Computers & Education 71 (2014) 279–288

Contents lists available at ScienceDirect

Computers & Education

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

Community knowledge assessment in a knowledge building environment

Huang-Yao Hong [a,*], Marlene Scardamalia [b]

[a] Department of Education, National Chengchi University, No. 64, Sec. 2, Zhinan Rd., Wenshan District, Taipei 11605, Taiwan
[b] Ontario Institute for Studies in Education of the University of Toronto, Canada

A R T I C L E I N F O

Article history:
Received 2 June 2013
Received in revised form
22 September 2013
Accepted 26 September 2013

Keywords:
Knowledge building
Computer-mediated communication
Collaborative learning
Elementary-level science
Learning communities

A B S T R A C T

This study investigates ways of using key terms to represent and assess community knowledge in an online knowledge building environment. Knowledge Forum – an environment specially designed to support advances in community knowledge – incorporates key-term analytic tools. In the current study these tools were used to determine if key-term measures complement conventional online behavioral measures in assessing community knowledge advances. Discourse rated as more reflective and depth-oriented showed higher percentages of shared key terms and higher frequency use of shared key terms than less reflective, shallower discourse. Limitations and possibilities for using key terms for automated assessment and visual representation of community knowledge are discussed.

1. Introduction

According to Stahl (2006) group cognition requires a theory of collaboration that takes the group as the unit of analysis rather than the individual. 'Learning by groups' is not the same as 'learning in groups' or individual learning through social processes. The new theme reflects the larger societal interest in knowledge creation/knowledge building. Knowledge Building is a group phenomenon, even when contributions come from identifiable individuals. It is the production of public knowledge of value to a community (Scardamalia & Bereiter, 2003). The community may be a research or design group or the world at large or it may be a group of learners in a class—in which case it is important to distinguish individual learning from the group's knowledge-building accomplishments. Neither one can be reliably inferred from the other and the interaction between the two is vital. Knowledge Building is defined by a set of knowledge-building principles which represent design challenges, ideals, and objects of continual improvement (Scardamalia & Bereiter, 2006). The principle of "community knowledge, collective responsibility" places value on contributions to shared, top-level goals of the community which are prized and rewarded over individual achievements (Scardamalia, 2004; Zhang, Scardamalia, Reeve, & Messina, 2009). Despite the importance of community knowledge and collaborative learning, the question of "How is community knowledge measured?" has remained largely unexplored (Hong, Teplovs, & Chai, 2007).

1.1. Community knowledge

To understand the nature of group or community knowledge, it is important to distinguish between a psychological concept of knowledge as something within an individual mind and a social concept of knowledge as conceptual artifacts that have a public life (Bereiter, 2002; Bereiter & Scardamalia, 1996; Hyman, 1999; Popper, 1972). Community knowledge is public knowledge—ideas made accessible to all community members through contributions to collective knowledge spaces. Thus ideas have a life beyond the individual mind and can be continually accessed and improved—much as is the case with publication in journals. Community knowledge involves a dynamic process—interactions between ideas and people knowledge (i.e., knowing people's expertise)—with participants monitoring who is working on what ideas or problems and advancing knowledge in the community (Hong & Lin-Siegler, 2012). To clarify community

* Corresponding author. Tel.: +886 2 29393091x66136.
E-mail address: hyhong@nccu.edu.tw (H.-Y. Hong).

0360-1315/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compedu.2013.09.009

knowledge, Scardamalia and Bereiter (2003) distinguish between the process through which the cultural capital of society is internalized/learned by individuals and the deliberate effort to improve community knowledge – the cultural or intellectual capital of the community.

1.2. Assessing community knowledge

Despite its time-consuming nature, a common way to assess group or community knowledge is to assess group products (Bourner, Hughes, & Bourner, 2001; Cheng & Warren, 2000; Lee, Chan, & van Aalst, 2006; Lejk & Wyvill, 2001). These group products can be a project report, a business plan, a technology tool, a design artifact, or so forth. While such outputs represent the most common way of assessing group work, there are several problems. Not all group learning results in a group product and not all group products signify contributions and understanding by all members. Especially in many online learning environments, the main learning activity tends to be online discussion. The "group product" is online activity logs and the content of individual contributions to discussions. In an effort to analyze group learning, researchers have employed peer assessment of community knowledge (Cheng & Warren, 2000; Lejk & Wyvill, 2001). For example, in a study by Lee et al. (2006) students were asked to identify important collaborative knowledge contributions to a community by collaboratively assessing one another's online notes in the computer discourse. A systematic review of empirical studies about peer assessment by Topping (1998) suggested that peer assessment was a reliable and valid assessment approach. However, like the assessment of group products, peer assessment can also be very time-consuming and reflect only one facet of community knowledge. If assessment tools could detect changes in group achievements and support users in advancing those achievements, that would be more convenient and advance our understanding of community knowledge.

One important advance in technological design for online learning environments has been automated assessment. Most online environments (e.g., Blackboard) use automated assessment measures such as average number of notes created, notes read, notes revised, words per note and so forth. These measures can be extremely useful in capturing the dynamics of online activities in a community and can be readily used on an on-demand basis by teachers and learners. But they are mainly designed to capture individual's behavioral patterns (e.g., reading or writing patterns) and/or group interaction patterns, but not the content knowledge the community is working with. To address this issue, content analysis is used to analyze notes and related content knowledge embedded in the database. This complements behavioral indicators but brings with it problems of its own.

A commonly used unit of content analysis has been theme. Themes can be either pre-specified based on an existing protocol or emergent from an open-coding process (Strauss & Corbin, 1990). But either way, there is an issue of subjectivity of theme selection. Second, content analysis has been used more often by researchers as a research tool than by users as an interaction/learning tool for assessing content in a database. Particularly for teachers (and students as well), it is a difficult and time-consuming technique to apply, thus limiting its use and usefulness for advancing the state of community knowledge. For a detailed account of content analysis, including lack of coherence in terms of theoretical base and choice for the unit of analysis, see the review by De Wever, Schellens, Valcke, and Van Keer (2006). A third issue is that it remains unclear what represents effective content analysis for assessing community knowledge. In part, this is because community knowledge is a complex construct that is not simply represented in individual members' personal contributions, but also in their interactions and collective understanding.

