Education Tech Research Dev
DOI 10.1007/s11423-012-9265-2

DEVELOPMENT ARTICLE

Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse

Jun Oshima · Ritsuko Oshima · Yoshiaki Matsuzawa

Abstract In recent studies of learning theories, a new methodology that integrates two prevailing metaphors of learning (acquisition and participation) has been discussed. However, current analytical techniques are insufficient for analyzing how social knowledge develops through learners' discourse and how individual learners contribute to this development. In this paper, we propose a novel approach to analyzing learning from an integrative perspective and present a social network analysis application that uses learner discourse as input data: Knowledge Building Discourse Explorer (KBDeX). To investigate the utility of this approach, discourse data analyzed in a previous study is re-examined through social network analysis supported by KBDeX. Results suggest that social network analysis can qualitatively and quantitatively support the conclusions from the previous study. In addition, social network analysis can reveal potential points that are pivotal for social knowledge advancement in groups, and can identify each individual's contribution to this advancement. On the basis of these results, we discuss how social network analysis could be integrated into existing in-depth discourse analysis.

Keywords New learning metaphor · Discourse analysis · Social network analysis · Knowledge Building Discourse Explorer

Recent learning theory studies have discussed a new approach that integrates two prevailing metaphors of learning: acquisition and participation (Paavola et al. 2004; Sfard 1998). However, current assessment techniques do not act in concert with the development of such a theoretical approach to learning, which this study addresses. In this section, we discuss a possible approach to analyzing learning from an integrative perspective by using network theory. A brief review is given of recent literature on learning theories that introduce a new perspective on learning: the knowledge-creation metaphor. By using the well-known knowledge building community model, current assessment techniques are

J. Oshima (✉) · R. Oshima · Y. Matsuzawa
Faculty of Informatics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu-shi, Shizuoka 432-8011, Japan
e-mail: joshima@inf.shizuoka.ac.jp

Published online: 30 June 2012

shown to be insufficiently able to describe learning in the knowledge-creation metaphor. Hence, to address this deficiency, social network analysis (SNA) is introduced as a novel assessment approach for learning in the knowledge-creation metaphor.

Metaphors of learning and corresponding assessment approaches

Sfard (1998) examined two different learning perspectives—the acquisition metaphor and the participation metaphor—and discussed the dangers in committing to only one of them. In the acquisition metaphor, according to the folk theory of mind, the mind is a container for holding knowledge. Learning is thus the process by which the container is filled, with knowledge internalized by learners themselves (i.e., constructivism) or transmitted from others who are more knowledgeable, for example, teachers. In this view, knowledge is understood as a property or capacity of an individual's mind, and learning is expressed by words such as construction, acquisition, and outcomes, through the process of applying existing knowledge to new contexts (i.e., knowledge transfer). The acquisition metaphor has influenced computational models of the mind and of knowledge which are used to investigate how the individual mind works (e.g., Anderson et al. 1996).

The main concern of the participation metaphor is with learning as a process of participation in cultural practices and shared learning activities. The focus of the metaphor is therefore on activities of "knowing" rather than "knowledge" as a product. In this way, knowledge is not present in individual human minds, but is an aspect of participation in cultural practices (Brown et al. 1989; Lave and Wenger 1991). Activities of knowing cannot be separated from the situations where they take place. Hence, learning takes place through participation in *practices and actions* (Anderson et al. 1996; Greeno 1997), *enculturation* (Brown et al. 1989), or *legitimate peripheral participation* (Lave and Wenger 1991). In this view, terms such as acquisition and accumulation are displaced in favor of concepts such as discourse, interaction, activity, and participation.

Through an intensive exchange between the acquisition and participation metaphors, several researchers (e.g., Anderson et al. 2000; Sfard 1998) have concluded that both metaphors are needed in order to enrich learning analysis, and that unified theories of learning could be developed by starting with a consensus approach that included both metaphors (Cobb and Bowers 1999; Greeno et al. 1997).

Following Sfard (1998), Paavola et al. (2004) continued this theoretical discussion of a unified learning theory by proposing a third learning metaphor: the knowledge-creation metaphor. They examined three prominent models of innovative knowledge communities (Bereiter 2002; Engeström 1999; Nonaka and Takeuchi 1995) to identify the aspects of the knowledge-creation metaphor necessary for establishing a theoretical framework that unifies the acquisition and participation metaphors. The knowledge-creation metaphor is critically different from the acquisition and participation metaphors in that it is concerned with the pursuit of newness; specifically, the creation of new knowledge based on given knowledge (Bereiter's model, and Nonaka and Takeuchi's model) or the transformation of a current activity system into a new systematic level (Engeström's model).

For instance, Scardamalia and Bereiter (2003) explain the nature of knowledge building as one of the prominent models of the knowledge-creation metaphor by referring to two modes of learning: the belief mode and the design mode. In the belief mode, learners are concerned with what they or others believe or ought to believe, namely, with the mental states of individuals. On the contrary, in the design mode, learners are concerned with the usefulness, adequacy, improvability, and developmental potential of ideas. Learners in the design mode should be aware of whether their ideas are good enough to solve the problems to be addressed, and how they should contribute to improving those ideas. Knowledge building is a social process that engages both modes of learning. The belief mode is used by learners to investigate the current state of their community knowledge level in order to highlight any problems. Learning in the design mode thus enables the creation of knowledge to solve problems. Exchange between learning modes is iterative, such that learners continuously participate in social practices of knowledge creation, and individuals generate knowledge that not only directly contributes to the advancement of community knowledge but also determines how best to contribute to this advancement.

Necessity of new assessment approaches in knowledge-creation metaphor

How do we assess learning in knowledge building? Many researchers rely on current assessment approaches based on either the acquisition metaphor or the participation metaphor. However, as discussed above, assessment based on the knowledge-creation metaphor requires higher complexity as compared with assessment based on the other metaphors or a combination of them.

In the acquisition metaphor, researchers are mainly concerned with what and how much knowledge is acquired by learners after learning experiences, and which activities are crucial to more profound knowledge acquisition (Sfard 1998). Often these researchers use quantitative analysis based on systematic observation to assess learning (Mercer 2005), and a typical assessment is conducted with a pre- and post-test design. By using a variety of methods, including paper-and-pencil tests, computerized tests, and presentations, learners' understanding in specific domains before and after instruction is gauged to identify knowledge acquisition. In addition, processes by which learners attain different levels or types of knowledge are investigated by coding their activities (including written or oral discourse) and counting frequencies for comparison across different conditions. For instance, van Aalst (2009) assessed secondary students' written discourse with a computer-supported collaborative learning (CSCL) system for supporting knowledge building, called Knowledge Forum (Scardamalia 2003). With the coding scheme he developed, van Aalst evaluated students' summary discourse after learning and, based on written discourse, how their activities during learning related to the quality of their summary discourse. The quality of students' summary discourse was found to be higher when they were engaged in knowledge-creation discourse.

