee suggested the Thinking Development module be divided into four categories (no, low, medium or high cognitive development). This suggestion was kept for further consideration in the next round of framework refinement.

Trial coding and FW V2

FW V1 was tried out with the page comments collected from the student Wikis, to verify the framework and further refine it to FW V2. To ensure coding reliability, 189 randomly selected units were coded with two independent coders. Cohen's kappa reached 0.821, 0.839 and 0.848 for Social Interaction, Meaning Construction and Thinking Development respectively, indicating very good interrater reliability (Altman, 1991). The refinement decisions made at this stage are summarized below.

In Social Interaction, the statistics of the coded data indicated SI-A (Giving Acts) accounted for about 70% of the comment units. To avoid such strong inclination, SI-A (Giving Acts) was segregated back to SI-1 (Give Suggestions), SI-2 (Give Opinions) and SI-3 (Give Information) in FW V2.

In Meaning Construction, the statistics of the coded data indicated MC-X (Responses) accounted for half of the comment units. Thus, the categories included in MC-X were rearranged in FW V2: MC-2 (Reply) was divided into two according to the complexity of the replies. The type of responses that are simple and straightforward in nature was combined with MC-10 (Support) to form the new category MC-B (Simple Replies). Those that are not simple was combined with MC-3 (Clarification), MC-4 (Interpretation), MC-6 (Assertion) and MC-8 (Judgment) to form another new category, MC-F (Argument) to capture comments that are argumentative in nature.

In Thinking Development, the question was whether to have three or four categories. The statistics of the coded data indicated the comment units reflecting the presence of cognitive development are of the small minority (around 30%). The small sample size makes it hard to judge whether finer categorization would enhance the expressiveness of the results. Therefore, no changes were made and three categories were kept in this module.

The framework developed up to this stage (FW V2) is presented in the Appendix. In comparison with those commonly adopted in online discussion research, this framework emphasizes teachers' practical needs, and the three modules are particularly helpful in supporting separate analyses of student comments from different perspectives of concern.

The distribution of the coded data across the categories in FW V2 is shown in Table 3. It is noteworthy that both Forms have similar distributions. For both Forms, student comments were dominated by two to three frequent categories in each module. In Social Interaction, over 70% of the units fall under the three categories of giving acts (i.e., SI-1: Give suggestions, SI-2: Give opinions, and SI-3: Give information). In Meaning Construction, MC-0 (N/A) and MC-F (Argument) were most frequent in both Forms, with each accounting for about 30% of all comments. In Thinking Development, the dominance of TD-0 (N/A) is obvious, followed by TD-H (High Cognitive Development).

The interaction/discussion patterns as reflected in the results are comparable to the findings of So (2009), where frequent online discussion activities among students were noted to include organizing work, reporting progress, sharing task, feedback seeking/giving, and group cohesion. These activities are related to the popular categories in the coded sample (Table 3) including SI-3 (Giving Information), MC-0 (N/A) and TD-0 (N/A). Different from Xie and Ke's findings (2011) that there tended to be more lower-level than higher-level knowledge construction behaviors in online discussion, there were more TD-H (High Cognitive Development) than TD-L (Low Cognitive Development) observed in the present study. This is possibly due to the different settings of the two studies. In Xie and Ke (2011), students were required to share information and participate in discussions as part of the course assessment whereas in the present study commenting was optional and was not the only channel for information sharing and interaction. This might have affected students' motivation in posting and the consequent interaction behaviors and discussion patterns.

Table 3. Distribution of coded categories within each module in FW V2*

Module	Category	Form 1	Form 2	Overall
Social interaction	SI-0: N/A	-	-	-
	SI-1: Give Suggestions	25%	23%	25%
	SI-2: Give Opinions	26%	28%	27%
	SI-3: Give Information	21%	21%	21%
	SI-B: Ask	12%	9%	11%
	SI-C: Agree	5%	5%	5%
	SI-D: Disagree	4%	3%	4%
	SI-E: Others	7%	10%	8%
Meaning construction	MC-0: N/A	30%	20%	27%
	MC-A: Questions	13%	12%	13%
	MC-B: Simple Replies	10%	13%	11%
	MC-C: Conflict	4%	2%	3%
	MC-D: Results	5%	0%	4%
	MC-E: Others	4%	6%	5%
	MC-F: Argument	33%	47%	37%
Thinking development	TD-0: N/A	68%	66%	67%
	TD-L: Low Cognitive Development	7%	6%	7%
	TD-H: High Cognitive Development	25%	27%	26%

Note. *Categories SI-1, SI-2, SI-3, MC-B and MC-F were recoded by two coders. The interrater reliability are: SI: $\kappa = 0.872$; MC: $\kappa = 0.851$; TD: $\kappa = 0.848$.

