

**Canadian Journal of Learning and Technology / La
revue canadienne de l'apprentissage et de la
technologie, V35(1) Winter / hiver, 2009**

**Participation in Knowledge-Building Discourse: An Analysis of Online
Discussions in Mainstream and Honours Social Studies Courses**

Hui Niu

Jan van Aalst

Authors

Hui Niu is with the Canadian Council on Learning, Vancouver, BC. Correspondence regarding this article can be sent to: hniu@ccl-cca.ca

Jan van Aalst is with the Faculty of Education, at The University of Hong Kong.

Abstract

Questions about the suitability of cognitively-oriented instructional approaches for students of different academic levels are frequently raised by teachers and researchers. This study examined student participation in knowledge-building discourse in two implementations of a short inquiry unit focusing on environmental problems. Participants in each implementation consisted of students taking a mainstream or an honours version of a tenth grade social studies course. We retrieved data about students' actions in Knowledge Forum® (e.g., the number of notes created and the percentage of notes with links), and conducted a content analysis of the discourse by each collaborative group. We suggest the findings provide cause for optimism about the use of knowledge-building discourse across academic levels: there was moderate to strong evidence of knowledge building in both classes by Implementation 2. We end with suggestions for focusing online work more directly on knowledge building.

Résumé

Les enseignants et les chercheurs soulèvent fréquemment des questions quant au caractère approprié des approches pédagogiques cognitives pour les élèves de différents niveaux scolaires. La présente étude a examiné la participation des étudiants à la coélaboration des connaissances lors de la formation, à deux reprises, d'une unité d'enquête de courte durée axée sur les problèmes environnementaux. Pour chacun des deux essais, les participants étaient des élèves qui suivaient un programme d'études de dixième année, soit général, soit spécialisé en sciences sociales. Nous avons récupéré des données sur les actions des élèves dans le Knowledge Forum (par exemple, le nombre de notes créées et

le pourcentage de notes avec des liens) et nous avons analysé le contenu du discours de chaque groupe de collaboration. Nous pensons que les résultats incitent à l'optimisme et qu'il est possible de parler de coélaboration des connaissances entre les niveaux scolaires : des données probantes moyennement rigoureuses ou rigoureuses montrant la coélaboration des connaissances ont été obtenues dans les deux classes lors du deuxième essai. Nous concluons avec des suggestions pour orienter plus directement les travaux en ligne sur la coélaboration de connaissances.

Introduction

In the last two decades there has been much interest in collaborative inquiry as an educational goal (National Research Council [NRC], 1996), and a number of technology-enhanced approaches to collaborative inquiry have emerged. Some examples are CoVis Collaboratory Notebook (Edelson, Pea, & Gomez, 1996), Learning by Design (Kolodner et al., 2003), and Web-based Inquiry Science Environment (WISE; Linn & Hsi, 2000). *Knowledge building* shares certain features with these approaches, including emphasis on collaboration, metacognition, distributed expertise, and use of computer-supported inquiry. As elaborated below in the section entitled "knowledge building," its distinctiveness follows from the commitment to make processes of expertise and innovation prominent in school. In a class operating as a knowledge-building community, students are agents of their own learning, work toward goals of collective knowledge advances, and treat ideas as real things that can be improved by means of discourse (Bereiter, Scardamalia, Cassells, & Hewitt, 1997). Advocates for knowledge building assert that it fosters a host of 21st century skills.

Though the studies that have informed the development of knowledge building as an educational possibility have involved students with a wide range of interest, prior knowledge, and ability (Lamon, Secules, Petrosino, Hackett, Bransford, & Goldman, 1996; Scardamalia, Bereiter, & Lamon, 1994; McAuley, this volume), teachers frequently express concern about the suitability of knowledge building for the majority of the students they teach. They question whether students are capable of engaging in the high level of agency, cognition, and metacognition that are needed. For example, experienced teachers interviewed by van Aalst and Hill (2001) after a six-week exploration of knowledge building in an in-service course made objections such as "the biggest thing is actually having it [knowledge building] in a group of 30 students where you have everyone engaged and excited about it," and "only a few [students] participate ... because they are good at the language thing and thinking on their feet and thinking quickly and they have lots of prior knowledge." Teachers and researchers also express concern about how to fit knowledge building into the curriculum and what to do about misconceptions. As a result, knowledge building is often assumed to be suitable for accelerated courses and optional after-school activities, but not for the mainstream curriculum. This is true even after teachers participate in extended professional development in which they examine student work and question students and their teachers about their extended knowledge-building experiences.

These concerns are common to a wide range of educational innovations, including cooperative learning and instructional approaches designed to foster higher-order thinking. Zohar and Dori (2003) concluded that many teachers who have taken their workshops on higher-order thinking express “the belief that instruction of higher order thinking is an appropriate goal mainly for high-achieving students and that low-achieving students, who have trouble with mastering even basic facts, are unable to deal with tasks that require thinking skills” (p. 146). Fishman, Marx, Blumenfeld, Krajcik, and Soloway (2004) argue that more attention must be given to the scalability and sustainability of technology-enhanced educational innovations. If knowledge building is to become a perspective capable of transforming education, it is necessary to address scalability issues such as participation by students of varying academic levels.