Researchers who prefer the participation metaphor predominantly consider detailed case studies in assessments (Mercer 2005), taking advantage of real cases, rather than abstracted categorization of data. Actual discourse remains throughout such analyses and the processes of constructing shared meaning can be examined in detail. Our research group (Oshima et al. 2006) examined how Japanese elementary students engage in knowledge building in a science class by placing the students' lab reports on Knowledge Forum, encouraging the students to read one another's experimental procedures and explanations of results. Comparing detailed descriptions from several groups in the classroom with data from the previous year, we found that students more frequently engaged in building larger communities consisting of several groups that share similar questions and problems through their intergroup communication by reading and commenting on other groups' reports. Students' intentional efforts at community building beyond their own groups enabled them to design appropriate experiments with greater complexity and to develop sophisticated explanatory models.

The research described above has shown that either metaphor-driven approach can assess certain aspects of knowledge building community models. However, there are disadvantages in using these assessment techniques. In the knowledge-creation metaphor, researchers are not only concerned with learners' comprehension of domain-specific knowledge but also with individuals' contributions to community knowledge. The learners' epistemic activities of utilizing their individual knowledge to improve their community knowledge cannot be analyzed by assessing static knowledge in a pre- and post-test design. Categorization of written or oral discourse during learning might be able to identify the epistemic activities, but such a coding scheme is content-free and what knowledge learners actually contribute cannot be examined.

The detailed description analysis of discourse in the participation metaphor can be used to complement the acquisition metaphor. Such analysis can reveal how a group of learners engage in their joint activities to construct shared knowledge in their groups as communities. However, as suggested by Mercer and colleagues (Mercer 2005; Wegerif and Mercer 1997), detailed description analysis of discourse can be impractical for large datasets because it is highly time-consuming. Consequently, Wegerif and Mercer (1997) developed a methodology to combine detailed description analysis of discourse and computerized discourse analysis to handle large datasets. Although their approach is useful in qualitative and quantitative analyses, the focus of the methodology is still on the participation metaphor, rather than a combination of the acquisition and participation metaphors or the knowledge-creation metaphor.

In this paper, SNA of discourse is proposed as an assessment approach in the knowledge-creation metaphor. In the next section, a brief review is given of literature on complex networks that investigates knowledge creation and on educational studies that apply SNA to collaborative learning.

SNA as a novel assessment approach in the knowledge-creation metaphor

In research outside of the educational field, knowledge creation is a major issue (e.g., Nonaka and Takeuchi 1995). In organizations creating knowledge, workers have a variety of resources and are linked to one another by multi-layer communication channels (Watts 2007). The social network generated through these channels is an infrastructure that can provide new insights and knowledge (Barabási 2005; Davis and Sumara 2006; Scardamalia and Bereiter 2003). Researchers studying complex networks have recently extended their interests toward how social networks influence knowledge-creating organizations or communities (e.g., Barabási 2005; Strogatz 2001). Guimera et al. (2005)investigated collaborative networks of creative teams (artists and scientists) to discover how team assembly mechanisms determine network structures. As a result, a self-assembly model was formulated based on the team size, the fraction of newcomers in a new project, and the tendency of incumbents to repeat previous collaborations. Model simulations suggested a phase transition occurs between the initial network state composed of numerous small isolated clusters and the emergence of a large connected community of practitioners. Furthermore, continual replacement of incumbents by newcomers was found to be needed for creative teams to remain highly influential in their field, while maintaining established relationships among incumbents was also needed to a certain degree.

In another study, Palla et al. (2007)developed an algorithm to investigate the time dependence of overlapping communities on a large scale. They attempted to ascertain the basic relationships characterizing community evolution in groups of scientists and of mobile telephone users. Results showed that (a) large groups are sustained when they can dynamically modify their membership and (b) small groups are stable when membership is kept unchanged. In addition, the time members commit to their community was found to be an indicator of the community's lifetime. Thus, computational power and human behavioral data can be used to explore the mechanisms of both community development and knowledge advancement. Lazer et al. (2009) called this direction of research *computational social science*.

In educational research, the graphical approach to representing and evaluating learners' knowledge structures in their minds has been examined for many years (e.g., Cañas et al. 2004; Pirnay-Dummer et al. 2008; Schvaneveldt 1990). In particular, *Concept Mapping Tools* developed by Cañas et al. (2004) provides researchers with visualization of students' semantic networks as their individual knowledge structures, and many researchers have used the technology in instructional design to improve students' conceptual understanding in science education. In most research, researchers have used concept maps in the acquisition metaphor. They have been concerned with how individual knowledge structures develop during students' learning. For instance, Ifenthaler et al. (2011) used *Concept Mapping Tools* for students to report their conceptual understanding across different times during their learning processes. Ifenthaler et al. (2011) developed an integrative framework for analyzing students' knowledge structures represented in concept maps based on graph theory, and examined whether their established measures were sensitive enough to detect students' learning progression.

On the other hand, in research on networked learning and computer-supported collaborative learning, there have been discussions on the advantages of using SNA to investigate community knowledge advancement and individual learners' engagement in this advancement from the perspective of the knowledge-creation metaphor (e.g., Martinez et al. 2003; Reffay et al. 2011; Reuven et al. 2003). Like concept maps, the representations used in SNA are also networks. The critical difference between the network representations in SNA from those in concept maps is the social nature of SNA. Researchers apply SNA to students' social activities such as knowledge building, and therefore attempt to examine the participatory structures in students' collaborative learning and their social knowledge advancement in that context.

de Laat et al. (2007) considered the application of SNA in CSCL research. They outlined an approach to synthesize and extend the understanding of CSCL teaching and learning processes so as to balance SNA, content analysis and critical event recall. In this complementary approach, SNA was used to study interaction patterns within a networked learning community, as well as to study how learners share and construct knowledge. de Laat et al. concluded that SNA should be included in any multi-method approach because of the following advantages: (a) researchers and learners are provided with a tool able to illustrate understanding and cohesion of group activities, and (b) a method is made available to researchers for selecting appropriate groups to study.

A limited number of studies have used SNA, especially in the knowledge-creation metaphor. Over a period of three years, Zhang et al. (2009) implemented a complementary approach that used SNA to visualize and compare classroom collaboration among fourth grade elementary school students through a CSCL environment designed to support them in knowledge building. An analysis of the students' online participatory patterns and knowledge advancement indicated that this learning process facilitated students' knowledge advancement effectively, and that this was the case through critical changes in organizations within the classroom: from fixed small groups in the first year of the study to appropriate collaboration through dynamic formation of small teams based on emergent goals.