Use of key-term measures may address a number of issues raised above and open up easier forms of assessment with greater potential for providing helpful feedback to users. First, as a fundamental unit of content analysis, key terms (or words) represent a fairly objective unit of analysis. Second, with current technology, it is possible to easily extract key terms from a database (comparing terms in the database against those in a pre-selected dictionary or curricular materials) and to create related measures for automated assessment. Thirdly, key terms may be used to capture and visually represent content knowledge recorded in a community knowledge space and to show relationships between community knowledge and individual contributions. Key terms have been widely used for subject-indexing in books and for idea search (e.g., in academic papers or on the Internet), and for visual knowledge representation (e.g., semantic or propositional network, see Anderson, 2000; knowledge or concept map, see Novak, 1998, and tag clouds, see Hassan-Montero & Herrero-Solana, 2006). Moreover, by examining the extent to which key terms are shared between individual members, it is possible to identify more frequently used concepts and ideas as opposed to those that are not-shared by community members.

It is worth noting that measurement of community knowledge does not mean that assessment of individual knowledge is no longer important. On the contrary, in a knowledge building community individual contributions are essential for advancing group understanding. Also, beliefs underlying assessment, as well as assessment methods, can have a tremendous impact on students' learning experiences and subsequently on what and how they contribute. If assessment is based only on a group outcome or product, it is difficult, if not impossible, to determine what individual students have learned (van Aalst, 2012). Or, if assessment is only focused on a group product, without engaging all students in individual as well as collective knowledge advances, the potential of some students to contribute may be lost, with group knowledge as a whole suffering in consequence. Evaluating both individual contributions and group knowledge can help overcome problems associated with awarding the same grade to all members of a group for collaborative work, while ignoring the unique contribution made by each individual member. Teachers and students should both benefit from more useful assessments of individual and collaborative knowledge advances, especially designs that make it possible to provide feedback as work proceeds rather than simply after the fact.

1.3. The present study

The present study aimed to improve a key-term-based assessment tool in Knowledge Forum, as part of a sustained effort to advance this platform as an effective, user-friendly knowledge building environment. Knowledge building requires sustained creative work with ideas; accordingly, Knowledge Forum is designed to support high-level knowledge processes and discourse that leads to generation and improvement of ideas. This is in contrast to learning platforms that support simpler forms of discourse and accumulation and organization of conceptual and procedural knowledge rather than knowledge creation. Knowledge Forum supports members of an online community in contributing their ideas in the form of notes that are entered into "views" (online problem-solving spaces for collaborative knowledge work among community members). Knowledge Forum allows users to build on and annotate the notes of other members, as well as co-author notes, generate problems, add keywords, and create higher level rise-above notes to integrate different notes with relevant ideas or concepts. Online knowledge-building activities (e.g., note contributing, note reading, or note revising) are recorded automatically into a database, and can be computed statistically by means of a built-in tool called the Analytic Toolkit (Burtis, 2002).

Design-based research with technology advances integral to classroom knowledge practices, has characterized Knowledge Forum development (Scardamalia, 2004; Scardamalia & Bereiter, 2010). Building on this effort, the present study aims to advance existing tools by investigating key-term online measures for assessing community knowledge. The key-term tool currently in Knowledge Forum was developed by Teplovs (Teplovs, 2008; Teplovs, & Scardamalia, 2007; Teplovs, Donoahue, Scardamalia, & Philip, 2007) and uses a "Term Extraction" application provided by Yahoo (for details regarding term extraction see http://www.programmableweb.com/api/yahoo-term-extraction). The key-term tool can be used to compare keywords/key terms extracted from different sets of notes in a database, and to identify shared or overlapping terms with any benchmark (e.g., a dictionary or curriculum document) that researchers or community members select. However, research is required to determine how various indicators of change over time relate to personal knowledge growth and community knowledge advances. The behavioral measures available (e.g., number of notes posted and read) are more suitable for assessing interaction patterns than idea improvement, but idea improvement is essential for knowledge building. Therefore, the purpose of this study is to explore possibilities of using key-term measures: (1) to help assess and visually represent both personal and community knowledge, and (2) to complement (not replace) traditional online behavioral measures, in order to better capture knowledge building dynamics.

2. Method

2.1. Design and data source

Fig. 1 shows the research design of this study. First, both the conventional measures and the proposed key-term measures were employed to compare two pre-selected databases (details below). Then similarities and differences between proposed key-term measures and conventional online measures were explored to see if the proposed measures capture aspects of knowledge-building not captured with conventional online measures.

The main data source was discourse among students in a Knowledge Forum database. Knowledge building activities (e.g., reading, linking, and editing) are recorded automatically in the database, and can be summarized statistically by means of the Analytic Toolkit software. Fig. 2 shows examples of a view and a note created by the participants in this study.

There were two datasets involved in this study, both in the same Knowledge Forum database and both generated by the same students from a school in downtown Toronto, Canada. The two datasets were selected for use because a previous study (Hong, Scardamalia, Messina, & Teo, 2008) found a significant difference in terms of the nature of inquiry in these two datasets. In this previous study, a laborious content analysis, using theme as unit of analysis, was conducted to explore the different nature of inquiry between the two phases. Three themes (including self-initiated inquiry, self-directed improvement, and self-assessment) were identified through an open coding procedure (Strauss & Corbin, 1990) and it was found that inquiry tended to be more reflective and depth-oriented in phase 2 than in phase 1. Using a different key-term assessment tool, the present study tests: (1) whether the use of new key-term measures will uncover the same results found in the previous study, and (2) whether key-term measures complement conventional online measures for assessing community knowledge by providing additional information.