The goal of this study was to examine participation in asynchronous online discourse as an aspect of knowledge building, with a view to understanding its scalability across courses differing in academic level. To this end, we analyzed server-log data and the content of students’ contributions to an online knowledge-building environment (Knowledge Forum®, see <http://www.knowledgeforum.com>) from two implementations of a short inquiry unit in which students investigated environmental problems. Each implementation involved a mainstream Grade 10 social studies course and an honours Grade10 social studies course taught concurrently by the same teacher. This arrangement made possible a quasi-experimental study in which the academic level of the course was an independent variable. Because the teacher was new to knowledge building and concerned about completing the curriculum he kept the inquiry unit short (three weeks). Thus, we would expect only a limited implementation of knowledge building, but one that may nonetheless provide a “starting place” where teachers can begin to explore knowledge-building pedagogy.

The study examined two kinds of questions in the context of two successive implementations of knowledge building: How do participation levels in the mainstream classes compare with those in the honours classes, and to what extent can we conclude students engage in knowledge-building discourse?

Conceptual Background

This section provides a brief description of Knowledge Building and the conceptual background of the study.

Knowledge Building

Knowledge Building is collaborative practice in a community analogous to scientific inquiry, in which participants work to advance the state of knowledge in the community (Bereiter, 2002; Scardamalia, 2002; Scardamalia & Bereiter, 2006). It involves question-driven inquiry and explanation-driven understanding in a progressive discourse (Hakkarainen, 2003). As knowledge builders, students assess the limits of knowledge in the community, develop and execute plans for

advancing this, and evaluate whether they are making progress. Such metacognitive tasks as planning and monitoring are often assumed to be the domain of teachers, but testimony from students and teachers involved in knowledge building suggests many students are capable of it (Bereiter et al., 1997). Another essential element of knowledge building is an epistemology in which ideas are treated as real objects that can be improved by means of discourse (Bereiter, 2002). In knowledge building, the main goal is not merely to understand the ideas of previous generations, but to transform them into powerful tools that the community uses to solve its problems. To do so, students critically examine sources of knowledge available to them, and propose, test, and evaluate their own ideas.

Though knowledge building involves many types of interactions, discourse in Knowledge Forum plays a fundamental role; it provides a reliable and permanent record of experiments, classroom activities, ideas and questions that can be used to review progress and to develop understanding at progressively more complex levels. However, as van Aalst (2006) points out, the online discourse needs to be much more than online versions of “conversations.” Students need to do considerable work to structure the database. Such work includes reviewing the database, creating new links among ideas, and identifying progress and emerging lines of inquiry that need further attention. Often this work results in notes that link new ideas or interpretations to previous work, and Knowledge Forum has a number of features that are designed to support work on ideas after these are entered into the database (e.g., views, rise-above notes, and ability to link notes to other notes and views). According to van Aalst, students need to think of their work in the Knowledge Forum database as building a communal learning resource that has lasting utility rather than as online conversations. By contrast, teachers most often use asynchronous environments to promote the sharing, discussion, or debate of ideas. While such uses can lead to cognitive benefits (e.g., Baker, 2003; Bell, 2004; Fjermestad, Hiltz, & Zhang, 2005), they rarely include the idea improvement as conceptualized in knowledge building (Bereiter, 2002). Participants commonly regard online discourse as discussions, and frequently say they prefer face-to-face discussions to online discussions (van Aalst & Hill, 2001). Studies indicate that the vast majority of discussion threads do not exceed more than a few contributions (Guzdial & Turns, 2000; Hewitt, 2003, 2005). Scardamalia and Bereiter (2007) propose a shift from discussion to discourse that “aims to rise above the initial knowledge and belief state of the participants” (p. 206).

Analyzing Participation in Knowledge-Building Discourse

We contrast two perspectives on participation in asynchronous online discourse: one focusing on individual students’ actions in the online environment and one focusing on the identification of evidence for emergent and collective phenomena within the discourse of a community.

Individual students’ actions in the online environment can be analyzed using server-log data—data obtained from the server of the online environment about

such variables as the number of notes created and read by individual students. Knowledge Forum's Analytic Toolkit (Burtis, 1998) provides such data, and is used by teachers and students to retrieve and reflect on information about their online discourse. From a teacher's perspective such information can be useful for assessing participation levels. For example, whereas writing in online environments can have cognitive benefits, such benefits are not available to students who do not contribute or read notes. In addition, the nature of individual students' contributions to online discourse can be examined by using rating scales, for example scales focusing on the epistemic level of questions or explanations (Hakkarainen, 2003; Hakkarainen, Lipponen, & Järvelä, 2002) or the scientific validity of explanations (Zhang, Scardamalia, Lamon, Messina, & Reeve, 2007). Taken together, such analyses can probe both the quantity and quality of individual students' contributions.