Relationships of codes

Table 4 shows strong association rules mined from the coded dataset. All of them have Lift values larger than 2.00. Those with cosine values larger than 0.65 are recognized as very strong rules according to Merceron and Yacef (2008).

Table 4. Association rules found from the coded sample (with redundant rules removed)

Antecedent	Consequence	Support	Confidence	Lift	Cosine
TD-0, SI-1	MC-0	0.13	0.61	2.21	0.54
SI-3	TD-0, MC-0	0.13	0.61	2.24	0.54
SI-2	MC-F, TD-H	0.15	0.56	3.01	0.67
TD-H	MC-F, SI-2	0.15	0.58	2.65	0.63
MC-F	SI-2	0.22	0.6	2.25	0.71
MC-F, SI-2	TD-H	0.15	0.68	2.65	0.63

The first two rules in Table 4 involve the associations among SI-1 (Give Suggestions), SI-3 (Give Information), MC-0 (N/A) and TD-0 (N/A). In particular, the rule “SI-3 -> MC-0, TD-0” suggests that comments in SI-3 (Give Information) are likely to be in MC-0 (N/A) and TD-0 (N/A) as well. The next four rules indicate that the interrelationships between SI-2 (Give Opinions), MC-F (Argument) and TD-H (High Cognitive Development) are strong and reciprocal, suggesting that comments in each of these categories are likely to be in the other two categories as well (except for MC-F -> SI-2, TD-H).

In line with these views that SI-1 (Give Suggestions) and SI-3 (Give Information) had similar implications, to enhance the filtering capacity of the scheme, SI-1 and SI-3 were combined, leading to FW V3 as shown in the Appendix.

Discussions

Results of this study have implications in teaching practice. The fact that most comments were in MC-0 (N/A) and TD-0 (N/A) categories suggests that the comment area in Wiki platform can be enhanced. Compared to fully-fledged discussion forums where deep discussions are more often, the Wiki comment areas usually have fewer functionalities and limited affordance. For example, the comment area in Google Sites where the sample data were collected did not support multiple discussion threads, subject line, or indention of replies. Thus it was difficult to tell the discussion structure (i.e., who responded to whom). Similar issues exist on other Wiki platforms often used in schools including PBworks, Wikispaces and the Wiki activity in Moodle, one of the most popular Learning Management Systems. In contrast, MediaWiki, the platform used by Wikipedia, has a more flexible discussion area that supports multiple threads, heading for each thread, indentions within each thread and the option of collapsing/expanding arbitrary levels of replies. As a result, many in-depth discussions on complex topics such as computer programming and knowledge engineering are supported in MediaWiki (Di Francescomarino, Ghidini & Rospocher, 2012). However, it is noteworthy that more structure built in the Wiki comment area may not necessarily bring in more benefits. After all, the purpose of Wiki commenting is to facilitate collaborative writing, while the discussion itself is not the targeted outcome. Therefore, a semi-formal approach to discussion could be more proper. Complex structures such as reference links and "build-on" relationships in the Knowledge Forum (Yang, van Aalst, Chan, & Tian, 2016) might possibly impose unnecessary cognitive load to the students, even though they can greatly facilitate collaborative knowledge co-construction and creation in the knowledge building context.

The dominance of MC-0 and TD-0 also suggests that interventions are desirable during the project period. Analytic tools can be developed using the developed framework to help teachers check student comments regularly. Early intervention can then be implemented once insufficient higher-level comments are detected. Similar tools can also be used by students as a means of obtaining timely feedback that can facilitate self-regulated learning (Carless, Salter, Yang & Lam, 2011).