2.2. Context and participants

The students in this study were engaged in a course titled "integrated studies" and the main topic to be explored was decided by the class as a whole to be "human body system." The aim of this course was to gain a deeper understanding, from an inter-disciplinary perspective, of the human body as a holistic system composed of many inter-related sub-systems. The course was divided into two major phases (thus two datasets): inquiry into the internal body as a system, e.g., studying how the brain, nerves, blood and cells work together (this phase covered the first half of the semester); and the physical body as a system, e.g., studying how the head, hands, legs, knees and feet coordinate to perform certain exercises such as a long jump (this phase spanned the second half of the semester). The two phases allow students to study human body system from both a biological perspective and a physical perspective.

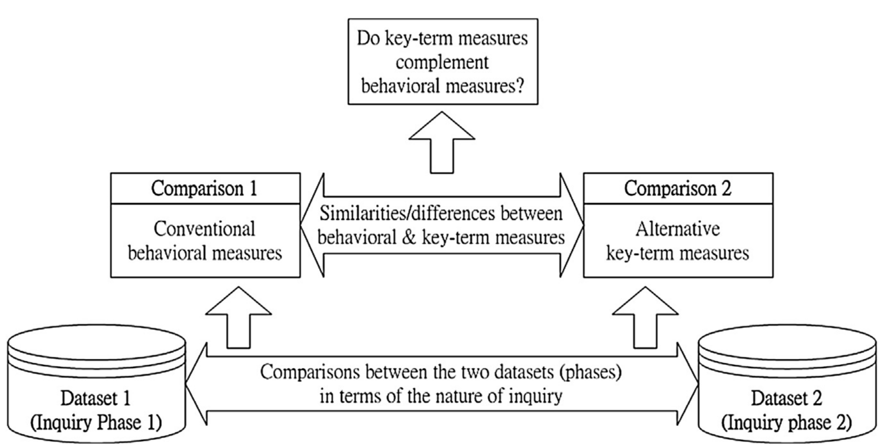

Fig. 1. Research design.

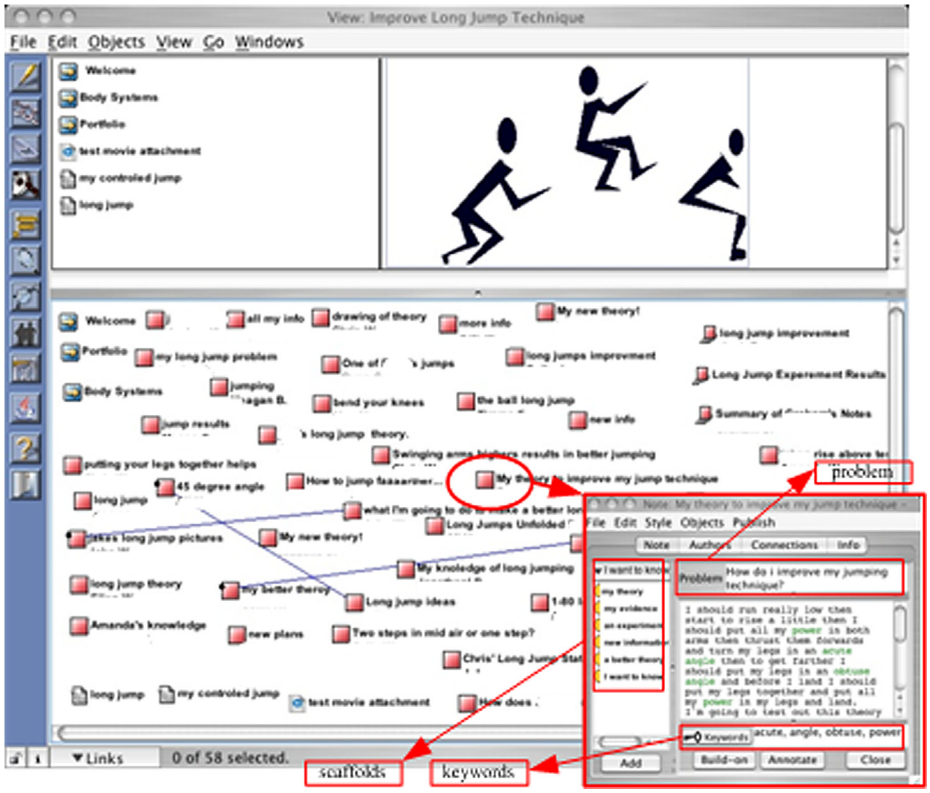

Fig. 2. Examples of a *view* and a *note* and other tool features (e.g., problem, scaffolds, and keywords). Note: A *view* (left) is a problem space created and designed by community members, to serve as a conceptual framework for ideas to be contributed in the form of *notes*. Members make use of header space, graphics palettes, and other tools to co-design the communal space for collective knowledge work. A *note* (bottom right) is where ideas are recorded, with title indicated. Within Knowledge Forum, users can identify the *problem* they want to work with, use *scaffolds* (e.g., I need to understand…), and highlight *keywords* to help convey the essence of their ideas.

The students in this class consisted of 22 fifth and sixth graders (10 girls and 12 boys). They were experienced knowledge builders as they had been engaged in knowledge building for five or six years depending on their grade level. The teacher also had several years of experience as a knowledge-building practitioner. As a part of knowledge building culture, students were accustomed to recording ideas in their community database; as they engage in face-to-face conversations they often construct plans to record their ideas in Knowledge Forum at some later point.

2.3. Method of key-term extraction

Key terms were extracted and compared in the following steps. First, all key terms within each student's notes in the database were automatically extracted by Knowledge Forum. Second, extracted key terms were compared with curriculum materials preloaded in the database—in the form of a note embedded in an independent Knowledge Forum view. This "curriculum note" was used for analysis purposes only—to identify key terms in student contributions that were also in the "curriculum note". Next, comparison was further performed between any two students' notes to identify shared/overlapping key terms. Then, the total number of shared/overlapping key terms and frequency of key-term use were calculated to indicate students' collective knowledge contribution in the community. In addition, all extracted key terms were checked by a predefined dictionary to identify misspelled words/terms. Fig. 3 provides an overview of procedures: key terms were spell-checked and different sets of notes selected and compared. It is important to note that when key terms are extracted, they are no longer viewed in their context of use. However, these key terms are still contextually linked with their original notes, so teachers can easily go back to check context to inform knowledge assessment.