Analysis of individual actions gives an incomplete picture. As many authors have pointed out, the actions in a discourse are mutually dependent (Sawyer, 2006; Stahl, 2002; Wells, 1999). For example, when students are asked to write notes to summarize what they have learned from their discourse, some comment that others have already stated their most salient learning and that they therefore do not state it again. Stahl (2002) points out that the meaning of the contribution to the database of an idea as a response to a question can easily be lost when the idea is analyzed out of context. Thus, although analysis of contributions by individual students is useful it is not sufficient for examining whether there is evidence for knowledge building in the discourse. As Sawyer (2006) comments, "knowledge and learning are often properties of groups, not only individuals" (p. 191).

A perspective on participation that can balance focus on individual students is to conceptualize participation at the group or community level. Thus, students can be said to participate in knowledge building if as a group they engage with the "life forms" of a knowledge-building community (also see Roth & Tobin, 2002, p. 157). For example, there must be evidence of discourse that examines the current state of knowledge in the community and that elaborates problems that are most promising for advancing this—"cutting edge" problems. Similarly, there needs to be evidence within the community for progressive problem solving (Bereiter & Scardamalia, 1993). Although not all students necessarily pose problems that are considered cutting edge by the community, all students work from the premise that it is important to advance the state of knowledge in the community. Similarly, not every important problem requires multiple layers of inquiry characteristic of progressive problem solving, but there must be sufficient evidence that progressive problem solving is a core value of the community. Therefore we propose that besides examining individual students' actions, it is necessary to examine knowledge-building phenomena using the community (or collaborative group) as the unit of analysis. In a knowledge-building community the actions by individual students need to occur in such a way as to produce emergent and collective effects. Bereiter (2002) suggests this is a question of self-organization rather than introducing rules. In this study, we do not focus on the mechanism by which collective effects are achieved but ask whether individual students' actions can be

said to produce discourse that has the qualitative features of knowledge-building discourse.

Variables that Influence Participation

A wide variety of variables can be expected to influence individual differences in participation in knowledge-building discourse, including prior domain knowledge, motivation, goal orientation, writing apprehension, epistemological beliefs, and ability to analyze arguments. However, in classroom studies it is often infeasible to measure all these variables. In this section we briefly describe the importance of two variables that we were able to measure for all students in this study: writing apprehension and ability to self-assess the contribution a computer note makes to the discourse.

Writing apprehension reflects a student's attitude and emotion towards writing tasks and written communication. According to Brand (1986), the role of emotion in writing processes is important to study, because the affective and cognitive components of composing are interrelated. This association of cognition with feeling is known as "hot" cognition (Abelson, 1963). In knowledge building, students write in an asynchronous environment to contribute and communicate ideas. Their writing contains information about their opinions, preferences, and evaluations. What they write is available for everyone to see and critique, which for some could create apprehension, causing them to write little and avoid spontaneity and sophisticated language (Faigley, Daly, & Witte, 1981). Using a questionnaire to measure writing apprehension, Daly (1978) found in a large study ($n = 3602$) that it influenced both the quantity and quality of what students write. Writing apprehension may thus have a negative impact on knowledge building and should be examined.

Ability to *reflect* on discourse is also important to knowledge building, especially for evaluating the progress of a line of inquiry and for setting communal learning goals. Making use of the ability in Knowledge Forum to link notes to other notes, van Aalst and Chan (2001) asked graduate students taking a course on knowledge building to create electronic portfolios based on their knowledge-building discourse. A portfolio consisted of a note in Knowledge Forum in which a student summarized evidence in support of four phrases describing knowledge building (*working at the cutting edge, progressive problem solving, collaborative effort, and identifying high points*), with hyperlinks to the notes used as evidence. The four phrases were based on Bereiter and Scardamalia's (1993) identification of progressive problem solving as essential to expertise, and exploratory studies of reflection and collaboration in CSILE (Computer supported Intentional Learning Environments) early in the 1990s. In subsequent studies with secondary school students in Hong Kong, students used these phrases not only for *describing* prior knowledge-building discourse, but also for guiding future contributions to it (Chan & van Aalst, 2003; Lee, Chan, & van Aalst, 2006; van Aalst & Chan, 2007). These studies revealed positive correlations between teacher scores of the portfolio narratives (measuring depth of reflection) and conceptual understanding as measured by essays and

conceptual questions from government examinations (van Aalst & Chan, 2007). In this article, which appears as part of a special issue on Knowledge Building, we refer to these four phrases as the *van Aalst-Chan-Lee (ACL) Principles* (ACL Principles for short) to distinguish them from the Scardamalia-Bereiter Principles (Scardamalia 2002; Scardamalia & Bereiter, 2006).

The Study

This study examined participation in online discussions in the context of two successive implementations of a three-week inquiry unit; each implementation involved a mainstream and an honours version of a tenth grade social studies course and the use of Knowledge Forum as the online knowledge-building environment. The research questions were:

1. To what extent do students in mainstream and honours social studies courses participate in online discussions?
2. To what extent can the online discussions in both academic levels be characterized as knowledge-building discourse?