In previous work (Oshima et al. 2011), we further extended the potential of SNA as a core assessment technique by describing a different type of social network. Ordinary SNA illustrates the social patterns of learners, namely, the learners' social network. As de Laat et al. (2007) suggested, this approach is thus informative when examining developments or changes in the participatory structure of a community. However, we argued that existing social network models are unable to examine how community knowledge advances through learners' collaboration (Oshima et al. 2007). Instead, we used a procedure similar to ordinary SNA, but proposed a different type of social network, one based on the *words* learners use in their discourse in a CSCL environment. We compared this social network, in which words were selected as nodes representing learners' knowledge or ideas during discourse on a study topic, with a network of words from the discourse of a group of experts on the same topic. The results showed that there were remarkable differences in the community knowledge of elementary school students and of experts in terms of the words centered on the networks. We concluded that SNA can provide a new type of representation of community knowledge building by learners, enabling researchers to adopt a new complementary assessment technique for investigating knowledge building community models.

Although educational studies have proposed the application of SNA to learning analysis as a new assessment technique in the knowledge-creation metaphor, an exact methodology has yet to be established. Several researchers familiar with SNA in education science, and especially those publishing results in scholarly journals, are involved in developing or refining software for SNA. SNA software that can easily explore discourse data is needed for those interested in using SNA to analyze discourse data for examining participation patterns and states of community knowledge. As a consequence, we have developed the Knowledge Building Discourse Explorer (KBDeX) software. In this paper, we introduce KBDeX and explain how three different types of discourse-based social networks are represented. Moreover, SNA conducted with KBDeX is demonstrated by using two sets of oral discourse data from small groups. Finally, the potential contribution of KBDeX to a new complementary assessment approach in the knowledge-creation metaphor is discussed.

KBDeX: an application of SNA of discourse

KBDeX runs on any platform where Java runtime environment is installed (Mac, Windows, Linux, etc.). The primary functions of KBDeX are as follows.

Data source for SNA with KBDeX

KBDeX visualizes discourse network structures based on a bipartite graph of *words* $\times$ *discourse units* (e.g., conversation turns, postings on an online forum, and sentences). To conduct the analysis, the original discourse data (e.g., transcriptions of oral discourse or written discourse in computer logs) are transformed by (a) merging singular and plural forms of nouns; (b) merging different words having the same meaning (e.g., "Net" and "Internet"); (c) merging the different conjugated forms of verbs if the user needs to analyze them; and (d) removing unrelated discourse units (e.g., social chatting and digressions from their problem solving).

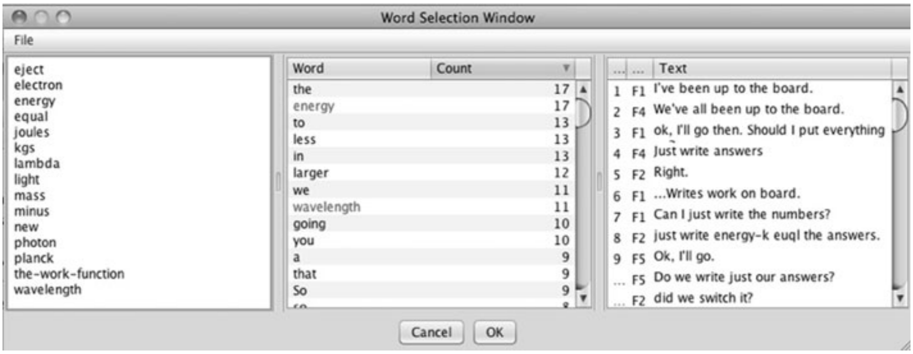

Fig. 1 The word selection window in KBDeX

Before analysis, a list of keywords must also be selected according to the analyst's objectives, for example, conceptual words for examining how learners use subject matter knowledge in their discourse or epistemic words for examining how learners attempt to regulate learning. A word selection window to support this process is built into KBDeX (Fig. 1). This window has three panels: (a) a list of the selected words (left); (b) a list of all words in descending order of frequency (center); and (c) the input discourse (right). The discourse units can be specified by the analyst as sentences, conversation turns, or other grain sizes of discourse. The analyst selects words by double-clicking on them within the word frequency list, and the selected words then appear in the list and are indicated in red on the monitor (e.g., "energy" and "wavelength") in panels (b) and (c).

Network building and basic features of KBDeX

KBDeX simultaneously visualizes three different network structures. The analyst can watch the unit-by-unit construction of these networks, and can pause and resume the process of network building. As seen in Fig. 2, the main view of the KBDeX graphical user interface has four windows: (a) a discourse viewer showing an overview of the discourse with the selected words marked in red on the monitor (top left); (b) the network structure of learners (top right); (c) the network structure of discourse units (bottom left); and (d) the network structure of selected words (bottom right). The networks of discourse units and selected words are created from the bipartite graph of *units* × *words*, with each network shown as a one-mode projection of the graph (Fig. 3). The network of learners is also a one-mode projection of the *words* × *learners* bipartite graph.

Although the analyst can select the layout algorithm for the three network structures, a circular layout is used by default for the networks of learners and discourse units, and for the latter network, the nodes are sorted by time of occurrence. For the network of selected words, a layout described by Fruchterman and Reingold (1991) is used in order to illustrate the links between words. In the layout algorithm, nodes are represented by steel rings, and edges are represented by springs; the network is then generated by attractive and repulsive forces corresponding to the spring force and the electrical force, respectively. The analyst sees that words with more links are pulled together whereas words with no links are pushed further apart until all the forces are in equilibrium.

The analyst can explore the three network structures seamlessly by clicking on a particular node, and KBDeX indicates related nodes in the other two networks; the color of the