Further, as a way to validate the accuracy and reliability of the tool's capacity for key-term extraction, an additional reliability check was performed. Two researchers, both with science teaching background and a deep knowledge of the database content, independently went through the whole set of key terms and checked the accuracy of all extracted key terms by removing those that were unrelated to the topics under investigation in both inquiry phases (i.e., biological concepts in phase 1 and physical concepts in phase 2). The inter-rater agreement between the two researchers was 0.95 (with differences resolved by discussion). Additionally, the total number of key terms ($n = 1614$, including repeated ones) that were automatically extracted by machine contained only thirty key terms raters identified as irrelevant (reliability rate for the tool's key-term extraction capacity = 0.98).

2.4. Proposed key-term measures

Fig. 4 shows how key terms can represent knowledge in students' knowledge spaces. In the oval to the left, overlapping that on the right, there are six key terms (cells, bones, nerves, blood, gene, DNA) These convey the full scope of student A's key-term contributions. The frequency of use of each of the six key terms (nine times) represents the intensity of key-term use by student A. Four of the six key terms are

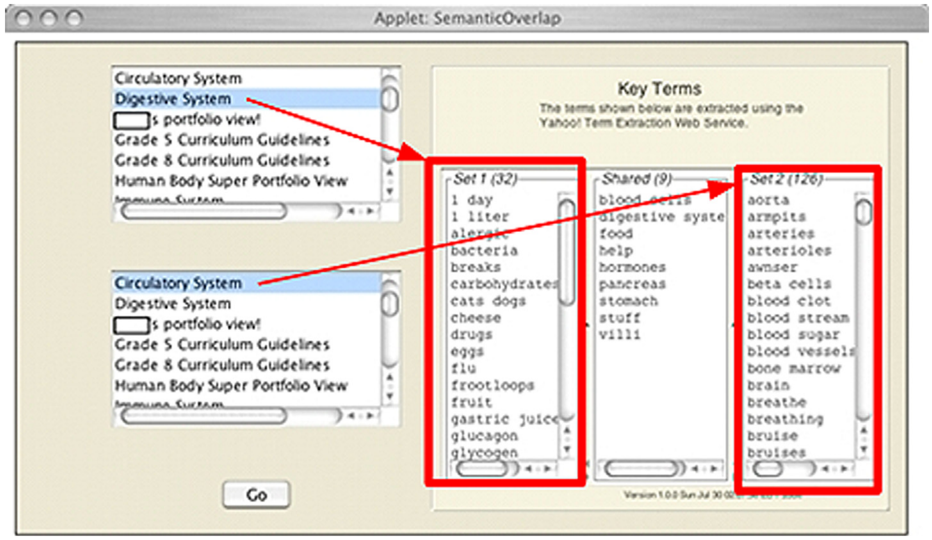

Fig. 3. Semantic Overlap Tool: Identifying key terms/shared ideas between *notes* in two different Knowledge Forum datasets. Note: The first step is to select two datasets to be compared. The selection is made through use of the view-selection windows shown to the left of the figure, top left for the first dataset, bottom left for the second. Then press go. As a first step misspelled words (words not found in the dictionary embedded in the Semantic Overlap Tool) are eliminated. Key terms then appear in the three columns within the space titled "Key Terms". The key terms extracted from all of the notes in the first dataset appear to the left, those in the second to the far right, and after automatic comparison, overlapping key-terms are shown in the middle column.

in the middle or shared area, overlapping terms used by student B. The frequency of use of these four shared key-term contributions for Students A and B is 16. Thus the intensity of key-term use in the shared space is 16. To illustrate interactions between Student A's unique and shared contributions, contributions to the shared space are represented as a percentage of the total number of key-term contributions (four out of six or 66.7%) and as a ratio of total instances (16 times over 9 times, which is 1.78). As suggested above, every student was paired with every other student to derive shared key terms for the group. Table 1 summarizes all key measures employed in this study and the statistics used to compare between the two datasets by using the conventional and new measures. The *t*-test was used for statistics computation in this study, as it is recognized to be a robust test with respect to the assumption of normality, meaning that deviations away from normality do not have a large impact on Type I error (Czado & Munk, 1998).

3. Results and discussion

3.1. Comparisons using conventional online behavioral measures

A number of traditional online measures were used to compare the two datasets for this study. A base line for subsequent work was first established; a *t*-test showed no significant difference between the two phases or datasets in terms of the total number of notes contributed ($t = 1.89$, df $= 21$, $p > .05$, with $N = 254$, $M = 11.55$, SD $= 4.39$ in phase 1, and $N = 212$, $M = 9.64$, SD $= 3.92$ in phase 2). More detailed analyses are shown in Table 2; *t*-tests indicate that there were significantly more problems being generated and worked on in the first phase. Also,

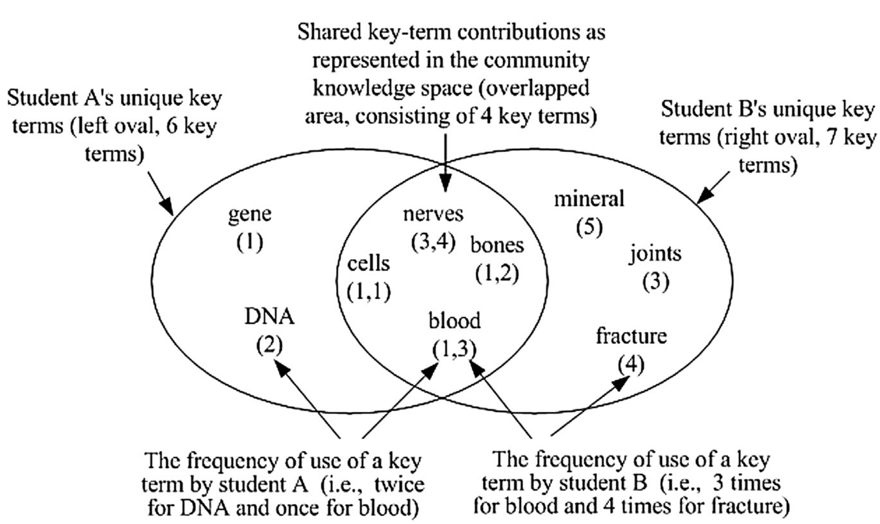

Fig. 4. An example of use of key terms for knowledge representation.