The first research question is closest to teacher concerns about gaps in participation between classes at the same grade level but different academic levels. Placement in mainstream and honours courses is determined by a wide range of psychological and social variables — not just ability. However, the distinction is important to the organization of schools; teachers seem to think differently about such issues as expected outcomes and student agency depending on the academic level of the course. The second research question examined the qualitative features of these discussions in light of Knowledge-Building theory.

Methodology

Implementation 1 was studied in “post-hoc” fashion, in part in response to teacher questions about scalability in the context of a larger project. As a result, we were able to study only the Knowledge Forum database for this implementation. With preliminary findings from this analysis, we designed Implementation 2 with the teacher and were able to analyze additional data. We administered an Epistemology Questionnaire (Conley, Pintrich, Vekiri, & Harrison, 2004) and Writing Apprehension Test (Daly, 1978) at the beginning of the inquiry unit. However, although we found a moderate difference between the mainstream and honours class for the Epistemology Questionnaire, there was little variance in the scores and we excluded the data from the study (see Niu, 2006). We also analyzed a portfolio task assigned by the teacher at the end of the inquiry unit to measure students’ ability to assess their contributions to the online discourse by identifying the high points.

As suggested in the conceptual background section, individual actions in online discussions are not statistically independent. To compensate for this, we used a more stringent alpha level of .01 for statistical tests involving server-log data from individual students (Stevens, 2002).

Setting and Participants

The school was located in a suburban area in British Columbia (Canada). According to statistics released by the British Columbia Ministry of Education (2005), the school had typical demographics for British Columbia, except that the proportion of students from homes where English was not the first language was high (48%, compared with 20% for the province), and the educational level of adults in the community was high (81% graduated from high school, compared to 68% for the province). It was a relatively new school with more than 1500 students. At the time of this study, the school and the community in which it was located were ethnically diverse. More than 300 students were enrolled in the English as a Second Language (ESL) program; 50% of the students were born outside of Canada. Major ethnic groups within the school included Persian, Chinese, Korean, and Canadian. The majority of students in the school were well motivated academically, and many were expected by their families to attend university following graduation. Parents of the students had high aspirations for their children and were supportive of the school's programs.

The teacher had nine years of teaching experience and had a Master's degree focusing on cognitive strategy instruction. Prior to the study, he had implemented learner-centered approaches, using goal-driven planning, emphasizing active learning and interactivity, and giving frequent support and feedback. The teacher participated in several workshops on knowledge building in which he was introduced to Knowledge-Building principles; between the two implementations, a delegation of students from Implementation 1 and he presented an informal analysis of their discourse at a Knowledge Building Institute held at a local university; he also read extensively from literature on knowledge-building and participated in a series of "virtual meetings" focusing on knowledge-building pedagogy sponsored by the Institute for Knowledge Innovation and Technology (IKIT, www.ikit.org). During an interview conducted after Implementation 1 of the inquiry project, he said that his motivation for exploring knowledge building in social studies was to have students examine "how they gain knowledge and research skills, and what they can do to process the information to deepen understanding."

In Implementation 1, the participants were 28 students (13 male and 15 female) taking a mainstream version of a tenth grade social studies course and 30 (15 male and 15 female) students taking an honours version. All students were new to knowledge building. In Implementation 2, there were 30 students (15 males and 15 females) in the mainstream class and 26 (eight males and 18 females) in the honours class, with similar demographics; in the mainstream class, one student had prior experience with Knowledge Forum, and in the honours class, five.

Curriculum and Procedures

The course consisted of five interrelated curriculum organizers that reflected the multidisciplinary nature of social studies: applications of social studies, society and culture, politics and law, economy and technology, and environment. The

instructional unit in which knowledge building was implemented focused on environmental studies. According to the prescribed curriculum, students were expected to learn geographical skills and apply them to enhance their understanding of natural environments, and apply this understanding to such areas as resource development, stewardship, and sustainability. We describe here aspects of the instructional design common to the two implementations; later in the paper we describe some changes in the procedures from Implementation 1 to Implementation 2.

The teacher framed the unit in both implementations by introducing a set of general environmental problems such as deforestation and pine beetle infestation, which he described as “real to the students’ everyday lives.” In an interview conducted after Implementation 1 he said, “None of these environmental problems had a generally accepted viable solution.” Students were expected to elaborate more specific problems starting from these general problems. To limit the amount of writing students would encounter in the online environment in a large class, students in each class collaborated in groups of approximately eight to investigate one of the problems; students joined a group based on interest in a general environmental problem, with the condition that all groups should have seven to eight students. The students refined the problem definitions; identified and studied relevant background documents, including the textbook and online resources such as government websites; discussed these; and made recommendations for what they thought should be done about the problems.

At the beginning of the unit, the teacher opened two Knowledge Forum databases (version 4.5) to support these inquiries, one for each class. He demonstrated several basic features of Knowledge Forum including how to create a note, how to respond to a note, and how to navigate between discussion areas (called “views” in Knowledge Forum) in the school computer lab. Subsequently, students worked on Knowledge Forum one full class (70 minutes) per week in the computer lab, and also worked on Knowledge Forum at home.

