Automatic Coding of Questioning Patterns in Knowledge Building Discourse

Jin Mu, Jan van Aalst, Carol Chan, Ella Fu
The University of Hong Kong, Pokfulam Road, Hong Kong
jinmu@hku.hk, vanaalst@hku.hk, ckkchan@hku.hk, ellafu@ied.edu.hk

Abstract: We propose a novel method for identifying questioning patterns, which are assumed to be one of the essential factors indicating the quality of knowledge-building discourse. The underlying principle of the proposed method is to extract syntactic and sematic information before segmenting the raw data and annotating them according to a multi-layer framework called ACODEA. As a bottom layer of the framework, the "pre-coding" phase makes it possible to translate the raw data into machine-readable and context-independent language, and to make Natural Language Processing tools aware of users' preferences and underpinning mechanisms of identifying the desired pattern. Explorative but promising evidence is reported toward a more comprehensive perspective by combining qualitative and quantitative methods to analyze the discourse data. Given those findings, we argue in favor of mixed methods of content analysis and they further generated directions for future methodological development and empirical applications.

Introduction

In computer-supported collaborative learning (CSCL) environments learners often communicate with each other via text-based, digital discussion boards (Rosé et al., 2008), and this has been argued to reflect socio-cognitive processes of knowledge construction (Vygotsky, 1986). During collaborative learning activities, individual learners interact with each other in a dynamic way, making it very difficult to measure and assess learning effects independently. This may be one reason why the focus of collaborative learning research has shifted from studying learning outcomes and products to studying learning processes (Dillenbourg, Baker, O'Malley, & Blaye, 1995). With an interest in the collaborative learning process, the focus has recently shifted again – this time from analyzing individual learning processes toward identifying collaborative patterns that positively influence learning. This shift is fundamentally grounded in our understanding of collaborative learning from socio-constructivist perspectives.

Although uncovering findings related to how collaborative knowledge creation is accomplished is useful, analyzing a huge body of discourse data manually is an arduous task that consumes much time and slows down the research progress substantially. Over the past decade, there has been a substantial effort to develop innovative technologies that enable automatic content analysis in the domain of CSCL. These techniques enhance the ability of traditional approached to extract patterns that are assumed to be essential in the cognitive and social processes of learning. Against this background, by using a Natural Language Processing (NLP) tool called TagHelper (Dönmez, Rosé, Stegmann, Weinberger, & Fischer, 2005) and its successor SIDE (Mayfield & Rosé, 2010), a multi-layer framework called ACODEA (Automatic Classification of Online Discussions with Extracted Attributes, Mu, Stegmann, Mayfield, Rosé, & Fischer, 2012) has been shown to be optimized for fully automatic segmentation and context-independent classification of the desired patterns—e.g., the quality of argumentation in a text-based CSCL discourse data. By extracting syntactic and semantic features during a pre-processing phase before content analysis, the framework allows a bottom-up specification of the in-depth information contained within the discourse corpus and it is therefore more precise and reliable than traditional approaches. The goal of the present study is to extend the previous work on automatic content analysis by applying the ACODEA framework to data from Knowledge Forum. The on-going efforts herein are assumed to extend the capabilities of the classification models with the outlined steps to be quickly customized for different contexts and alterative coding dimensions of interest in the field of CSCL.

Knowledge Building and Automatic Content Analysis

Knowledge building refers to the development of innovative and sustained knowledge within a community (Scardamalia & Bereiter, 1994, 2006). The major objective of this pedagogical approach is to initiate students into a knowledge-creating civilization by encouraging them to engage in sustained idea improvement and advance the knowledge collectively as a community (Scardamalia & Bereiter, 2006; Zhang, Scardamalia, Reeve, & Messina, 2009). Consequently, it turn out to be essential to conduct content analysis which is capable of revealing what is developed through the continuous process of idea improvement and knowledge advancement at both individual and collective levels.