One of the important goals of education is to help students acquire higher-order cognitive skills as embodied by TD-H (High Cognitive Development) in the framework. As such, the association rules found (Table 4) suggest that giving information (SI-1) or suggestions (SI-3) only may not need to be particularly encouraged during collaborative projects whereas it is worthwhile to encourage students to post opinionated comments (SI-2), to present arguments (MC-F), or both in order to increase the chance of using higher-order cognitive skills (TD-H).

The relationships among categories also support the possibility of using the outcomes from one module to predict the general outcomes of the other two. For instance, an observed inclination toward SI-1 (Give Suggestions) and SI-3 (Give Information) can be taken as indicators of the possible low frequencies of MC-F (Argument) and TD-H (High Cognitive Development). In this way, schools may selectively implement one module instead of the full framework, depending on the workload of their teachers and resources available. This reflects the flexibility of the framework.

Last but not least, the similar distribution of categories across the two Forms (Table 3) can be exploited to build automated categorization tools (Rosé et al., 2008). Although Form 1 groups had more MC-0 (N/A) whereas Form 2 groups had more MC-F (Argument), results of Mann-Whitney *U* Tests (as the sample did not follow normal distribution) showed no significant difference between the two Forms throughout all categories. Comments collected from one Form can be manually annotated and used to train the categorization model. The model then can be applied to comments of the other Form to automatically generate the categories.

Conclusion and future work

Student comments in Wikis are important evidence revealing students' ongoing collaborative writing processes. The goal of this study is to develop a framework for analyzing students' page comments posted in Wikis of group writing to inform affirmative learning assessment and feedback. A three-module framework was developed from current literature and enhanced with interviews of in-service teachers. The refined scheme was tested with real-life data collected from group Wikis of students in a secondary school in Hong Kong. The coded sample was analyzed statistically in terms of distributions. Relationships across categories in different modules were also detected with association rule mining. The final product is, therefore, verified to be practically useful in providing readily interpretable information about the quality of student interaction, meaning construction and thinking development as reflected in their Wiki page comments.

This study has several limitations. First, it is contextually confined to Hong Kong where student behaviors may be distinctive as a consequent of its socio-cultural uniqueness. Second, this study was informed by data from junior secondary school students from one local school in Hong Kong. The results may be different among other student populations (e.g., senior secondary school students, students in international schools). Third, while Hong Kong is a multilingual city, a majority of its population speaks Cantonese. The data of this study comes from native speakers of Cantonese. The comment pattern might be different if they used Chinese rather than English predominantly in their comments. Therefore, more empirical studies are necessary to test the applicability of the findings in this study to other regions, other languages and other student populations.

Undoubtedly, this study focuses only on one facet of student activities in Wiki. As future work, this study will be further developed in conjunction with existing studies on Wiki page edits for a more comprehensive model of assessing Wiki activities. In addition, the evolution of student commenting behaviors during the project period may disclose meaningful patterns of the learning process and is worth further investigation.

Acknowledgements

This study was supported by an Early Career Scheme grant from the Research Grants Council of the Hong Kong Special Administrative Region, China. (Project No. HKU 27401114) and the Seed Funding Programme for Applied Research in the University of Hong Kong.

References

Altman, D. G. (1991). Inter-rater agreement. *Practical statistics for medical research*, *5*, 403-409.

Aydin, Z., & Yildiz, S. (2014). Using wikis to promote collaborative EFL writing. *Language Learning & Technology*, *18*(1), 160-180.

Baker, R. S. J. D. (2010). Data mining for education. In B. McGaw, P. Peterson, & E. Baker (Eds.), *International Encyclopedia of Education* (3rd ed., pp. 112-118). Oxford, UK: Elsevier.

Bales, R. F. (1950). A Set of categories for the analysis of small group interaction. *American Sociological Review*, *15*(2), 257-263.

Berland, M., Martin, T., Benton, T., Smith, C. P., & Davis, D. (2013). Using learning analytics to understand the learning pathways of novice programmers. *Journal of the Learning Sciences*, *22*(4), 564-599.

Brooks, C., Erickson, G., Greer, J., & Gutwin, C. (2014). Modelling and quantifying the behaviours of students in lecture capture environments. *Computers & Education*, *75*, 282-292.