The teacher designed the collaborative work to proceed in several phases: (a) showing the area of concern on a world map; (b) identifying the problem with historical and current information; (c) identifying causes, consequences, and solutions to the problem; (d) and explaining difficulties one might face in implementing a proposed solution. He had used this design for several years, and now created a view in Knowledge Forum for each phase. After setting up the database this way, the teacher did not systematically analyze the discussions or comment on them in Knowledge Forum. However, he regularly read notes when students had them open during class and asked students if they were making progress or needed assistance.

Participation in Knowledge Forum was not included in the formal assessment scheme for the unit. Instead, students were required to individually create portfolios using several of their own notes as artifacts; these were assigned at the end of the unit. Students were asked to identify two to three of their own notes and explain

why they considered these notes as exemplary knowledge-building contributions. Thus, we may assume that students' productivity in Knowledge Forum was not influenced by the need to meet a quota for note creation and reading. When students began preparing their portfolios, the teacher related the topics of investigation to the prescribed learning outcomes provided by the Ministry of Education to provide synthesis across the work by different groups.

In an interview conducted after Implementation 1 of this inquiry unit, the teacher said he felt the unit was an appropriate way to address the prescribed curriculum: It covered the required geography topics and research skills, and involved work with information technology and library resources. He felt that although he provided general problems, the students still needed to articulate these more fully and they were open-ended; in his view, there also was opportunity to see a range of viewpoints discussed and opportunity for all participants to contribute to the discussions.

From the researchers' perspective the design was expected to enable only a limited form of knowledge building. Perhaps the most significant limitation was that the course commenced only a few weeks before the inquiry unit, thus there was little time for students to develop as a community and to acquire values and practices conducive to knowledge building; the decision to assign the students to groups (thought necessary by the researchers in a large class) also limited community development. In addition, three weeks seemed short for observing emergent knowledge-building phenomena such as progressive problem solving and the articulation of general principles from the solutions proposed by the various groups. Thus, this study examines knowledge building in collaborative groups and during relatively short periods of time. Despite these limitations and given the constraints of the curriculum and the risks involved in embarking on a new direction for the teacher, the instructional design appeared to be a reasonable one with which a teacher could begin to explore knowledge building in the classroom.

Measures and Analysis

Server-log data

To address the first research question the Analytic Toolkit was used to retrieve and analyze summary statistics on individual students' activity in Knowledge Forum. Version v4.6 provides up to 27 analyses of how students interact with each other in the Knowledge Forum database. We selected the analyses that have been used most frequently in prior studies. These analyses yielded five measures, to which we refer as "Analytic Toolkit indices:"

Notes Created: This is a productivity measure. Building a database minimally requires writing notes; each note represents at least one thought or information unit. Previous studies suggest that the amount of writing is correlated with depth of explanation (Hakkarainen et al., 2002) and with gains in basic literacy (Scardamalia et al., 1994).

Percentage of Notes Read: This also is a general productivity measure: the total number of notes opened as a percentage of the total number of notes in the database. Opening a note does not imply that it is read carefully, but one cannot read a note without opening it. A low read level would suggest a low level of familiarity with the content of the database, especially if students do not meet face to face regularly to discuss their collaborative inquiry.

Percentage of Notes with Links: This measure is the percentage of notes that respond to, quote, or reference another note. Such linkages among notes produce networks with notes, in which links signify relationships between the notes (e.g., direct response versus the use of one note as a reference in another note). In other words, this measure is not a strict productivity measure but a measure of how students make contributions to the database. A database with a high percentage of linked notes indicates that students are attempting to relate their ideas to ideas already represented in the database. This process is essential to improving the community's ideas. Links provide multiple pathways to ideas.

Note Revision: An important notion in knowledge building is that ideas are seen as improvable objects (Bereiter, 2002). If students treat ideas as improvable, one way this may be evident in a database is through a high number of note revisions. Of course note revision is not the only way idea improvement can be evident in the database, and "revisions" captured by the Analytic Toolkit are not necessarily substantive.

Scaffold Use: Scaffolds are metacognitive prompts that guide knowledge construction. The inclusion of scaffolds in notes is an effort to make the database more useful as a knowledge-building resource, because scaffolds can be used to search the database and assist the members in maintaining focus on theory building.

Writing Apprehension Test and Portfolio Task (Implementation 2)

In Implementation 2, The Writing Apprehension Test (Daly, 1978) was administered at the start of the inquiry unit. It consists of 20 Likert scale items, including the following: "My mind seems to go blank when I start to work on a composition," "I would enjoy giving my writing to magazines for evaluation and publication," and "Discussing my writing with others is an enjoyable experience." Such items reflect the extent of anxiety students have when faced with a writing task. Although some research has shown that some anxious writers are good writers (Bloom, 1980), most researchers agree that the writing apprehension test is an accurate tool in surveying writing apprehension (Reed, Burton, & Vandett, 1988). The scale reliability for the questionnaires completed by the participants was .92 (Cronbach alpha).