In the last decade, while various assessment approaches have been developed so intensively that some of the tools have been even integrated with Knowledge Forum - a technology-mediated learning environment to foster knowledge building - it has proven challenging to grasp the overall picture of the community-based learning process. In fact, the majority of current automatic approaches are still in early stages of development; previous research has mostly focused on detection of simple patterns rather than in-depth content analysis of discourse data. For instance, the Analytic Toolkit (Burtis, 2002) provides summary statistics on student participation and interaction in Knowledge Forum databases, by counting the instances of note creation, note reading, and note linking. Similarly, applet tools (Zhang, Hong, Scardamalia, Teo, & Morley, 2011) for social network analysis (SNA) have been used to explore the social structure of collaborative discourse by offering quantitative indices, such as network centrality in networks based on reading behaviors (i.e., who has read whose notes). However, little attention has been given to the quality of knowledge advancement and reflection on the depth of cognitive and social processes taking place during the collaborative learning. Recently, van Aalst et al. (2012) took one step forward to explicitly analyse the quality of knowledge-building discourse by developing a tool for formative assessment – the Knowledge Connections Analyzer (KCA). The KCA was designed to create a model for the collaborative and epistemic patterns of collaborative knowledge construction by retrieving evidence on four key questions: 1) Are we a community that collaborates? 2) Are we putting our knowledge together? 3) How do our ideas develop over time? And finally 4) What is happening to my own ideas? Using this model, van Aalst et. al. (2012) began to illustrate the collective (Q3) and individual (Q4) aspects of idea improvement by extracting key words which were used most frequently to trace the awareness and use of new concepts appeared in the database.

Adapted from previous efforts within knowledge-building communities to conduct qualitative content analysis either manually or automatically (Carol Chan & Lam, 2010; van Aalst, 2009; Zhang et al., 2011), in the current study we intended to go beyond the existing approaches to further identify critical features of knowledge-building discourse by using advanced NLP technologies. Briefly, the NLP tool SIDE can automatically extract features like line length, unigrams, bigrams and part-of-speech bigrams from the annotated data to build models (Mayfield & Rosé, 2011). The process is similar to linear regression that expresses the classification categories as a linear combination of the attributes (extracted syntactic or semantic features) with predetermined weights (coefficients). We assume that the appropriate value of the predicted weights is dependent on the importance of the extracted features to reflect on the underlying epistemic and collective aspects of knowledge-building discourse instead of the simple accounting frequency.

Questioning in Knowledge Building Discourse

Questioning is a core function and a key feature of both learning and teaching, and good questions can stimulate students to think at higher cognitive levels (Dillon, 1988). Furthermore, the questioning behavior in learning has consistently elicited elaborated explanations, inferences, justifications, speculations, and other essential signs of complex knowledge construction (King, 1994). While asking and answering questions are among the most common human activities, it is remarkable how little is known systematically about questioning , especially about the methods for measuring and analyzing the desired questioning patterns in CSCL.

It has been reported that over 75% of the questions posed in both elementary and secondary classrooms are "recalling" questions (Dillon, 1988). Approximately 3.5% of the questions are asked to check for understanding of procedures, routines, and only slightly more than 1% questions are at a higher cognitive level, such as evaluation and synthesis questions (Craig & Cairo III, 2005). In addition, learners are rarely observed to ask self-generated questions of the teacher or other peer pupils. Hence, the majority of studies in researching the effectiveness of questioning focus on teacher-generated questions and examine the relationship between such questioning behaviour and student achievement (Craig & Cairo III, 2005). However, when learners engage in knowledge-building discourse, in which learners play more central roles (Scardamalia & Bereiter, 1994), it would be useful to know whether students can generate higher-order questions, which lead students to think, analyse and synthesize the discussion topic at higher cognitive levels.

Craig and Cairo III (2005) identified six types of questions: Recall (facts from memory); Check for understanding of procedures and routines; Use (using knowledge to comprehend, apply, or analyse); Teacher repeats the question two or more times; Create (synthesizing to arrive at a conclusion) and Teacher asks multiple questions. According to King (1994), we need to differentiate "memory" questions which refer to those requiring learners to simply remember and repeat what they had heard and memorized from the lesson and "thinking" questions. The latter ones require learners to not only remember information from the lesson but also think about that information. Thinking questions were further classified into comprehension questions and connection questions. King (1994)) stated that comprehension questions "check how well you understand the lesson" and "ask you for a definition in your own words or ask you to tell about something you learned about- but in your own words, not the teacher's words" (p.346). Connection questions are thought provoking because they require students to go beyond what was explicitly stated in the lesson by linking two or more ideas together in some way. As a result, during a discussion the learners tended to make those connections between and among ideas, which may reflect the mental representations they constructed the links in mind. Such highly elaborated and richly integrated questions could account for the improved comprehension of the instructional material.