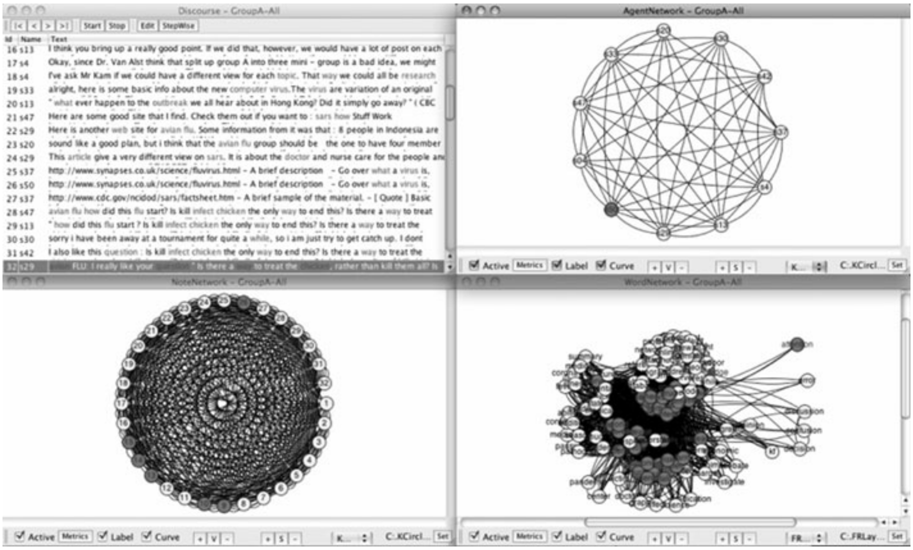

Fig. 2 The main view of KBDeX

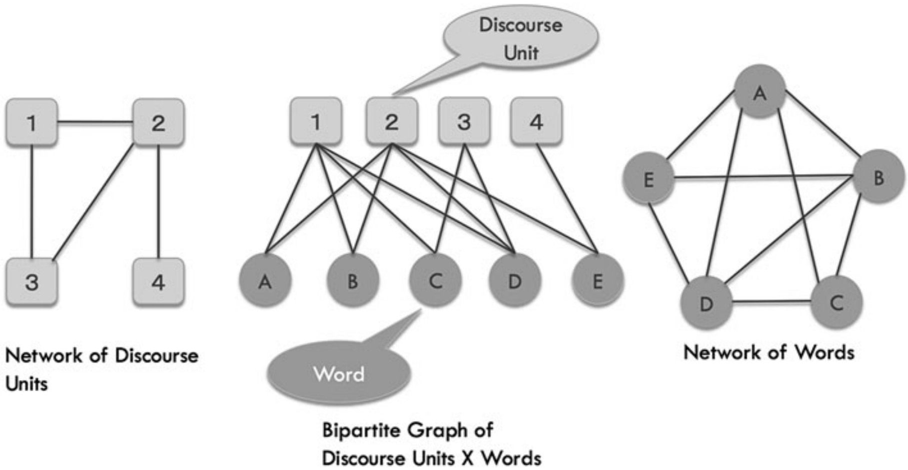

Fig. 3 Networks of discourse units and selected words as two different projections of the discourse units × selected words bipartite graph

selected and related nodes changes from yellow to red. For instance, when the analyst selects a particular learner in the network of learners, the discourse units given and words used by that learner are automatically identified.

Analysis of network characteristics by using coefficients

During SNA, KBDeX calculates conventional network measures, such as the betweenness centrality coefficient, the degree centrality coefficient, and the closeness centrality coefficient (Fig. 4).

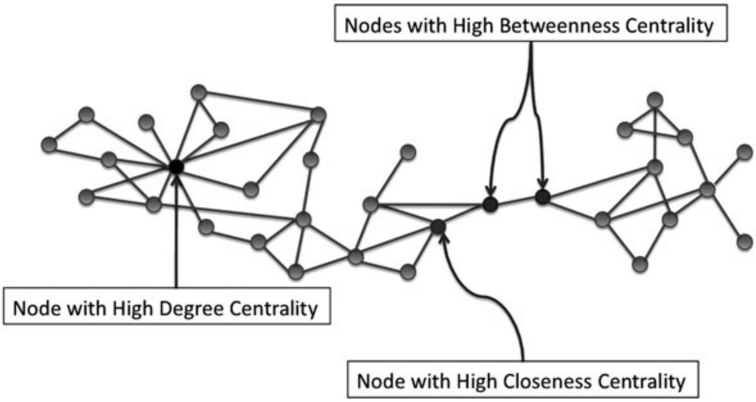

Fig. 4 Three centralities in a network

Betweenness centrality is a measure of the number of node pairs for which the shortest path between them passes through a selected node. High betweenness centrality suggests that the selected node works as a key mediator in linking other nodes. For a network with n nodes, the betweenness centrality, $C_b(i)$, for node i is

$$C_b(i) = \sum_{i \neq j \neq k} \frac{g_{jk}(i)}{g_{jk}},$$

where g_{jk} is the number of shortest paths (i.e., paths with minimal numbers of nodes) from j to k, and $g_{jk}(i)$ is the number of shortest paths from j to k that pass through node i. The normalized betweenness centrality, $C'_b(i)$, in an undirected network that has values lying between zero and one, is thus

$$C'_b(i) = \frac{2}{(n-1)(n-2)} \sum_{i \neq j \neq k} \frac{g_{jk}(i)}{g_{jk}}.$$

Degree centrality is a straightforward concept that indicates cumulative path lengths by which each node is linked to other nodes in the network. High degree centrality means that the node is at the center of the network as a whole, or near the center of a local cluster in the network. The degree of a node in a network is the number of edges connected to it. Thus, for a network with n nodes, the degree centrality, $C_d(i)$, for node i is

$$C_d(i) = \sum_{j=1}^{n} a_{ij},$$

where a_{ij} is the adjacency matrix for each pair of nodes (i, j), and the normalized degree centrality, $C'_d(i)$, is

$$C'_d(i) = \frac{1}{n-1} \sum_{j=1}^{n} a_{ij}.$$

Closeness centrality is a more sophisticated measure of how close the node is to other nodes in a network, based on the geodesic distance. For a network with n nodes, the closeness centrality, $C_c(i)$, for node i is

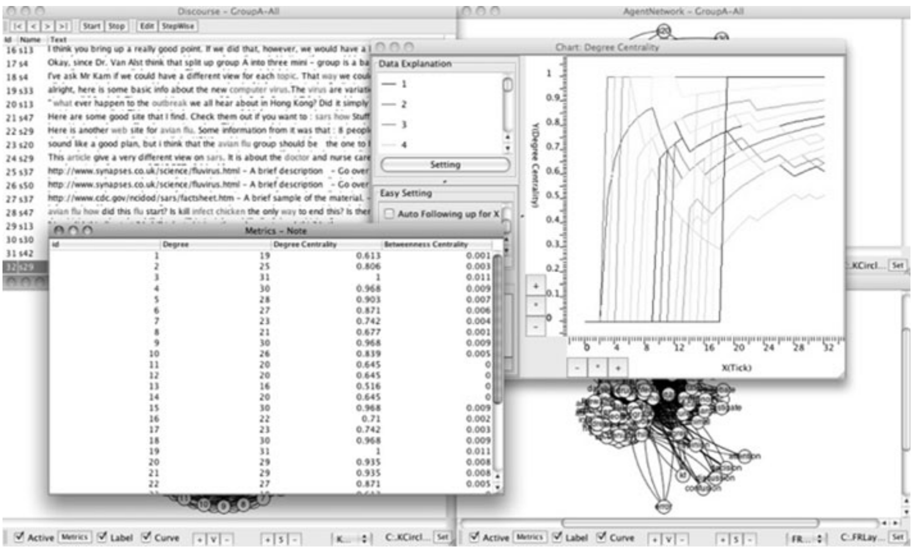

Fig. 5 Coefficient values and plots in KBDeX

$$C_c(i) = \frac{1}{\sum_{j=1}^{n} d_{ij}},$$

where d_{ij} is the length of a geodesic path from i to j, explicitly, the number of edges along the path. The normalized closeness centrality, $C'_c(i)$, is

$$C'_c(i) = \frac{n-1}{\sum_{j=1}^{n} d_{ij}}.$$

In KBDeX, coefficients of each node are calculated and dynamically plotted with the progress of the discourse (Fig. 5). Hence, when analysts find notable changes in values, they can return to the input discourse data and three networks to speculate why such a change occurred.