Bruett, K. (2006). Why American business demands twenty-first century skills: An Industry perspective. *New Directions for Youth Development*, *2006*(10), 25-30, 9-10.

Cabiness, C., Donovan, L., & Green, T. (2013). Integrating wikis in the support and practice of historical analysis skills. *TechTrends: Linking Research & Practice to Improve Learning*, *57*(6), 38-48.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, *36*(4), 395-407.

Cullen, R., Kullman, J., & Wild, C. (2013). Online collaborative learning on an ESL teacher education programme. *ELT Journal*, *67*(4), 425-434.

DeWever, B., van Keer, H., Schellens, T., & Valcke, M. (2011). Assessing collaboration in wiki: The Reliability of university students' peer assessment. *The Internet and Higher Education*, *14*(4), 201-206.

Di Francescomarino, C., Ghidini, C., & Rospocher, M. (2012). Evaluating wiki-enhanced ontology authoring. *Knowledge Engineering and Knowledge Management*, *7603*, 292-301.

Du, H., Chu, S. K. W. & Chan, R. C. H. (2016). Collaborative writing with Wikis: An Empirical investigation on group-based learning. *Online Information Review*, *40*(3), 380-399.

Hakkarainen, K., & Sintonen, M. (2002). The Interrogative model of inquiry and computer-supported collaborative learning. *Science & Education*, *11*(1), 25-43.

Han, J. (2012). *Data mining: Concepts and techniques*. Waltham, MA: Morgan Kaufmann Publishers.

Judd, T., Kennedy, G., & Cropper, S. (2010). Using wikis for collaborative learning: Assessing collaboration through contribution. *Australasian Journal of Educational Technology*, *26*(3), 341-354.

Krathwohl, D. R. (2002). A Revision of Bloom's taxonomy: An Overview. *Theory into Practice*, *41*(4), 212-218.

Kubincová, Z., Homola, M., & Janajev, R. (2012, April). *Tool-supported assessment of wiki-based assignments*. Paper presented at the 4th International Conference on Computer Supported Education (CSEDU 2012), Porto, Portugal. Retrieved from http://dai.fmph.uniba.sk/~homola/papers/csedu2012.pdf

Li, X., Chu, S. K. W., & Ki, W. W. (2014). The Effects of a wiki-based collaborative process writing pedagogy on writing ability and attitudes among upper primary school students in Mainland China. *Computers & Education*, *77*, 151-169.

Lo, H. C. (2013). Design of online report writing based on constructive and cooperative learning for a course on traditional general physics experiments. *Educational Technology & Society*, *16*(1), 380-391.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A Proof of concept. *Computers & Education*, *54*, 588-599.

McKenzie, W. A., Perini, E., Rohlf, V., Toukhsati, S., Conduit, R., & Sanson, G. (2013). A Blended learning lecture delivery model for large and diverse undergraduate cohorts. *Computers & Education*, *64*, 116-126.

Merceron, A., & Yacef, K. (2008). Interestingness measures for associations rules in educational data. *Educational Data Mining*, *8*, 57-66.

Notari, M., & Doebeli, B. (2012). Wiki: An Archetypical tool for collaborative learning in the digital age. In S. Bocconi & G. Trentin (Eds.), *Wiki Supporting Formal and Informal Learning* (pp. 21-40). New York, NY: Nova Publishers.

O’loughlin, M. (1992). Rethinking science education: Beyond Piagetian constructivism toward a sociocultural model of teaching and learning. *Journal of Research in Science and Teaching*, *29*(8), 791–820.

Palomo-Duarte, M., Dodero, J. M., Medina-Bulo, I., Rodríguez-Posada, E., & Ruiz-Rube, I. (2014). Assessment of collaborative learning experiences by graphical analysis of wiki contributions. *Interactive Learning Environments*, *22*(4), 444-466.

Partnership for 21st Century Skills. (2009). *P21 Framework Definitions*. Retrieved from http://p21.org/storage/documents/P21_Framework_Definitions.pdf

Pena-Shaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, *42*, 243-265.

Romero, C., López, M.-I., Luna, J.-M., & Ventura, S. (2013). Predicting students’ final performance from participation in on-line discussion forums. *Computers & Education*, *68*, 458-472.