Learners have been regarded as being capable to ask and recognize two types of questions, namely text-based question promoted by text and higher-order knowledge-based questions stimulated by event (Carol Chan, Burtis, Scardamalia, & Bereiter, 1992; Scardamalia & Bereiter, 1992). In line with the previous research on knowledge building, questioning patterns have also been classified using two categories determined by the cognitive goals: "fact-seeking" and "explanation-seeking" questions. Explanation-seeking questions are embedded in the process of inquiry by asking "why" and "how", whereas fact-seeking questions are looking for "fragmented pieces of knowledge" (Hakkarainen, 2003, p. 1075). In another study (Lee, Chan, & van Aalst, 2006), further differentiated questions based on the nature of the information sought: 1) definitions and simple clarifications; 2) factual, topical and general information; 3) specific gaps in terms of open-ended responses and different viewpoints; and 4) explanation-based questions that focus on problems instead of topics and identify sources of inconsistencies; generates conjectures and possible explanations.

Three functions of question can, therefore, be identified in the present study. Simple statements of information or facts gleaned directly from the lesson, prior knowledge, or experience are coded as fact-seeking questions (e.g., "What is meant by zone of proximal development?"). Thinking questions (e.g., "What is the role of assessment in a learning community?") ask for deeper understanding of by translating into a student's own words and they are often elaborated upon by connecting with other conceptual ideas. Using questions that integrate aspects of the contextual information outside the learning environment assume to go beyond other question functions in some manner. An Example is "How can we make use of Knowledge Forum and really help the students to construct knowledge through reading?"). This kind of questions assume to effectively prompt students to connect learning content with their prior knowledge and personal experience with the purpose of resolving the authentic problems raised from real contexts. As mentioned above, a major concern of CSCL research focuses on in-depth analysis of collaborative learning processes. In the following, we will present an advanced approach to automate the content analysis of questioning behavior in knowledge-building discourse. The main question addressed in this study is: How does the ACODEA framework perform in automatically analyzing knowledge-building discourse data? We divided this question into three sub-questions: (a) to examine the reliability of capturing key patterns of questioning behaviours in an automatic way, (b) to explore the function and degree of questioning patterns in a knowledge-building community by applying the developed approach of automatic content analysis, and (c) to determine the effects of this automatic content analysis by comparing with other automatic approaches integrated in Knowledge Forum.

Research Questions

RQ1: *Can the automatic content analysis be implemented reliably to extract the key patterns of questioning behavior in knowledge-building discourse?* We expected to achieve an acceptable level of agreement between automatically generated codes by SIDE and human codes when we automate the text classification on the multi-layer ACODEA framework by extracting the desired attributes of questioning behavior in a systematic way.

RQ2: *Which function and degree of questioning behaviour would be more often exhibited in the knowledge-building discourse?* To answer this explorative question we needed to describe the frequency, type and quality of the questioning behaviors coded through the automatic approach investigated in RQ1.

RQ3: *To what extent the results of the automatic content analysis are related to those results reported by other automatic measurement approaches, such as the Analytic Toolkit (ATK)?* We hypothesized that notes embedded with higher-order questions are expected to be more widely read and built-on during the discussion. By examining the features of such notes, we hoped to gain some insight into why some of notes have more impact than others to be read or receive more build-on notes. In other words, we mainly concern how much variance in the number of reading and building-on can be explained by the extracted patterns of questions.