Phase and stepwise analysis

A feature of KBDeX is that the analyst can deactivate nodes in the networks. The deactivated nodes are shown in gray, and the bipartite connections related to the nodes are dimmed (Fig. 6). This function may be used for two analytic purposes. First, by activating or deactivating discourse units for different time periods in the network, the analyst can divide the entire discourse into a number of phases of collaborative learning, and compare coefficients across these different phases. Second, the analyst can deactivate specific learners to evaluate each individual's contribution to the networks. For instance, in our previous work, we calculated each learner's contribution by omitting his or her written discourse from the dataset and finding the averaged absolute difference in the centrality coefficients of the network of selected words (Oshima et al. 2007). In other words, a comparison is made between the centralities calculated when the network structure includes the target learner's discourse and when the network does not include this

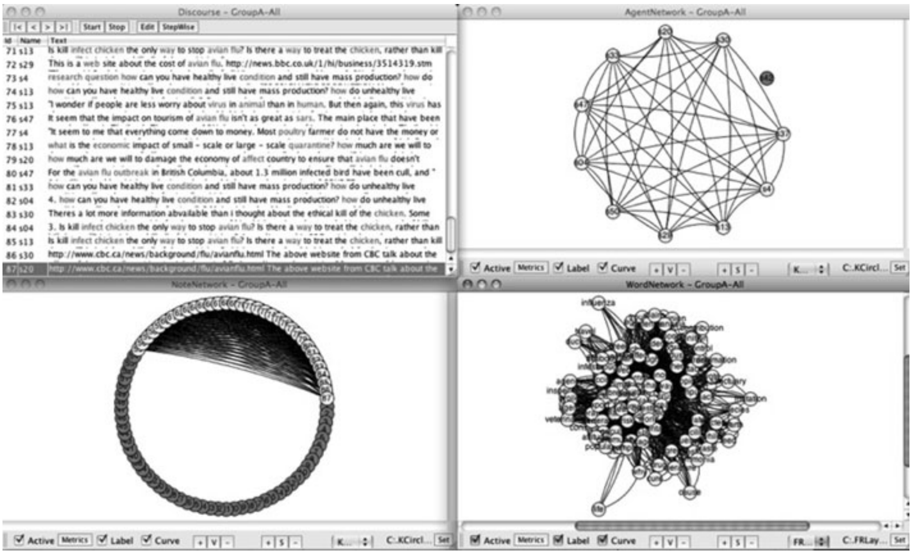

Fig. 6 The main view during stepwise analysis

discourse. In the KBDeX environment, the analyst can easily carry out this process by deactivating the learner who is the target of the analysis.

Methods and results

Example of discourse data for SNA with KBDeX

As an example, an analysis was conducted by using KBDeX on discourse data from two groups of three university students taken from a previous analytic study (Sawyer et al., in press). From the results of an in-depth discourse analysis, Sawyer et al. concluded that the two groups were significantly different in their strategic approaches to solving a general chemistry problem, namely, calculating the wavelength of an electron discharged from an object by utilizing formulas related to the photoelectric effect and de Broglie wavelength. On the basis of conversation analysis for the two groups, Sawyer et al. described profiles of the groups (Gillian versus Matt). A summary of these profiles is as follows.

The Gillian group went beyond pure calculation by discussing conceptual ideas about what they had learned and engaged in collaborative knowledge construction through mutual reflection of ideas. Conversely, the Matt group was involved in calculation activities without articulate recognition of what they had learned (Sawyer et al., in press). In terms of metaphors of learning, the Gillian group was more oriented toward the knowledge-creation metaphor (i.e., engagement in social knowledge advancement) in their problem solving, whereas the Matt group was oriented toward the acquisition metaphor (i.e., application of their procedural knowledge) in solving the same problem. In fact, Sawyer found that Gillian group acquired higher conceptual understanding than did Matt group (R. K. Sawyer, personal communication, March 7, 2012). Thus, results suggest that once a collaborative group culture oriented toward the knowledge-creation metaphor has emerged, it can facilitate students' engagement in knowledge building and in constructing understandings that support integration and application of the content (Blumenfeld et al. 2006; Engle and Conant 2002; Greeno 2006; Sawyer 2006; Scardamalia and Bereiter 2006).

From the same dataset, (i.e., transcriptions from video recordings of students' problem solving activities), SNA using KBDeX was conducted to assess the engagement of the two groups of students in their social knowledge advancement through collaborative learning. The focus of SNA was on collaborative learning processes themselves rather than final solutions. Firstly, time course changes in networks of selected words were visually inspected to examine whether social knowledge advancement differed between the two student groups and find pivotal points that lead to those differences. Secondly, three coefficients were used to analyze network characteristics and examine whether results from visual inspection were supported. Lastly, a stepwise analysis of each learner's contribution to the group discourse was conducted to characterize cognitive group dynamics in the two groups.

Bipartite graphs based on discourse data

Transcriptions of the discourse data provided by Sawyer et al. (in press; see Table 1 for an example) were used to create bipartite graphs for the two groups. The original discourse was approximately 9 min long between members of the Gillian group and 11 min long between those of the Matt group. Moreover, there were 87 turns in conversation for the Gillian group and 35 for the Matt group. To examine social knowledge advancement during their collaborative learning when solving the general chemistry problem, words conceptually related to solving the problem (i.e., those used in formulas) were selected. Students used 18 conceptual words in the Gillian group and 14 in the Matt group (Table 2).

Table 1 An example discourse between members A1, A2, and A3 in the Gillian Group [transcription taken from Sawyer et al. (in press)]

A1	So we need lambda and its give us the *** so with the wavelength we can find velocity. And with…
A3	And we don't have to works [sic] with the-work-function right away
A1	Not right away but we do need the-work-function at the end.
A3	Ok to find the…
A1	Because the important thing that for the lambda we are find the wavelength of the electron not of light
A3	Right, exactly. So, first for the electron we use Planck/mass velocity because we know the mass of an electron
A1	We need to find the mass of an electron
A3	No, we know the mass of an electron. It's an electron.
A1	Very true, very true
A3	But we don't know…
A1	With this wavelength we would be find velocity
A2	What did they give us. For the following wavelength. So, well lambda equal [sic] Planck/mass velocity right?
A1	Yes, its tell us to use lambda equal Planck/mass velocity
A1	You can find energy-k
A3	Energy-k of a photon
A1	Use energy-k of the-work-function equal [sic] energy-k equal Planck nu minus the-work-function equal energy-k of a photon and use energy-k of a photon to find the wavelength of light

Transcription taken from Sawyer et al. (in press)

Table 2 Word lists selected for SNA with KBDeX

Gillian Group	Eject, electron, energy-k, energy-p, equal, joule, kg, lambda, light, mass, meter, minus, nu, photon, Planck, the-work-function, velocity, wavelength
Matt Group	Energy-p, the-work-function, equation, energy-k, velocity, mass, electron, square-root, minus, eject, lambda, nm, angstrom, energy-x

Visual inspection of network of words

A feature of KBDeX is its visualization support. The analyst can visually inspect each event that occurs in learners' discourse at a network level and discover pivotal points (or key conversation turns) within the discourse by observing changes in all three networks. Here, an example is given of such a network visualization analysis.