Table 1
Key measures used in this study and related statistics.

	Descriptions	Statistics
1. Conventional measures	(1) Number of notes created, (2) percentage of notes read, (3) number of notes edited, (4) percentage of notes linked, (5) number of problems worked on, (6) number of build-on notes, (7) number of note revisions, (8) number of total words, and (9) number of words per note.	*t*-tests between phases 1 and 2
2. Proposed new key-term measures	(1) Number of unique key-terms, (2) Frequency of unique key-term use, (3) Number of shared key-terms, (4) Frequency of shared key-term use, (5) Number of shared key-terms divided by number of unique key-terms, and (6) Frequency of total shared key-term use in a community divided by frequency of unique key-term use.	*t*-tests between phases 1 and 2

there were significantly more notes being linked and built-on in the first phase. But there are higher numbers of note revisions and more words (per note) being generated by students in phase 2. In other words, in the second phase there was (1) less time spent on problem generation and connectivity (i.e., note-linking and building-on), and (2) more time on elaborating the content of each note (e.g., more note revisions and more words per note). This suggests a change from more breadth-oriented inquiry in phase 1 to more reflective, depth-oriented inquiry in phase 2. This confirms findings from the previous study mentioned above (Hong et al., 2008). It is worth noting that the standard deviations seemed rather high (e.g., the number of note revisions). This may be because knowledge building highlights student autonomy, whereas traditional didactic teaching tends to maintain a relatively equal learning agenda and pace for a group of students. Therefore, it is possible that variations in performance reflect greater diversity in the activities engaged in by knowledge builders.

Of course, knowing of these behavioral changes in knowledge building dynamics is important, but these measures do not tell much about the actual change in terms of content knowledge recorded in a community database. For that we use different measures, as elaborated below.

3.2. Comparisons using alternative key-term measures

Key-terms can be constructed as a key-term cloud to help visually represent one aspect of community knowledge (see Fig. 5 for the cloud constructed for phase 1; the key terms in the figure were alphabetically ordered and visually weighted by font size). In the two phases of the study, there were in total 371 key terms contributed in the database ($n = 231$ in phase 1 and $n = 140$ in phase 2). The number of key terms was relatively low in phase 2 (as compared with phase 1); however, the frequency of their use was relatively high, indicating a shift of inquiry effort to more knowledge elaboration in phase 2. The frequency of key-term use ranges from 1 to 75 in phase 1 and ranges from 1 to 352 in phase 2. In the figure, key terms used once by one person (i.e., non-shared key terms) were excluded from visual representation as there were many. Key terms with higher frequency (bigger font size) represent terms more frequently used and shared and thus presumably more valuable concepts or ideas in a community. The idea behind such an analysis is to complement conventional online measures and capture and represent additional facets of content knowledge recorded in the community space.

Accordingly, the analysis used both the number of key terms and the frequency of key-term use (1) to differentiate shared key terms from key terms unique to an individual and then (2) to assess key-term contributions to community knowledge. First, in terms of unique key terms when comparing the mean number of key terms each student contributed between the two phases, it was found that there was no significant difference ($t = 2.065$, df $= 21$, $P > .05$). On average, each student generated 40.10 key terms (SD = 17.74) in phase 1 and 33.27 key terms (SD = 11.35) in phase 2 (see Table 3). However, when comparing the mean frequency of key-term use by each student between the two phases, there was significant difference ($t = -2.483$, df $= 21$, $P < .05$) in that the mean frequency of key-term use is 78.41 (SD = 38.74) in phase 1 and 100.14 (SD = 40.11) in phase 2.

Second, in terms of community knowledge, when comparing the mean number of key terms in the community knowledge space between the two phases, there was no significant difference ($t = -0.617$, df $= 21$, $P > .05$; with $M = 29.10$, SD $= 12.55$, in phase 1 and $M = 20.64$, SD $= 9.35$, in phase 2). However, when comparing the mean frequency of use of shared key terms in the community knowledge space between the two phases, there was a significant difference ($t = -7.686$, df $= 21$, $P < .001$), with the second phase ($M = 324.36$, SD $= 68.0$) having a higher value than the first phase ($M = 203.77$, SD $= 63.77$). Further, how each member's key-terms interact (overlap) with one another's key-terms in the community knowledge spaces was investigated. First, in terms of percentage of the total number of key-term contributions (i.e., the number of shared key-terms divided by the number of unique key-terms), the results showed a significant difference between the two phases ($t = -6.207$, df $= 21$, $P < .001$) in that the mean percentage in phase 2 ($M = 93.3\%$, SD $= 6.0\%$) is higher than in

Table 2
Changes between phase 1 and 2: conventional measures ($N = 22$).