Romero-Zaldivar, V.-A., Pardo, A., Burgos, D., & Kloos, C. D. (2012). Monitoring student progress using virtual appliances: A Case study. *Computers & Education*, *58*, 1058-1067.

Rosé, C., Wang, Y. C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, *3*(3), 237-271.

Šerbec, I. N., Strnad, M., & Rugelj, J. (2010). Assessment of wiki-supported collaborative learning in higher education. In *Proceedings of the 2010 9th International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 79-85). doi:10.1109/ITHET.2010.5480060

So, H.-J. (2009). When groups decide to use asynchronous online discussions: Collaborative learning and social presence under a voluntary participation structure. *Journal of Computer Assisted Learning*, *25*, 143-160.

Tirado, R., Aguaded, I., & Hernando, A. (2011). Collaborative learning processes in an asynchronous environment: An Analysis through discourse and social networks. *Journal of Latin American Communication Research*, *2*(1), 115-146.

Tobarra, L., Robles-Gómez, A., Ros, S., Hernández, R., & Caminero, A. C. (2014). Analyzing the students’ behavior and relevant topics in virtual learning communities. *Computers in Human Behavior*, *31*, 659-669.

Wake, D. G., & Modla, V. B. (2012). Using wikis with teacher candidates: Promoting collaborative practice and contextual analysis. *Journal of Research on Technology in Education (International Society for Technology in Education)*, *44*(3), 243-265.

Wang, Y. (2014). Using wikis to facilitate interaction and collaboration among EFL learners: A Social constructivist approach to language teaching. *System*, *42*, 383-390.

Weltzer-Ward, L. (2011). Content analysis coding schemes for online asynchronous discussion. *Campus-Wide Information Systems*, *28*(1), 56-74.

Williams, P. (2014). Squaring the circle: A New alternative to alternative-assessment. *Teaching in Higher Education*, *19*(5), 565-577.

Woo, M., Chu, S., & Li, X. (2013). Peer-feedback and revision process in a wiki mediated collaborative writing. *Educational Technology Research & Development*, *61*(2), 279-309.

Xie, K., & Ke, F. (2011). The Role of student’s motivation in peer-moderated asynchronous online discussions. *British Journal of Educational Technology*, *42*(6), 916-930.

Yang, Y., van Aalst, J., Chan, C. K., & Tian, W. (2016). Reflective assessment in knowledge building by students with low academic achievement. *International Journal of Computer-Supported Collaborative Learning*, *11*(3), 281-311.

Appendix: Coding Scheme Evolution*

Source	V0 category	V1 category	V2 category	V3 category
Social Interaction (SI) Objective: To measure students' efforts devoted to maintaining and sustaining the interaction and collaboration				
Bales (1950)	SI-0: N/A No social interaction reflected		SI-0	SI-0
	SI-1: Give suggestions Propose directions, possible ways of actions	SI-A: Giving acts	SI-1	SI-A
	SI-3: Give information (1) Inform, repeat, clarify, confirm		SI-3	
	SI-2: Give opinions Express thoughts, feelings, wishes; evaluate; analyze		SI-2	SI-2
	SI-4: Ask for information (1) Request for information, repetition, clarification, confirmation	SI-B: Make requests	SI-B	SI-B
	SI-5: Ask for opinions Request for thoughts, expressions of feelings and wishes, evaluation, analysis			
	SI-6: Ask for suggestions Request for directions, possible ways of action			
	SI-7: Ask for help (2) Request for support and assistance			
	SI-8: Agree Show passive acceptance, understanding, compliance	SI-C: Agree	SI-C	SI-C
	SI-9: Disagree Show passive rejection, withhold help	SI-D: Disagree Show passive rejection, withhold help, deflate others, or defend/assert self	SI-D	SI-D
	SI-10: Show antagonism Deflate other's status, defend or assert self			
Tirado, Aguaded, & Hernando (2011)	SI-11: Encouragement (3) Cheer up other, express support, aiming to promote group cohesion	SI-E: Others Building comradeship, socialize	SI-E	SI-E
	SI-12: Others(3) Social statements such as greetings and apologies that facilitates communication			