At the end of Implementation 2, participants in both classes individually completed a (paper-format) portfolio task designed by the teacher. Students were asked to

identify two to three of their own notes and explain why they considered these notes as exemplary knowledge building contributions. Each note was accompanied by an explanation of its function in knowledge-building discussions from three perspectives: content, context, and role. By *content*, students showed evidence of their learning process and understanding of domain knowledge. Students were asked to demonstrate the evolution of their understanding from earlier notes to later notes in the portfolio. By context, students provided explanations of how each selected note helped to build the class's collective knowledge. This required that students explained how their notes functioned within the discussion by placing the notes in the context of the thread in which they appeared. By role, students were asked to describe the roles their notes played in the discussion; explain how the notes clarified, elaborated, or extended the discussion; and/or provide a new way of looking at the issue under discussion. The portfolio also included a summary paragraph, in which students wrote about what they had learned and whether they found Knowledge Forum effective for learning about environmental issues.

The portfolio task was assigned by the teacher as a summative assessment of what students had learned from their inquiry. However, we suggest that it is better interpreted as a measure of students' abilities to reflect on and summarize their contributions to the database at the time of their inquiry. For example, with respect to "content," the main issue addressed by the scoring was not the correctness of claims relating to domain knowledge, but whether students were able to formulate cogent arguments about the evolution of their ideas. It is unlikely that the abilities measured by this task change significantly as a result of such a short inquiry unit, so we treated the measure as an independent variable. Though it is a messy measure, we suggest that it provides a useful indicator of a cognitive and metacognitive performance relevant to knowledge building.

All 46 portfolios were rated by the teacher, using a marking scheme he designed. A second social studies teacher was trained on a small set of portfolios from Implementation 1 (a few that had not been collected by their authors), and then independently rated all the portfolios for Implementation 2, using the same evaluation criteria used by the teacher. She was not given any information about the class from which a portfolio came. The inter-rater reliability was 0.66 (Pearson correlation). Although low, this is not unusually low for portfolios (Koretz, Stecher, Klein, & McCaffrey, 1994). See Appendix D in Niu (2006) for a detailed analysis of the reliability of the coding of the portfolios.

Content analysis of knowledge-building discourse

We conducted content analysis of the online discourse, using all the writing by a collaborative group—the *group discourse*—as the unit of analysis. Although the classroom work was framed in terms of a broad set of Knowledge-Building principles (see MacKinnon, McAuley, this volume), five were used for analysis because we did not achieve high inter-rater reliability with the larger set using Law and Wong (2003) procedures. Therefore, we used the same set as in a recent knowledge-building portfolio study (Lee et al., 2006): with one addition—*constructive use of authoritative sources*. This set provides a good lens for

examining major features of knowledge building. Since this research, Zhang et al. (2007) analyzed knowledge-building discourse by a single class of Grade 4 students. In addition to constructive use of authoritative sources, they used *idea improvement*, *real ideas/authentic problems*, and *collective responsibility/community knowledge*. This set provides similar coverage of knowledge building as the set we used. Below, we summarize the five principles we used; the most relevant Scardamalia (2002) Principles are stated in parentheses.

Working at the cutting edge. A scholarly community works to advance its collective knowledge. For example, scientists do not only work on problems of personal interest, but on problems that can contribute something new to a field. Such problems may emerge from conflicting models, theories, and findings that require further explanation. By “working at the cutting edge” we mean that there is a community value to advance the state of knowledge, which produces an advancing knowledge frontier or cutting edge. Indications of this include evidence that students propose problems that can advance the state of knowledge in the community and evidence that the community takes up such problems. (*Collective responsibility, community knowledge; epistemic agency; real ideas, authentic problems*)

Progressive problem solving. When an expert understands a problem at one level, he or she reinvests learning resources into new learning (Bereiter & Scardamalia, 1993). In a scholarly community, we often find one study raises new questions that are explored in follow-up studies. Indicators of progressive problem solving in computer discourse would include instances when students have solved certain problems but then reinvest their efforts in formulating and inquiring into other problems for deeper understanding. Often students document the history of the problem and mark the progress of the idea. (*Improvable ideas; rise above*)

Collaborative effort. We consider collaboration as joint activity aimed at shared understanding (Dillenbourg, 1999); “collaborative effort” then is the effort students make to help others understand ideas. Besides writing and responding to notes it can include service to the community, in which students synthesize a line of inquiry and integrate perspectives; students may also add keywords to notes and link notes to make it easier to locate ideas. Collaborative effort can be found in many types of learning communities; in a knowledge-building community it needs to be directed at advancing the state of knowledge. (*Collective responsibility, community knowledge; idea diversity; democratizing knowledge*)

Identifying high points. Knowledge building requires metacognitive understanding. Specifically, students need to have insight into their own and the community’s knowledge advancement processes. Whereas the principle progressive problem solving can be used to examine the history of problems in the community, this principle focuses on students’ personal insight. For

example, students may identify events that help them understand something differently. (*Epistemic agency; rise above*)

Constructive uses of authoritative sources (Scardamalia, 2002; Scardamalia & Bereiter, 2006). This principle highlights the importance of keeping in touch with the present state and growing edge of knowledge in the field. Whereas it is commonplace for students to refer to the Internet or websites, knowledge building emphasizes the constructive and evaluative uses of resources in scientific inquiry. Indicators in the computer discourse include students identifying inconsistencies and gaps in knowledge sources and using resources effectively for extending communal understanding.