Basic indicators	Phase 1		Phase 2		*t* Values
	Mean	SD	Mean	SD	
Number of notes created	11.55	4.39	9.64	3.92	1.89
Number of notes read	12%	7%	8%	4%	1.96
Number of notes edited	5.32	2.15	4.05	2.66	1.89
Number of problems worked on	9.41	4.35	2.00	2.16	8.35***
Percent of notes linked	25%	20%	10%	7%	3.33**
Number of build-on notes	3.00	3.13	1.00	0.76	3.24**
Number of note revisions	8.09	7.57	29.04	25.72	−3.57**
Number of total words	574.77	227.20	699.82	300.26	−1.93
Number of words per note	51.03	17.73	72.61	15.83	−5.66***

**$P < .01$; ***$P < .001$.

Fig. 5. Community knowledge as represented by a key-term cloud in Phase 1.

phase 1 ($M = 73.2\%$, SD $= 12.5\%$). This suggests that students contributed more key terms to the community knowledge space in phase 2. Moreover, in terms of how members increase the frequency of shared key-term use in the community knowledge space (i.e., frequency of total shared key-term use in the community divided by frequency of unique key-term use), the results showed a marginally significant difference between the two phases ($t = -1.976$, df $= 21$, $P < .10$) in that phase 2 has a higher ratio ($M = 3.68$, SD $= 1.41$) than phase 1 ($M = 2.97$, SD $= 1.05$), showing key terms are more likely to be repeatedly elaborated or referred in the community knowledge space in phase 2. The results here suggest a more reflective or depth-oriented inquiry in phase 2. Conventional online behavioral measures were not as sensitive as these key-term measures in differentiating and illustrating patterns involving individual contributions and community knowledge. Table 3 summarizes the results.

3.3. Analysis on relationships between the conventional and new key-term measures

The relationships between traditional online measures and the proposed new key-term measures were investigated using Pearson correlations for each phase (see Tables 4 and 5). Three observations follow: First, in both phases, almost all key-terms measures are significantly correlated with three sets of conventional measures which are # of notes created, # of problems worked on and # of total words generated per person. This was expected as typically the more notes/problems/words generated, the more key terms (or ideas/concepts). Second, it was found that in both phases, the majority of key-term measures were not correlated with the following three sets of conventional online measures: # of notes read, % of notes linked, and # of build-on notes. This seems to reflect the fact that connecting (i.e., linking and building-on) and reading are behaviors that are useful in creating connections between ideas but they do not directly generate ideas (or key terms), although it will be important to analyze note/problem/word generation patterns to better understand entry of unique terms into knowledge spaces and reading, linking, building-on patterns that lead to their uptake and advancement. Third, it was found that key-term measures are significantly correlated with two sets of conventional measures (i.e., # of notes edited and # of note revisions) in phase 2, but not in phase 1. Why? As suggested from the above findings, there is relatively stronger reflective activity going on in phase 2 than in phase 1, therefore, it is possible that the number of notes being edited and the number of total note revisions also increased in phase 2, thus significant positive correlations.

4. Summary and conclusion

In summary, findings based on the first comparison using conventional behavioral measures suggested a change of interaction patterns from more superficial inquiry (in phase 1) to more reflective, deeper inquiry (in phase 2). Findings based on the second comparison using key-term related measures further showed that there were higher percentages of shared key terms and higher frequency of use of shared

Table 3
Changes between phases 1 and 2: key-term measures ($N = 22$).

	Phase 1		Phase 2		t Values
	M	SD	M	SD	
Unique key-term contributions					
Number of unique key-terms	40.10	16.74	33.27	11.35	2.065
Frequency of unique key-term use	78.41	38.74	100.14	40.11	-2.483^{**}
Shared key-term contributions					
Number of shared key-terms	29.1	12.55	30.64	9.35	-0.617
Frequency of shared key-term use	203.77	63.77	324.36	68.02	-7.686^{***}
Interactions between unique and shared key-term contributions					
Number of shared key-terms divided by number of unique key-terms	73.2%	12.5%	93.3%	6.0%	-6.207^{***}
Frequency of total shared key-term use in a community divided by frequency of unique key-term use	2.97	1.05	3.68	1.41	-1.976^{*}

$^{*}p < .10$; $^{**}p < .05$; $^{***}p < .001$.

Table 4
Correlations between conventional behavioral and key-term measures in phase 1 ($N = 22$).

Behavioral measures	Key-term measures			
	# of total key terms	Frequency – use of total key terms	# of shared key terms	Frequency – use of shared key terms
# of notes created	0.42*	0.43**	0.53**	0.47**
# of problems worked on	0.37*	0.31	0.44**	0.38*
# of total words (per person)	0.65**	0.82***	0.73***	0.60**
# of words per note	0.36	0.54***	0.40*	0.26
# of notes read	0.29	0.25	0.42*	0.39*
% of notes linked	0.15	0.06	0.16	0.18
# of build-on notes	0.28	0.24	0.40*	0.34
# of notes edited	0.20	0.27	0.28	0.35
# of note revisions	−0.06	0.01	−0.03	0.06

*$p < .10$; **$p < .05$; ***$p < .01$.

key terms in the community knowledge space in phase 2 (as compared with phase 1). This change is important as having metacognitive awareness of keywords, concepts, or ideas discussed, exchanged, referenced, elaborated, or improved in the community space can play an important role in productive community knowledge building. This represents a central area where key-term measures can help complement conventional behavioral measures; and vice versa, key-term measures need to be complemented by online behavioral measures in identifying interaction patterns. Findings based on the third comparative analysis also confirmed complementary relationships between the conventional and new key-term measures.

Overall, similarities between the conventional and key-term measures showed the following important patterns: First, both types of measures can be used for automated assessment. Second, both can be easily integrated into online learning environments. Third, both are easy to implement, and with careful instructional design, can be used by teachers and even young students (see, e.g., Hong et al., 2008). On the other hand, there are also differences between the conventional and key-term measures. There are two important aspects of knowledge-building that conventional online measures do not support but that can be supported through complementary use of key-term measures. One is that key terms can be used to capture and represent important key concepts/ideas recorded in a community database. The other is that key-term measures can help to uncover relationships between community knowledge and individual contributions. Having deeper understanding of assessment about group or community knowledge is essential for understanding the dynamic nature of group learning, or learning by groups. The present study is a step forward in this direction.