Methods

Participants and Learning Task

The participants consisted of more than 40 teachers, researchers, and graduate students who were part of the BCHK Network in 2002 (CKK Chan & Van Aalst, 2003). The Knowledge Forum database for this networks hosted a course on knowledge building, but also contained online discourse of teachers who were attempting to implement other higher-order thinking strategies in classrooms, in line with a recent curriculum reform in Hong Kong that emphasized "learning how to learn" (CDC, 2001). Participants were required to contribute to online discussion on Knowledge Forum, which mainly focused on a set of independent but closely connected topics to acquire deeper understanding of knowledge-building and related theories, classroom implementation, the role of teacher, and instructional designs.

Data Source and Coding Processes

Altogether, there were 1742 notes and 65,535 words in the corpus collected from 5 Knowledge Forum views. Two human coders analyzed almost all of the raw data. About half of the human-coded data were used as the training materials on which a few automatic models can be built by SIDE (Mayfield & Rosé, 2010). The resulting model could then be easily applied to classify un-annotated data, and then the assigned codes could be further reviewed on the annotation interface that facilitates the process of humans correcting errors made by the automatic coding. The remaining manually coded dataset were further used for testing the training models. SIDE employs a consistent evaluation methodology referred to as 10-fold cross-validation, where the data for training the models can be randomly distributed into 10 piles. Nine piles are combined to train a model. One pile is used to test the model. This is done 10 times so that each segment is used as a test set once. And then the performance values are averaged to obtain to final performance value (Rosé et al., 2008).

By following the Automatic Classification of Online Discussions with Extracted Attributes (ACODEA, Mu et al., 2012), the coding process implemented in the present study consists of three layers. The general idea underlying the multi-layer framework of automatic content analysis is to extract features at the lower layer that assume to contribute to the text classification at the upper layer. For instance, a unit of analysis can be identified as fact seeking question (at the upper layer) by combining both contextual facts and question words (at the lower layer). (i) Regarding the semantic attributes extracted at the lower layer, each single word in the text was separated into one of the following categories: (a) Core Concept, keywords from knowledge-building theory and principles; (b) Peripheral Concept, keywords form relevant theories and learning sciences; and (c) Contextual Information from the learning environment and local settings. In addition, there were other attributes being of importance in reflecting the (d) Question Words and (e) Thinking Verbs as the key indicators of higher-order questions that were distinct from other (f) General Verbs. Examples of the extracted attributes are illustrated in the Table 1. (ii) The unit of analysis was defined as a sentence or part of a compound sentence that can be regarded as "syntactically meaningful in structure" (Strijbos, Martens, Prins, & Jochems, 2006). For instance, according to these rules of segmentation, punctuation and the special symbol like question mark are boundaries that can be used to segment compound sentences if the parts before and after the boundary are 'syntactically meaningful'. The segments can be further identified either as statement or question. (iii) The last coding layer was designed to capture the patterns of questioning behavior in knowledge-building discourse. In the present study we are mainly concerned with the categories (summarized in Table 2 below) by using a machine-readable coding system that captures the function of the questions ranging from low to high in complexity, and roughly corresponding to the two degrees for each function of questions.

Table 1 Extracted attributes and examples

Extracted attributes	Example
Core Concept	Knowledge Building, Principle, Collective, Expert, KF
Peripheral Concept	Ability, Gifted, Pedagogy, Notes, Views
Contextual Information	Hong Kong, Primary School, Mathematics, Physics
Question Words	Why, How, Where, When, Who, What, What if
Thinking Verb	Think, Wonder, Reflect, Test, Hypothesize
General Verb	Do, Have, Be, Can

Table 2: Coding schema to analyze the questioning patterns within knowledge-building discourse

Function of Questions		Degree of Question	
Fact-seeking Questions	Requiring learners to simply recall and repeat explicit and factual information	Low-level Yes or No Trivial Question	by asking for yes-or-no responses in terms of trial facts with brief wording
		High-level Open-ended Question	by asking for open-ended responses
Thinking Questions	Requiring learners to think about conceptual ideas for deeper understanding	Low-level Elaboration Question	by clarifying, elaborating and explain a conceptual idea
		High-level Connection Question	by linking two or more conceptual ideas together
Using Questions	Requiring learners to apply conceptual ideas to certain real context with the purpose of practice	Low-level Utilization Question	by applying peripheral concept to context
		High-level Application Question	by applying core concept ideas to context

Measuring Reliabilities of Automatic Content Analysis and other Variables

The reliability of the coding was measured using Cohen's Kappa value and percent agreement. Both of the indexes have been regarded as widely used standards for measuring coding reliability. Percent agreement is the most simple and most popular reliability coefficient (De Wever, Schellens, Valcke, & Van Keer, 2006).