To identify the problem solving methods of the two groups, the differences between their strategic approaches, and their pivotal conversation turns, we visually inspected the development of the network of words, in particular, for each conversation turn. Notable differences were found between the two groups, including differences between the pivotal conversation turns within their discourses.

One major difference between the two groups was the cohesiveness of the networks. Although the groups were solving the same problem, their usage of conceptual words during conversational turns was considerably different. The network structure of conceptual words for the Gillian group was cohesive; while one word was isolated, the others were clustered into a large structure (Fig. 7). This clustering suggests that the learners related these conceptual words during their discussion. In contrast, the network of words in the Matt group was segmented; the network consisted of two completely separate clusters of words (Fig. 8). This result suggests that learners in the Matt group were not having a cohesive discussion at the conceptual level of problem solving.

Points that were pivotal for the cohesiveness of network were found in the Gillian group discourse. Through observation of changes in the network for each conversation turn, it was found that segmented clusters of conceptual words merged as one cohesive cluster in a specific conversation turn. At the pivotal point when the network structure became

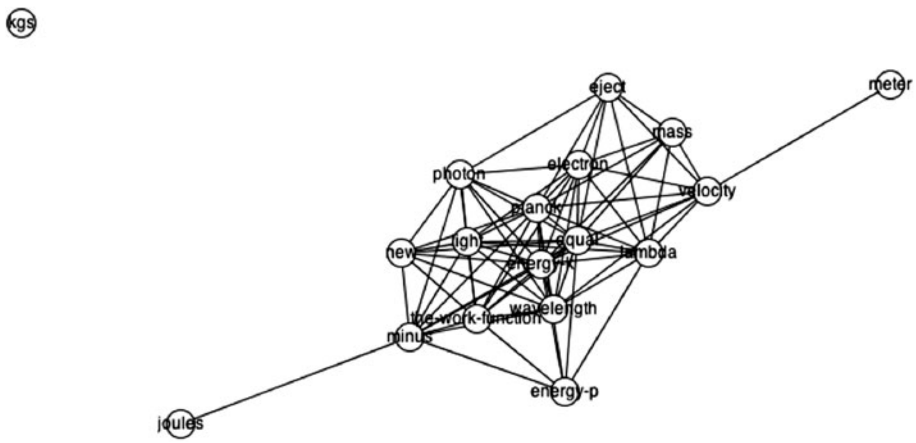

Fig. 7 The network of conceptual words for the Gillian group

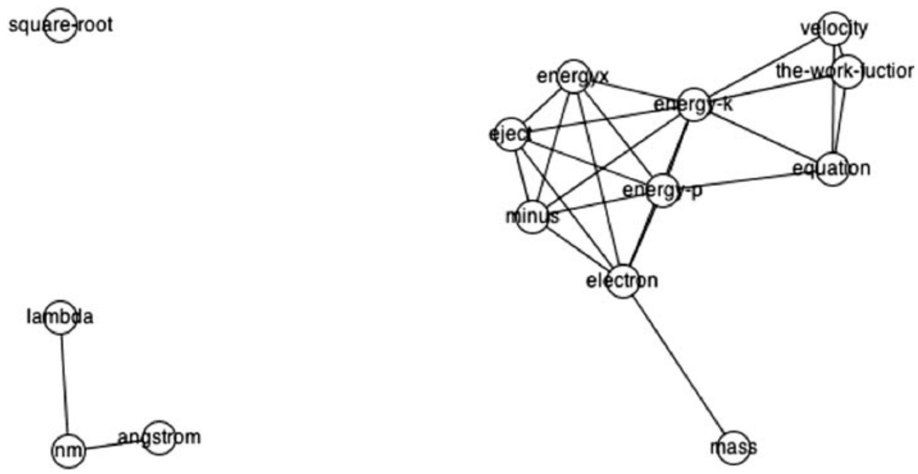

Fig. 8 The network of conceptual words for the Matt group

cohesive, a student suggested relating the different formulas at the planning stage, before actually calculating the answer by using the formulas (the last conversation turn in Table 1).

Analysis of network characteristics by using coefficients

Using the same data as above, we now demonstrate the analysis of discourse in collaborative learning by using centrality coefficients. The aim in this section is to establish an indicator for social knowledge advancement in collaborative learning based on the knowledge building community model. In particular, one aspect of social knowledge advancement is idea improvement (Scardamalia 2002). Under the basic assumption that a learner's idea can be represented as a cluster of linked words in a network, we attempt to illustrate how learners' ideas develop by observing transitions in the degree centrality coefficients.

Figure 9 shows how the cumulative degree centrality (the summation of the degree centrality coefficients) of conceptual words changed over conversational turns. As discussed in the visualization analysis, a notable difference was found in the cohesiveness of the network structure between the two groups. The final cumulative degree centrality was much higher for the Gillian group than for the Matt group. The following characteristics are seen in the transition of the cumulative degree centrality for the Gillian group: (a) after reaching a high value in the initial phase of their discourse, the cumulative degree centrality was sustained at a high level; and (b) there were no large fluctuations in the coefficients. Conversely, for the Matt group, it was found that (a) the cumulative degree centrality fell steeply in the final phase and (b) there were a number of large fluctuations in the coefficients. The gradual increase in total value of the coefficients and the sustainment of the values suggest that the Gillian group was consistently involved in knowledge integration through convergence of their ideas. The large fluctuations in the coefficients for the Matt group are evidence that group members were using assorted pieces of ideas in different segments of their discourse and were not able to relate their ideas to one another.

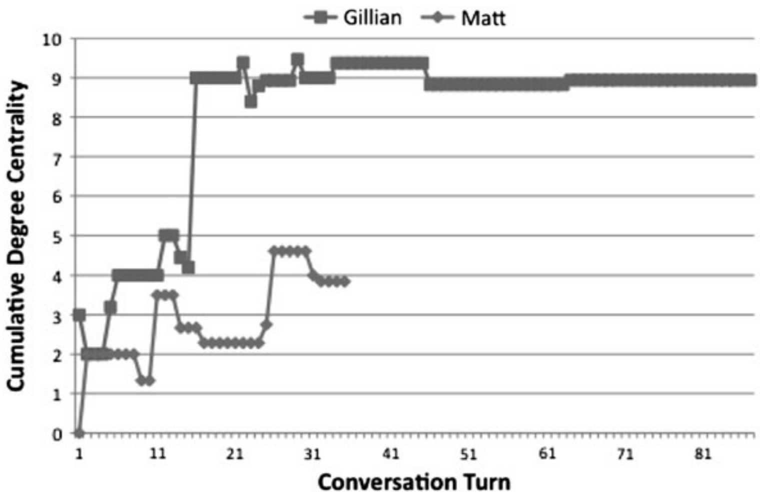

Fig. 9 Transition of the cumulative degree centrality for Gillian group and Matt group