Each group discourse was first separated into excerpts that provided evidence for at least one principle, keeping enough surrounding text to keep a sense of the context of the episode. We developed a four-point rating scale for each principle, based on a set of guidelines we had provided to students in prior research to help them select evidence in support of the principles when preparing knowledge-building portfolios (Study 2, van Aalst & Chan, 2007). The score reported is the average of the scores obtained from all the excerpts within a group discourse.

For *Working at the cutting edge*, a discourse excerpt receiving a score of “4” would provide strong evidence that the group identified gaps in collective understanding and posed questions with potential for closing such gaps; a discourse excerpt receiving a score of “3” would have less convincing but still adequate evidence for these things. A score of “2” would be assigned to a discourse excerpt in which members asked questions that were somewhat relevant towards extending the community’s knowledge, but were rarely taken up by the majority of group members. Finally, a score of “1” would result from members’ not working on identifying gaps in community’s knowledge, or asking questions that did not necessarily extend the community’s understanding.

In total 15 group discourses were rated by the researcher (the first author). A second rater (a graduate student with extensive background in knowledge building) independently rated seven of the group discourses. Inter-rater reliability was established by calculating the Pearson correlation coefficient between the average scores assigned by the two independent raters (based on 35 ratings – seven group discourses × 5 principles), leading to an inter-rater reliability of 0.80. The appendix provides examples of ratings.

Results: Implementation 1

Server-log Data

The analysis in this section addresses the first research question using server-log data obtained with the Analytical Toolkit. Multivariate analyses of variance (ANOVA) of these data were conducted, using class membership as the independent variable.

The mainstream class wrote 327 notes and the honours class 623. Of the 52 discussion threads created by the mainstream class, 43 (83%) had fewer than 6 notes; this percentage was smaller for the honours class (56%). These general features of the databases are consistent with the teacher's impressions about participation levels.

Table 1 shows means and standard deviations for the Analytical Toolkit indices for the two classes. For convenience of presentation, the results for Implementation 2 are shown in the same table. In Implementation 1, the honours class had larger means for Notes Created and Percentage of Notes with Links. However, the standard deviations for Percentage of Notes with Links were larger for the mainstream class than the honours class. Individual students in the mainstream class, on average, wrote 11.7 notes, or almost four notes per week; approximately one in two notes, was linked to at least one other note. The percentage of notes read seemed low, as did Scaffold Use and Note Revision. For example, both classes used scaffolds infrequently compared to the number of notes written: on average, students in the mainstream class used approximately one scaffold in two notes. The honours class used scaffolds less frequently than the mainstream class—one in three notes.

Table 1. Analytic Toolkit Indicators per Student for Mainstream Class and Honours Class

	Implementation 1				Implementation 2			
	Mainstream		Honours		Mainstream		Honours	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Notes Created	11.7*	11.1	20.8*	10.1	13.4	8.2	13.2	9.9
% Notes Read	17.5	19.2	21.0	10.8	12.9	8.2	12.1	5.8
% Notes with Links	41.9**	34.2	82.8**	16.2	41.9*	32.7	62.0*	22.8
Note Revision	4.0	5.5	4.0	3.5	3.3**	3.9	8.9**	9.2
Scaffold Use	5.8	10.2	3.2	3.5	1.2**	2.4	6.7**	7.5

* $p < .01$ ** $p < .001$

A multivariate analysis of variance (MANOVA) showed that the five Analytical

Toolkit indices significantly differentiated the two classes, $F(5, 52) = 14.13$, $p < .001$, Wilks' $\Lambda = .42$, $\eta^2 = .58$. Accompanying this overall effect was a significant univariate effect for Percentage of Notes with Links: 41.9% for the mainstream class, compared with 82.8% for the honours class, $F(1, 56) = 34.40$, $p < .001$, $\eta^2 = .38$. The difference for Notes Created was not significant at the .01 level, $F(1, 56) = 10.48$, $p < .05$, $\eta^2 = .16$.

Content Analysis

The databases were segmented into eight group discourses (the mainstream courses and the honours classes each had four groups). Each group discourse was then rated for the five Knowledge-Building principles, as explained in the methodology section. The mean ratings for each are shown in Table 2. For convenience of presentation, results for Implementation 2 are shown in the same table. In Implementation 1, the mean scores for the mainstream class were generally 0.5 less than for the honours class; the total score was 10.9 (54.5% of the maximum possible) for the mainstream class and 13.3 (66.5%) for the honours class. Observe that when the principles are ordered from the highest score to the lowest score, the same order is obtained for the two classes; for example, for both classes the evidence for collaborative effort was strongest and the evidence for identifying high points weakest. Though the mean scores were not high, they do suggest moderate evidence for participation by the groups in the mainstream class for three of the principles: *working at the cutting edge*, *progressive problem solving*, and *collaborative effort*. It is worth noting that while the scores were higher for the honours class, the data did not suggest *large* differences between the two classes compared with the large between-class differences for the server-log data. (Due to the small number of groups no statistical tests are done for the content analysis. The findings must therefore be interpreted with caution.)