Statistically, the inter-rater agreement is determined by dividing the number of codes that are agreed upon by the total number (agree and disagree all inclusive) of codes. Supplemental criterion for success is reaching a level of inter-rater reliability with a gold standard as measured by Cohen's Kappa that is .7 or higher (Strijbos et al., 2006). Here it is worthwhile to further clarify that the present study was undertaken to evaluate different types of Kappa including (1) inter-rater agreement between human coders Kappa (Human-Human) to evidence the initial reliability of training examples; (2) inter-rater agreement generated by the 10-fold cross-validation to certify the internal reliability of the SIDE training models. The 10 results from comparing the coding between SIDE and manually coded training materials then can be averaged to produce a single estimation Kappa (SIDE-Training); and finally (3) the conclusive Kappa (SIDE-Testing) between SIDE and human coders calculated with the additional testing materials.

With respect to other variables measured in the present study, both of the categorical variables *Function of Questions* (Fact-seeking, Thinking vs. Using) and *Degree of Questions* (Low vs. High) were coded by applying the approach of automatic content analysis developed for the present study. Another analytic tool Analytic Toolkit (ATK) that is integrated within Knowledge Forum provided information for reflecting on the *Number of Reading and Building-on* which refer to how many times the notes were read or replied by other members within the Knowledge Building community during the online discussion.

Results

Two different analyses were conducted in the present study. First, reliability of the various coding categories in the multi-layer framework was calculated and table displayed. Second, linear regression analyses were conducted to assess the degrees of association between automatically coded questioning behaviours and the number of reading and building on the notes as assessed by ATK.

RQ1: Can the automatic content analysis be implemented reliably to extract the key patterns of questioning behavior in knowledge-building discourse? Two coders created the training material for SIDE. The overall value of kappa on segmenting and identifying questions was statistically highly significant; Cohen´s Kappa (Human-Human) was 1.00 with 100 percent agreement that indicated a good degree of inter-rater reliability beyond chance. Additionally the human coders achieved a high value of Cohen's Kappa (Human-Human) = .89 (Percent Agreement = 91.7%) for the final coding layer. These results indicate acceptable human baseline performances for SIDE to be trained to analyze the un-annotated data regarding the extracted attributes, segmentation and coding layers.

SIDE achieved an internal Cohen's Kappa (SIDE-Training) = .73 (Percent Agreement = 96.7%) on the layer of segmentation. The reliability comparing SIDE with a human coder (based on raw text) was sufficiently high (Cohen's Kappa (SIDE-Testing) = .71; Percent Agreement = 89.0%). As shown in Table 3, sufficient inter-rater agreement values were also achieved for the second layer with Cohen's Kappa (SIDE-Training) = .94 (Percent Agreement = 99.1%) and Cohen's Kappa (SIDE-Testing) = .96 (Percent Agreement = 99.4%). Internal Cohen's Kappa (SIDE-Training) = .73 (Percent Agreement = 82.9%) was achieved by SIDE when it attempted to automatically code the questions with respect to the function and the degree. A human coder and SIDE achieved an agreement of Cohen's Kappa (SIDE-Testing) = .77 (Percent Agreement = 85.5%).