Stepwise analysis

To evaluate each learner's contribution to the group discourse, stepwise analyses were conducted. Figures 10 and 11 show averaged absolute differences of the centrality coefficients calculated for the network of words for each learner in the two groups. For the two groups, 3 (Learners) $\times$ 3 (Centralities) ANOVAs with the averaged absolute difference as the dependent variable showed that the averaged differences of the closeness and degree centrality coefficients for student A1 in the Gillian group were significantly higher than those for the two other students: $F(2, 51) = 7.52$, $p < 0.01$ and $F(2, 51) = 8.01$, $p < 0.01$, respectively; however, no significant differences were found for the Matt group. Although a direct comparison of the coefficients between the groups was not made, the results suggest that learners in the Gillian group contributed more to the building of their network structure. Furthermore, the significantly higher averaged absolute differences for student A1 in the Gillian group show that student's unique contribution to the group discourse,

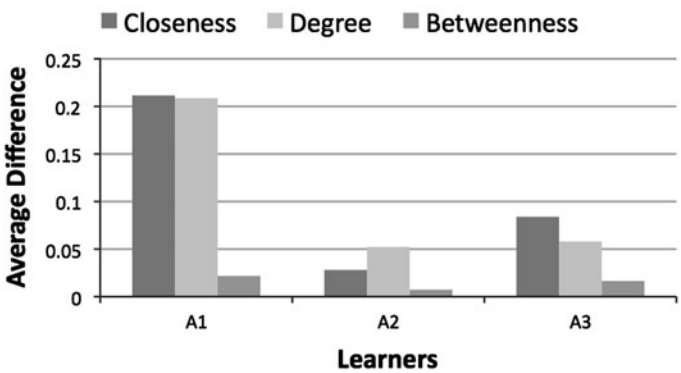

Fig. 10 Averaged absolute differences of centrality coefficients for learners in the Gillian group

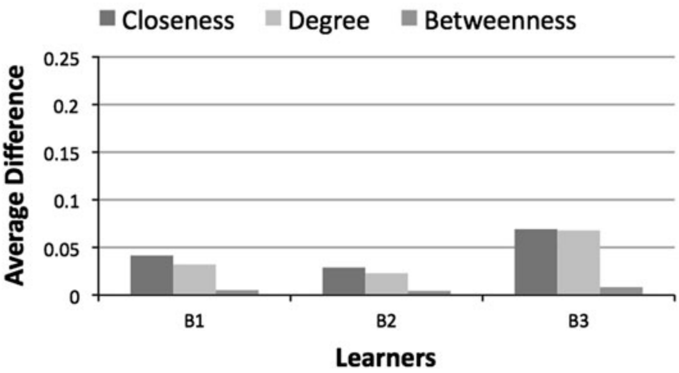

Fig. 11 Averaged absolute differences of centrality coefficients for learners in the Matt group

such as linking and summarizing small clusters of ideas. Thus, individual contributions to the network, which represent the state of the learners' social knowledge, could be evaluated by breaking down the discourse datasets for each individual.

Discussion

SNA using KBDeX for in-depth discourse analysis

The purpose of this study was to examine whether SNA supported by an application for discourse analysis, such as KBDeX, could provide a new approach to assessing learning processes in the knowledge-creation metaphor. Even though the datasets used in this study were small, it was possible to discuss the potential of this proposed approach by analyzing datasets for which in-depth qualitative discourse analysis had been already conducted (Sawyer et al., in press). Results from an SNA using KBDeX as a support tool revealed similar conclusions to the in-depth analysis, in that one group (Gillian) was more cognitively oriented to social knowledge advancement than the other (Matt). The SNA in this study further revealed the pivotal point in the students' discourse that facilitated this social knowledge advancement.

A unique contribution of SNA using KBDeX is the quantitative examination of each individual learner's contribution to social knowledge advancement. Stepwise analyses revealed that one student in particular in the Gillian group played a cognitively important role that the two other students did not. This relationship between social knowledge advancement in a group and individual contributions has not been simultaneously examined in previous studies. KBDeX's features of representing social knowledge advancement through visualization of three different types of networks (learners, discourse units, and selected words) and automatically calculating centrality coefficients (betweenness, degree, and closeness) for each learner's turn in a discourse should make unified analysis possible.

SNA as a new approach in the knowledge-creation metaphor: future research

As we have demonstrated, KBDeX opens up new possibilities for handling discourse data in order to evaluate learning from the perspective of social knowledge advancement.

Through visualization tools, KBDeX supports qualitative analysis by enabling investigation of how learners use words and how the words are connected to one another during discourse. The time course of changes in the network of words, for instance, may give insights into identifying how ideas emerge and are integrated into social knowledge. The analyst can also quantitatively examine the degree to which a network is coherently structured by calculating different centrality coefficients. The quantitative analysis tools in KBDeX are particularly advantageous when analyzing large and complex discourse datasets for which detection of network characteristics through visual inspection is challenging.

Beyond the attempt in this work to establish a unifying analytic framework in the knowledge-creation metaphor, further studies are under way. First, in collaboration with other research groups, we are continuing to investigate the validity of SNA supported by KBDeX for large discourse datasets. In one of the studies (Oshima et al. 2011), we analyzed high school students' discourse across several months in a CSCL environment that had already been analyzed in depth by van Aalst (2009). To deal effectively with such large discourse datasets, it is necessary to determine suitable indicators of social knowledge advancement. The measures used in SNA are appropriate for analysts to discuss network characteristics; however, they are not sufficient for educational researchers to examine learning in the knowledge-creation metaphor. The approach that we took in the establishment of indicators was to create a number of potential indicators based on one of the innovative models in the knowledge-creation metaphor (i.e., the knowledge building community), to apply these indicators to discourse datasets, and to compare the results with those in previously conducted in-depth analyses.

Second, to analyze learning processes where groups of students discuss their ideas in well-structured tasks such as the physical chemistry problem considered in this work (i.e., calculation of a photon's wavelength in a certain state), further studies are needed to assess students' social knowledge advancement by comparing their learning processes with experts' problem solving processes and examining their own networks qualitatively and quantitatively. Such comparisons could improve the predictive validity of analysts' interpretations on students' social knowledge advancement. Unfortunately, there were no such benchmark discourse data in this work, and the interpretation of the SNA results was limited in this regard.

Third, we have not sufficiently discussed alternative SNA-based measures (other than the centrality measure) that can be used to examine learning processes related to different metaphors of learning or different outcomes of learning. In this study, we used only the centrality measure to detect a few patterns of productive learning processes related to the knowledge-creation metaphor. There are certainly many other measures that can be used to identify patterns that support and inhibit student learning. Further systematic analyses with larger data sets are needed to establish guidelines and/or theories to help researchers determine and select appropriate measures.