A weakness of using key terms in instructional contexts is that the context in which the terms are used is not evident in key-term overviews. However, it was never an intention of this study to imply that context should be ignored. The current development of key-term measures is meant to serve as an easy-to-use instrument to help students and teachers summarize key ideas or concepts for building community knowledge in an online knowledge space. Students and teachers are the creators of the content in their community knowledge spaces, so use of the key-term tool by them provides an abstracted representation of their contributions from which they can easily return to the context of use. The key-term tool represents one of several metadiscourse tools available within Knowledge Forum (see also Resendes, Chen, Chuy, & Scardamalia, 2012). Constructive use of key-term measures depends on effective pedagogical designs and implementation. For example, a productive use of the key-term tool under exploration is asking students to compare the key terms in their notes with key terms in curriculum documents and identify conceptual gaps or advances for further knowledge advancement. We have seen students identify important advances in their notes that are not represented in the curriculum guidelines, leading to students' recommendations for improving the curriculum guidelines. Further studies are necessary to gain better understanding of how key-term measures can support knowledge building pedagogy.

Community knowledge as viewed via key-term measures in the present study may be thought of as providing an interesting perspective to support group reflection regarding the content space within which the community is advancing, including the key ideas, concepts, and problems collectively explored by the community. We aim to determine if the growth of key terms over time reflects evolution in the state of individual and community knowledge and to use indicators of advances and gaps to empower teachers and students. As long as ideas or concepts are recorded in a public knowledge space, key-terms can help capture and assess knowledge advances. Deeper understanding of use of these ideas, concepts, and problems for knowledge advancement may be more related to procedural and interaction patterns, and compared with key-term measures, conventional online behavioral measures play an important role in assessing how ideas are shared, linked, built on, and revised to form community knowledge.

Table 5
Correlations between conventional behavioral and key-term measures in phase 2 ($N = 22$).

Behavioral measures	Key-term measures			
	# of total key terms	Frequency – use of total key terms	# of shared key terms	Frequency – use of shared key terms
# of notes created	0.68**	0.64***	0.65***	0.54**
# of problems worked on	0.37*	0.51**	0.40*	0.40*
# of total words (per person)	0.78***	0.82***	0.77***	0.64***
# of words per note	0.35	0.58***	0.36*	0.34
# of notes read	0.10	−0.02	0.13	0.08
% of notes linked	0.11	0.17	0.09	0.31
# of build-on notes	0.39*	0.41*	0.34	0.46**
# of notes edited	0.62**	0.54**	0.60***	0.47**
# of note revisions	0.43**	0.47**	0.42**	0.29

*$p < .10$; **$p < .05$; ***$p < .01$.

As indicated in the introduction, it is important to assess 'learning by groups' not just 'learning in groups' or individual learning through social processes. Toward this end it is important to understand how an initial idea or concept is collaboratively pursued in depth over time to form a progressively more coherent idea or solution for addressing a problem. From a knowledge creation perspective, knowledge assessment and representation tools must show the improvement process of ideas and capture/represent promising ideas. This is of great significance for cultivating students' knowledge building capacities, as the sense of promisingness represents a type of knowledge that is found to be especially valuable as a resource for creative experts. Previous studies suggest that after continuously solving many problems in their area of expertise, expert learners (or teams of expert learners) possess a stronger sense of which ideas are more promising to pursue in finding a solution to a problem, and/or of how to improve, refine, or redesign the problem space (e.g., Schwartz & Martin, 2004). A tool that can document collaborative learning processes and capture promising ideas would be very useful to help learners develop a strong sense of collaboration and promisingness for sustained community knowledge advancement. Therefore, an optimal assessment tool should not just retrospectively assess content and progress in a community knowledge space, but should also foster the collective development of promising ideas among members to support their community knowledge work. Important work and tools are under development to help students assess the promisingness of ideas in their collaborative spaces (Chen, Scardamalia, Resendes, Chuy, & Bereiter, 2012). Given such tools, visualizations of key-term cloud, the evolution of key terms in "promising idea" selections, and other features built into a key-term tool may help teachers and students visually (1) summarize the content (like a concept map) in a database, (2) search for and trace related ideas/concepts to enhance knowledge building discourse (Sinclair & Cardew-Hall, 2008), (3) encourage more opportunistic collaboration on emergent ideas (Hong, 2011; Hong, Chen, Chai, & Chan, 2011; Zhang et al., 2009), and (4) compare and link ideas between groups in the community (Oshima, Oshima, & Matsuzawa, 2012). To further enhance the capacity of the key-term tool in order to better capture collaborative dynamics and promising ideas, a more integrated technological design approach needs to be considered. Given the complementary nature of both the conventional behavioral and the key-term measures, it is necessary to take the complex interactional processes among community members into consideration when developing key-term measures. The knowledge building discourse analyzer, KBDeX (Oshima et al., 2012) represents an important advance in this direction. A fruitful future development of the key-term tool would be to integrate this work with visualization and comparison capacity over a broad range of conventional online behavioral measures. In so doing, the collaborative process of developing promising ideas may be more readily identified and traced to support sustained community knowledge building. Overall, results suggest that use of key-term measures for visualizing community knowledge advances and for fostering students' individual contributions to collective efforts may well lead to more effective assessment as well as more powerful supports for knowledge creation.

Acknowledgments

Support for writing this article was provided, in part, from a Social Sciences and Humanities Research Council of Canada grant (512-2002-1016) and a Taiwan's National Science Council grant (NSC # 101-2628-S-004-001-MY3).

References

van Aalst, J. (2012). Assessment in collaborative learning. In C. Hmelo-Silver, C. Chinn, C. K. K. Chan, & A. O'Donnell (Eds.), *The international handbook of collaborative learning*. Taylor & Francis/Routledge.

Anderson, J. R. (2000). *Cognitive psychology and its implications* (5th ed.). New York: Worth.

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bereiter, C., & Scardamalia, M. (1996). Rethinking learning. In D. R. Olson, & N. Torrance (Eds.), *The handbook of education and human development: New models of learning, teaching and schooling* (pp. 485–513). Cambridge, MA: Basil Blackwell.