Table 3: Reliability of the multiple layers of automatic content analysis

Multiple Layers of Automatic Content Analysis		Cohen's Kappa	Percent Agreement
Layer i	Segmenting		
	Training (SIDE)	0.73	95.7%
	Testing (Human vs. SIDE)	0.71	89.0%
Layer ii	Identifying Questions		
	Training (SIDE)	0.94	99.1%
	Testing (Human vs. SIDE)	0.96	99.0%
Layer iii	Coding Questions		
	Training (SIDE)	0.73	82.9%
	Testing (Human vs. SIDE)	0.77	85.5%

RQ2: Which function and degree of questioning behavior would be more often exhibited in the knowledge-building discourse? Upon initial impression, there were 3465 single segments in total, and 263 of them were identified as questions. The results indicate that the community members did generate a number of questions spontaneously. Among them, slightly less than half (44.7%) of questions generated in the Knowledge Building discourse were thinking-oriented, only 28.2% of questions were seeking for factual information, and a rather low percentage of 15.3% linked to the using questions. The frequency percentage of various degrees of questions did not appear to be significantly different cross three functions. But participant appeared to be able to generate higher-order questions. For instance, in the questions asking for factual knowledge, roughly two third

of them were open-ended. The most frequently asked question was connection question at 25.2%, followed by elaboration questions at 19.5%. Perhaps not surprisingly, the participants tended to be more often to apply the knowledge-building theory at the lower level, given that 15.3% of the generated questions were classified as the utilization questions, followed by the higher degree of application questions (11.8%). The main patterns of the questioning behaviour in the KB discourse are summarised in Table 4.

Table 4: Frequency of various functions and degrees of questions

Questions Categories		Frequency	Percentage
Fact-seeking Questions		74	28.2%
Low Degree	Yes or No Question	*26*	*10.0%*
High Degree	Open-ended Question	*48*	*18.3%*
Thinking Questions		117	44.7%
Low Degree	Comprehension Question	*51*	*19.5%*
High Degree	Connection Question	*66*	*25.2%*
Using Questions		71	27.1%
Low Degree	Utilization Question	*40*	*15.3%*
High Degree	Application Question	*31*	*11.8%*
Total		262	100%

RQ3: To what extent the results of the automatic content analysis are related to those results reported by other automatic measurement approaches, such as the Analytic Toolkit (ATK)? A multiple regression analysis was performed between the dependent variables (separately, the frequency of Build-on and Reading Notes) and the independent variables (simultaneously, the Function of Questions in terms of two dummy coding variables Thinking and Using, the Degree of Questions Low vs. High, and the Authority of the Authors Researchers vs. Teachers). Analysis was performed using SPSS Linear Regress.

Table 5: The number of reading and building-on by other community members

Function of Q		Fact-seeking Questions		Thinking Questions		Using Questions		Total
Degree of Q		Low Degree	High Degree	Low Degree	High Degree	Low Degree	High Degree	
Reading	*Mean*	28.58	24.08	25.82	44.27	14.88	28.13	29.03
	SD	20.92	15.82	16.23	39.09	8.61	22.33	26.13
Building On	*Mean*	1.00	1.06	1.00	1.23	0.45	0.61	0.94
	SD	1.06	2.04	1.11	1.32	1.11	0.99	1.38
	N	26	48	51	66	40	31	262

Regression analysis revealed that the model significantly predicted the number of reading and building-on by other community members. The model using the 4 predictors explained about 36.5% of the variance of the number of reading by others, F (4,257) = 36.89, p < .001. The predictor Thinking had a significant positive effect on the number of reading, β = .14, p < .05. Notes containing higher order thinking question were more often read than notes with fact-seeking questions, while another dummy coding factor Using was not a significant predictor of the number of reading during the online knowledge building discussion. β = -.04, p > .05. The Degree of Questions had a significant regression coefficient, β = .16, p < .01, indicating that notes with higher degree of questions were expected to be more read. Meanwhile, the Authority of the note author was of similar magnitude to predict the number of reading, β = .52, p < .01. Not surprisingly, researchers still provided focus to the discussion by posting more impactful notes read by other participants.

The R square value of the Build-on Model was lower, which was able to account for 19.7% of the variance in the model, F(4,257) = 15.78, p < .001. Different form the model of reading, the factor of Using had a significantly negative effect on the number of building-on, β = -.15, p < .05. Notes with the questions asked for utilization and application on authentic problems were expected to be surprisingly less desired regarding the number of building-on than notes with fact-seeking questions. Different from reading a note, building-on other's note is a more active and challenging task, especially when it is required for applying the core conceptual ideas to real context. Hence, teachers engaged in the knowledge building discourse shunned to respond to more difficult questions. Other variables did not contribute to the model.