Lastly, we hope that discourse analysts can easily extend their analyses by integrating SNA of discourse data into their current approaches. KBDeX has been developed for this purpose, and the first version of KBDeX will soon be made publicly available via the Internet so that discourse analysts interested in SNA can access it. On the basis of feedback from KBDeX users, we will refine its usability and functions.

Acknowledgments This research was supported by a Grant-in-Aid for Scientific Research (B) and (A) to Jun Oshima.

References

Anderson, J. R., Greeno, J. G., Reder, L. M., & Simon, H. (2000). Perspectives on learning, thinking, and activity. *Educational Researcher, 29*(4), 11–13.

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational Researcher, 25*(4), 5–11.

Barabási, A.-L. (2005). Network theory: The emergence of the creative enterprise. *Science, 308*, 639–641.

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

Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2006). Motivation and cognitive engagement in learning environments. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences*. New York: Cambridge University Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Cañas, A. J., Hill, R., Carff, R., Suri, N., Lott, J., & Eskridge, T. (2004). CmapTools: A knowledge modeling and sharing environment. In A. J. Cañas, J. D. Novak, & F. M. González (Eds.), *Concept maps: Theory, methodology, technology. Proceedings of the first international conference on concept mapping* (pp. 125–133). Pamplona: Universidad Pública de Navarra.

Cobb, P., & Bowers, J. S. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher, 28*(2), 4–15.

Davis, B., & Sumara, D. J. (2006). *Complexity and education: Inquiries into learning, teaching, and research*. Mahwah, New Jersey: Lawrence Erlbaum.

de Laat, M., Lally, V., Lipponen, L., & Simons, R.-J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *International Journal of Computer-Supported Collaborative Learning, 2*(1), 87–103.

Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamäki (Eds.), *Perspectives on activity theory* (pp. 19–38). Cambridge: Cambridge University Press.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition & Instruction, 20*(4), 399–483.

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software—Practice and Experience, 21*(11), 1129–1164.

Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher, 26*(1), 5–17.

Greeno, J. G. (2006). Learning in activity. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 79–96). New York: Cambridge University Press.

Greeno, J. G., & The Middle-School Mathematics Through Applications Project Group. (1997). Theories and practices of thinking and learning to think. *American Journal of Education, 106*(1), 85–126.

Guimera, R., Uzzi, B., Spiro, J., & Nunes Amaral, L. A. (2005). Team assembly mechanisms determine collaboration network structure and team performance. *Science, 308*, 697–702.

Ifenthaler, D., Masduki, I., & Seel, N. M. (2011). The mystery of cognitive structure and how we can detect it: Tracking the development of cognitive structures over time. *Instructional Science, 39*, 41–61.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. -L., Brewer, D.,…Van Alstyne, M. (2009). *Computational Social Science, 323*, 721–723.

Martinez, A., Dimitriadis, Y., Rubia, B., Gomez, E., & de la Fuente, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions. *Computers & Education, 41*(4), 353–368.

Mercer, N. (2005). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics, 1*(2), 137–168.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

Oshima, J., Oshima, R., & Knowledge Forum Japan Research Group. (2007). Complex network theory approach to the assessment on collective knowledge advancement through scientific discourse in CSCL. In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of CSCL 2007* (pp. 563–565). Mahwah, NJ: Lawrence Erlbaum.

Oshima, J., Oshima, R., Matsuzawa, Y., van Aalst, J. C., & Chan, C. C. K. (2011, April). *Network structure analysis for knowledge building: A macroscopic view of collaborative learning discourse*. Paper presented at the annual meeting of educational research association, New Orleans, LA.

Oshima, J., Oshima, R., Murayama, I., Inagaki, S., Takenaka, M. Yamamoto, T.,…Nakayama, H. (2006). Knowledge-building activity structures in Japanese elementary science pedagogy. *The International Journal of Computer-Supported Collaborative Learning, 1*(2), 229–246.
Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of Educational Research, 74*(4), 557–576.
Palla, G., Barabási, A.-L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature, 446*, 664–667.
Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. M. (2008). Highly integrated model assessment technology and tools. In Kinshuk, Sampson, D. G., Spector, J. M., Isaias P., & Ifenthaler D. (Eds.), *Proceedings of the IADIS international conference on cognition and exploratory learning in the digital age* (pp. 18–28). Freiburg: IADIS.
Reffay, C., Teplovs, C., & Blondel, F.-M. (2011). Productive re-use of CSCL data and analytic tools to provide a new perspective on group cohesion. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Proceedings of CSCL 2011* (pp. 846–850). Mahwah, NJ: Lawrence Erlbaum.
Reuven, A., Zippy, E., Gilad, R., & Aviva, G. (2003). Network analysis of knowledge construction in asynchronous learning networks. *Journal of Asynchronous Learning Networks, 7*(3), 1–23.
Sawyer, R. K. (2006). Analyzing collaborative discourse. In K. R. Sawyer (Ed.), *The Cambridge handbook of the learning science* (pp. 187–204). New York: Cambridge University Press.
Sawyer, K., Frey, R., & Brown, P. (in press). Dataset description: Peer-led team learning in general chemistry. In D. Suthers, K. Lund, C. Rose, N. Law, C. Teplovs, & G. Dyke (Eds.), *Productive multivocality in the analysis of collaborative learning*.
Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Jones (Ed.), *Liberal education in a knowledge society* (pp. 67–98). Chicago, IL: Open Court.
Scardamalia, M. (2003). Knowledge building environments: Extending the limits of the possible in education and knowledge work. In A. DiStefano, K. E. Rudestam, & R. Silverman (Eds.), *Encyclopedia of distributed learning* (pp. 269–272). Thousand Oaks, CA: Sage Publications.
Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In J. W. Guthrie (Ed.), *Encyclopedia of education* (2nd ed., pp. 1370–1373). New York: Macmillan.
Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 97–118). New York: Cambridge University Press.
Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex Publishing Corporation.
Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4–13.
Strogatz, S. (2001). Exploring complex network. *Nature, 410*, 268–276.
van Aalst, J. (2009). Distinguishing knowledge sharing, knowledge construction, and knowledge creation discourse. *The International Journal of Computer-Supported Collaborative Learning, 4*(3), 259–287.
Watts, D. J. (2007). A twenty-first century science. *Nature, 445*, 489.
Wegerif, R., & Mercer, N. (1997). Using computer-based text analysis to integrate qualitative and quantitative methods in research on collaborative learning. *Language and Education, 11*, 271–286.
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*, 7–44.

Jun Oshima is a professor in the Faculty of Informatics at Shizuoka University, Japan. His interests include developing new analytic approaches to evaluating collective knowledge and aiding in its advancement.

Ritsuko Oshima is a professor in the Faculty of Informatics at Shizuoka University, Japan. Her main research topic is the development of learning environments for knowledge building.

Yoshiaki Matsuzawa is assistant professor in the Faculty of Informatics at Shizuoka University, Japan. His academic focus is on applying conceptual complexity to the evaluation of learning processes.