<table>
<tr><td colspan="5">Meaning Construction (MC)
Objective: To measure the contribution and relevancy of the comments to the overall discussion about the project in question</td></tr>
<tr><td rowspan="11">Pena-Shaff& Nicholls (2004)</td><td>MC-0: N/A
Comments with no observed contribution / relevancy to the discussion</td><td>MC-0</td><td>MC-0</td><td>MC-0</td></tr>
<tr><td>MC-1: Question
Gather unknown information, inquire, start a discussion, reflect on problems raised</td><td>MC-A: Question</td><td>MC-A</td><td>MC-A</td></tr>
<tr><td>MC-2: Reply
Respond to or elaborate on other’s questions or statements</td><td rowspan="6">MC-X: Response
Provide feedback or address previous questions, statements or perspectives</td><td rowspan="2">MC-B: Simple Replies
Make simple responses to other’s questions or statements, establish rapport, share feelings, agree with other’s ideas either directly or indirectly</td><td rowspan="2">MC-B</td></tr>
<tr><td>MC-10: Support
Establish rapport, share feelings, agree with other’s ideas either directly or indirectly, provide feedback to other’s comments</td></tr>
<tr><td>MC-3: Clarification
Identify and elaborate on ideas and thoughts</td><td rowspan="4">MC-F: Argument
Make statements or responses that are argumentative in nature or contribute to the argument for or against a thought or idea</td><td rowspan="4">MC-F</td></tr>
<tr><td>MC-4: Interpretation
Analyze, make predictions, build hypotheses</td></tr>
<tr><td>MC-6: Assertion
Provide explanations and arguments to maintain and defend ideas questioned by other</td></tr>
<tr><td>MC-8: Judgment
Make decisions, appreciations, evaluation and criticisms of ideas, facts and solutions discussed; evaluate text orientation and authors’ positions</td></tr>
<tr><td>MC-5: Conflict
Debate other’s viewpoints, show disagreement, show friction with others</td><td>MC-C: Conflict</td><td>MC-C</td><td>MC-C</td></tr>
<tr><td>MC-7: Consensus building
Try to attain a common understanding of the issues in debate</td><td rowspan="2">MC-D: Result
Conclude a discussion, reflect upon the learning process of the project</td><td rowspan="2">MC-D</td><td rowspan="2">MC-D</td></tr>
<tr><td>MC-9: Reflection
Acknowledge learning something new, judge importance of discussion topic in relation to learning</td></tr>
</table>

	MC-11: Others Social statements, messages with functions unable to categorize	MC-E: Others	MC-E	MC-E
Thinking Development (TD) Objective: To measure the quality of student learning as reflected in their cognitive development				
The Revised Bloom's Taxonomy in the cognitive domain (Krathwohl, 2002)	TD-0: N/A No cognitive development observed	TD-0: N/A	TD-0	TD-0
	TD-1: Remember Retrieving relevant knowledge from long-term memory	TD-L: Low cognitive development • Retrieving relevant knowledge from long-term memory • Determining the meaning of instructional messages, including oral, written, and graphic communication • Carrying out or use a procedure in a given situation	TD-L	TD-L
	TD-2: Understand Determining the meaning of instructional messages, including oral, written, and graphic communication			
	TD-3: Apply Carrying out or use a procedure in a given situation			
	TD-4: Analyze Breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose	TD-H: High cognitive development • Breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose • Making judgments based on criteria and standards • Putting elements together to form a novel, coherent whole or make an original product	TD-H	TD-H
	TD-5: Evaluate Making judgments based on criteria and standards			
	TD-6: Create Putting elements together to form a novel, coherent whole or make an original product			

Notes:

(1) The categories "Give Orientation" and "Ask for Orientation" in Bales' categorization (1950) were rephrased into "Give Information" and "Ask for Information" to highlight the act of information exchange seemingly central to a constructivist learning process.

(2) The scope of the category 'Show Tension' in Bales' categorization (1950) were narrowed down to 'Ask for Help' to highlight the act of seeking help that is commonly seen in a learning process.

(3) These two categories are adapted from Tirado, Aguaded, & Hernando (2011) to capture acts of team building and socialization that are deemed essential to facilitate communication and collaboration.

* The category definitions are adapted from the designated sources with slight modifications from the present authors. Examples of categories are presented in this linked document: http://ccmir.cite.hku.hk/data/FW_examples.pdf