Conclusions and Future Work

This study found promising evidence that questioning patterns in Knowledge Forum databases can be coded automatically using the ACODEA framework acceptable reliability. Moreover, although previous research only evaluated the reliability for argumentation data (Mu et al., 2012), this study suggests that the method can also be applied to data created within a different theoretical framework—knowledge building. It suggests that the framework can be applied successfully for automatic content analysis on different construct of interest and crossing different domains. The particular strength of the method lies in the clear understanding of how the discourse properties of interest manifest themselves via a variety of linguistic terms (either syntactic or semantic), which can be further viewed as a natural extension of keywords targeted at the machine-readable and context-independent language to build text classification algorithms that are consequently more powerful than text classification directly based on raw data.

We also presented two methods for assessing a real-world data set of knowledge-building discourse. While qualitative content-based analysis appeared to be more effective to detect and analyze the desired discourse patterns than quantitative analysis of counting the reading and building-on behaviors in a superficial manner, when used in isolation the methods may not identify all of the aspects of the Knowledge Building discourse. For this reason, we combined both of the qualitative and quantitative methods to provide a full picture of what happened during the online discussion. Our classifiers can reliably identify multiple patterns of questioning behavior, which have been further shown to be able to explain and predict if notes can be more read or built-on as assessed by the analytic tool (ATK). In this way, integrating results from different evaluations into a global consideration brought new insight for us to analyze the discourse data comprehensively and deeply.

We now briefly discuss some avenues for future work. As one of the major contribution of the present study, automatic analysis not only intends to speed up research projects, it also brings insights into essentially changing the way how teachers and educators design learning environments and scaffold the desired collaborative learning. Specifically, automatic analysis of online discussion can provide instructors with the capability to monitor the real-time learning progress occurring in large classes, indicate what the specific and personalized need should be addressed and consequently enable the adaptive intervention, which is assumed to be more efficient in promoting productive collaboration and knowledge building, than the static, one-size-fits-all scaffolds (Gweon, Rosé, Carey, & Zaiss, 2006; Kumar, Rosé, Wang, Joshi, & Robinson, 2007; Stegmann, Mu, Gehlen-Baum, & Fischer, 2011). Practically, integrating the automatic assessment in Knowledge Forum can be of valuable assistance for teachers to get to know how well their students are learning with a much lower investment of efforts. Therefore teachers can scaffold individuals and groups of learners more effectively in formative assessment. Based on the current study, the implementation of a well-controlled, randomized experiment is needed to examine the efficacy of the automatic content analysis as an effective formative assessment technique.

In addition the newly developed approach seems to be promising to develop domain insensitive coding schemas to model similar behavioral patterns occurs knowledge-building discussion. In other words, it enables researchers to address the urgent need for the re-use of coding schemas in diverse contexts. While being a well-established tradition to reanalyze quantitative data in social sciences, conducting secondary analysis of qualitative resource collected by other researchers, e.g. text-based discourse is relatively scarce in the field of CSCL. Hence, the general goal of the preliminary investigation aims at developing a feasible model in a manner that allows content analysis focusing on context-independent perspectives. At the same time, we will try to promote the discussion among the researchers within knowledge-building communities to facilitate the secondary analysis cross various learning settings.

References

Burtis, J. (2002). *Analytic toolkit for Knowledge Forum*. Paper presented at the Institute for Knowledge Innovation and Technology, Toronto, Ontario, Canada.

CDC. (2001). Learning to learn - the way forward in curriculum. Hong Kong, SAR, China: Government Printer.

Chan, C., Burtis, P. J., Scardamalia, M., & Bereiter, C. (1992). Constructive activity in learning from text. *American Educational Research Journal, 29*(1), 97-118.

Chan, C., & Lam, C. K. (2010). *Conceptual change and epistemic growth through reflective assessment in computer-supported knowledge building*. Paper presented at the International Conference of Learning Sciences, Chicago.

Chan, C., & Van Aalst, J. (2003). Assessing and scaffolding knowledge building: Pedagogical knowledge building principles and electronic portfolios. In U. Hoppe, B. Wasson & S. Ludvigsen (Eds.), *Support for Collaborative Learning 2003 -Designing for change in networked learning environments*: Kluwer Academic Publishers.