Bourner, J., Hughes, M., & Bourner, T. (2001). First-year undergraduate experiences of group project work. *Assessment & Evaluation in Higher Education, 26*(1), 19–39.

Burtis, J. (2002). *Analytic toolkit for knowledge forum (Version 4.0)*. Toronto, ON: Institute for Knowledge Innovation and Technology, Ontario Institute for Studies in Education/University of Toronto.

Cheng, W., & Warren, M. (2000). Making a difference: using peers to assess individual students' contributions to a group project. *Teaching in Higher Education, 5*(2), 243–255.

Chen, B., Scardamalia, M., Resendes, M., Chuy, M., & Bereiter, C. (2012). Students' intuitive understanding of promisingness and promisingness judgments to facilitate knowledge advancement. In J. van Aalst, K. Thompson, M. J. Jacobson, & P. Reimann (Eds.), *The future of learning: Proceedings of the 10th international conference of the learning sciences (ICLS 2012)* (Vol. 1; pp. 111–118). Sydney, Australia: ISLS. Full Papers.

Czado, C., & Munk, A. (1998). Assessing the similarity of distributions-finite sample performance of the empirical Mallows distance. *Journal of Statistical Computation and Simulation, 60*(4), 319–346.

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysiss schemes to analyze transcripts of online asynchronous discussion groups: a review. *Computer & Education, 46*, 6–28.

Hassan-Montero, Y., & Herrero-Solana, V. (2006). Improving tag-clouds as visual information retrieval interfaces. In *Paper presented at the international conference on multidisciplinary information sciences and technologies (InSciT), Merida, Spain*.

Hong, H. Y. (2011). Beyond group collaboration: facilitating an idea-centered view of collaboration through knowledge building in a science class of fifth-graders. *The Asia-pacific Education Researcher, 20*(2), 248–262.

Hong, H. Y., Chen, F. C., Chai, C. S., & Chan, W. C. (2011). Teacher-education students' views about knowledge building theory and practice. *Instructional Science, 39*(4), 467–482.

Hong, H.-Y., & Lin-Siegler, X. (2012). How learning about scientists' struggles influences students' interest and learning in Physics. *Journal of Educational Psychology, 104*(2), 469–484.

Hong, H.-Y., Scardamalia, M., Messina, R., & Teo, C. L. (2008). Principle-based design to foster adaptive use of technology for building community knowledge. In G. Kanselaar, V. Jonker, P. A. Kirschner, & F. J. Prins (Eds.), *ICLS 2008* (Vol. 1; pp. 374–381). Utrecht, the Netherlands: International Society of the Learning Sciences, Inc.

Hong, H.-Y., Teplovs, C., & Chai, C. S. (2007). On community knowledge. In B. Chong, A. Kashihara, J. Lee, T. Matsui, R. Okamoto, D. Suthers, et al. (Eds.), *ICCE 2007 Proceedings* (pp. 292–294). Tokyo: IOS Press.

Hyman, J. (1999). How knowledge works. *The Philosophical Quarterly, 49*(197), 433–451.

Lee, E. Y., Chan, C. K., & van Aalst, J. (2006). Students assessing their own collaborative knowledge building. *International Journal of Computer-supported Collaborative Learning, 1*(1), 57–87.

Lejk, M., & Wyvill, M. (2001). Peer assessment of contributions to a group project: a comparison of holistic and category-based approaches. *Assessment & Evaluation in Higher Education, 26*(1), 61–72.

Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Mahwah, NJ: Lawrence Erlbaum Associates.

Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge building discourse explorer: a social network analysis application for knowledge building discourse. *Educational Technology Research and Development, 60*, 903–921.

Popper, K. R. (1972). *Objective knowledge: An evolutionary approach*. London: Oxford Univ. Press.

Resendes, M., Chen, B., Chuy, M., & Scardamalia, M. (2012, August). The effect of meta-discourse on ways of contributing to an explanation-seeking dialogue in grade 2. In *Paper presented at the 16th Knowledge Building Summer Institute (KBSI2012), Toronto, Canada*. Web: ikit.org/SummerInstitute2012/Papers/3024-Resendes.pdf.

Scardamalia, M. (2004). CSILE/Knowledge Forum®. In A. Kovalchick, & K. Dawson (Eds.), *Education and technology: An encyclopedia* (pp. 183–192). Santa Barbara, CA: ABC-CLIO.
Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In *Encyclopedia of education* (pp. 1370–1373). New York: Macmillan Reference, USA.
Scardamalia, M., & Bereiter, C. (2006). Knowledge building: theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 97–118). New York: Cambridge University Press.
Scardamalia, M., & Bereiter, C. (2010). A brief history of knowledge building. *Canadian Journal of Learning and Technology, 36*(1). Published online at http://www.cjlt.ca/index.php/cjlt/article/view/574.
Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: the hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129–184.
Sinclair, J., & Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful? *Journal of Information Science, 34*(1), 15–29.
Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press. Retrieved from http://www.cis.drexel.edu/faculty/gerry/mit/.
Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications.
Teplovs, C. (2008). The knowledge space visualizer: a tool for visualizing online discourse. In *Paper presented at the common framework for CSCL interaction analysis workshop, international conference of the learning sciences*. Utrecht: NL.
Teplovs, C., Donoahue, Z., Scardamalia, M., & Philip, D. (2007). Tools for concurrent, embedded, and transformative assessment of knowledge building processes and progress. In Clark A. Chinn, Gijsbert Erkens, & Sadhana Puntambekar (Eds.), *Proceedings of the 8th international conference on computer supported collaborative learning (CSCL'07)* (pp. 721–723). International Society of the Learning Sciences.
Teplovs, C., & Scardamalia, M. (2007). *Visualizations for knowledge building assessment*. AgileViz workshop: CSCL2007.
Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249–276.
Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge-building communities. *The Journal of the Learning Sciences, 18*(1), 7–44.