Craig, J., & Cairo III, L. (2005). Assessing the relationship between questioning and understanding to improve learning and thinking (QUILT) and student achievement in mathematics: A pilot study. *Appalachia Educational Laboratory at Edvantia*.

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education, 46*(1), 6-28. doi: 10.1016/j.compedu.2005.04.005

Dillon, J. T. (1988). *Questioning and teaching. A manual of practice*: ERIC.
Dönmez, P., Rosé, C., Stegmann, K., Weinberger, A., & Fischer, F. (2005). *Supporting CSCL with automatic corpus analysis technology*. Paper presented at the Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!, Taipei, Taiwan.
Gweon, G., Rosé, C., Carey, R., & Zaiss, Z. (2006). *Providing support for adaptive scripting in an on-line collaborative learning environment*. Paper presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems, Montreal, Quebec, Canada.
Hakkarainen, K. (2003). Progressive inquiry in a computer‐supported biology class. *JOURNAL OF RESEARCH IN SCIENCE TEACHING, 40*(10), 1072-1088.
King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368.
Kumar, R., Rosé, C., Wang, Y.-C., Joshi, M., & Robinson, A. (2007). *Tutorial Dialogue as Adaptive Collaborative Learning Support*. Paper presented at the Proceeding of the 2007 conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work, Los Angeles, California, USA.
Lee, E. Y. C., Chan, C. K. K., & van Aalst, J. (2006). Students assessing their own collaborative knowledge building. *International Journal of Computer-Supported Collaborative Learning, 1*(278-307).
Mayfield, E., & Rosé, C. (2010). *An interactive tool for supporting error analysis for text mining*. Paper presented at the Proceedings of the NAACL HLT 2010 Demonstration Session, Los Angeles, California.
Mayfield, E., & Rosé, C. (2011). *Recognizing authority in dialogue with an integer linear programming constrained model*. Paper presented at the Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, Portland, Oregon.
Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *International Journal of Computer-Supported Collaborative Learning, 7*(2), 285-305. doi: 10.1007/s11412-012-9147-y
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. doi: 10.1007/s11412-007-9034-0
Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge based questioning by children. *COGNITION AND INSTRUCTION, 9*(3), 177-199.
Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences, 3*(3), 265-283. doi: 10.1207/s15327809jls0303_3
Scardamalia, M., & Bereiter, C. (2006). Knowledge building. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 97-115). Cambridge, UK: University Press Cambridge.
Stegmann, K., Mu, J., Gehlen-Baum, V., & Fischer, F. (2011). The myth of over-scripting: Can novices be supported too much? . In H. Spada, G. Stahl, N. Miyake & N. Law (Eds.), *Connecting computer-supported collaborative learning to policy and practice: CSCL2011 Conference proceedings* (Vol. I - Long papers pp. 406-413). Hong Kong, China: International Society of the Learning Sciences.
Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education, 46*(1), 29-48.
van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 259-287. doi: 10.1007/s11412-009-9069-5
van Aalst, J., Chan, C., Tian, S. W., Teplovs, C., Chan, Y. Y., & Wan, W.-S. (2012). *The knowledge connections analyzer*. Paper presented at the The future of learning: Proceedings of the 10th international conference of the learning sciences (ICLS 2012) Sydney, Australia.
Vygotsky, L. S. (1986). *Thought and language* (2nd ed.). Cambridge, MA: MIT Press.
Zhang, J., Hong, H. Y., Scardamalia, M., Teo, C. L., & Morley, E. (2011). Sustaining knowledge building as a principle-based innovation at an elementary school. *Journal of the Learning Sciences, 20*(2), 262-307.
Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge building communities. *Journal of the Learning Sciences, 18*(1), 7-44.

Acknowledgement

The data collection for this study was supported by a grant from the Universities Grants Council of Hong Kong "Developing a Teacher Community for Classroom Innovation through Knowledge Building" (Grant HKU 740809H) and a Faculty Research Fund (Grant 201209176200).