Table 2. Scores for Five Principles for Analyzing Knowledge-Building Discourse

	Implementation 1				Implementation 2			
	Mainstream		Honours		Mainstream		Honours	
	<i>n</i> = 4		<i>n</i> = 4		<i>n</i> = 4		<i>n</i> = 3	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Working at the cutting edge	2.3	.5	2.8	.5	3.0	.8	3.0	.0
Progressive problem solving	2.3	.5	3.0	.0	3.0	.8	3.3	.6
Collaborative effort	2.5	.6	3.0	.0	2.5	.6	3.3	.6
Identifying high points	1.8	1.0	2.0	.8	2.5	.6	2.7	.6
Constructive use of authoritative sources	2.0	.0	2.5	.6	2.3	.5	2.7	.6
Total	10.9	2.6	13.3	1.9	13.3	3.3	15.0	2.4

Lessons Learned

In summary, in Implementation 1 there were large differences between the mainstream and honours classes for the Analytical Toolkit indices. The honours class wrote nearly twice as many notes as the mainstream class (but with much within-class variation), and the proportion of notes that were linked was also nearly twice that for the mainstream class. These effects are very observable in the various views in the databases, and consistent with the teacher's impression of the relative performance of the two classes. However, the content analysis suggests that these large differences were not accompanied by large differences in the qualitative evidence for knowledge building. Indeed, perusal of the database indicated that in many of the long threads students kept asking similar questions, there were few notes that responded to more than one previous note, and there was little evidence for *branching* in the threads (i.e., for sustained but emergent lines of inquiry).

Results: Implementation 2

The goal of Implementation 2 was to examine the consistency of the findings of Implementation 1 with new mainstream and honours classes, and to measure additional variables that could influence participation in knowledge building. The same teacher again taught mainstream and honours tenth grade social studies classes and used essentially the same instructional design. However, the teacher now had a deeper understanding of knowledge building and the researchers also asked for several minor changes in the procedures. In brief, the changes to the procedures were as follows. First, the teacher provided more similar training on Knowledge Forum prior to the start of the Inquiry unit. We had observed that in Implementation 1 the honours class had had a general discussion in a practice view but the mainstream class had not. Second, both classes now shared a database (but still with a specific area for each group); students were encouraged

to examine discourses by groups not in their own class. The rationale for this change was to effect “social comparison” (Festinger, 1954) leading to more equivalent participation for the two classes. (A minor amount of inter-class interaction occurred, which is excluded from our analysis.) Third, we asked the teacher to encourage the mainstream class more to contribute to the database.

Server-log Data

In Implementation 2, the mainstream class wrote 388 notes and the honours class 339. Of the 58 threads created by the mainstream class, 42 had fewer than 6 notes (72%); of the 54 threads created by honours class, 36 had fewer than 6 notes (67%). These statistics suggest that differences in participation between the classes were smaller than in Implementation 1. Note that the honours class created many fewer notes than in Implementation 1 (339, compared with 623).

Table 1 shows the means and standard deviations for the Analytic Toolkit indices for the mainstream and honours classes. In Implementation 2, the large differences between the classes for Notes Created were not reproduced. Instead, there were large differences for Note Revision and Scaffold Uses. However, these indicators were still low. For example, while in the honours class individual students on average used 6.7 scaffolds, this amounted to only one scaffold per two notes; if scaffolds are used consistently, one would expect approximately one scaffold use per note.

A MANOVA showed that the five Analytic Toolkit indices significantly differentiated the two classes, $F(5, 49) = 10.94$, $p < .001$, Wilks' $\Lambda = .47$, $\eta^2 = .53$. Accompanying this overall effect were small univariate effects for: Percentage of Notes with Links, $F(1, 53) = 6.86$, $p = .01$, $\eta^2 = .12$; Note Revision, $F(1, 53) = 8.86$, $p < .005$, $\eta^2 = .14$; and Scaffold Use, $F(1, 53) = 14.08$, $p < .001$, $\eta^2 = .21$.

Writing apprehension and portfolio scores

Means and standard deviations for the Writing Apprehension Test and Portfolio Task are reported together in Table 3. The results are similar for both measures, with the honours students outperforming the mainstream students with effect sizes of approximately 20% (η^2).

Table 3. Mean (SD) for Writing Apprehension Test and Portfolio Task

	Mainstream			Honours		
	n	M	SD	n	M	SD
Writing Apprehension Test	27	60.3	12.2	25	72.0	13.4
Portfolio Task	23	33.6	9.7	23	40.6	7.3

For the Writing Apprehension Test a higher score indicates less anxiety toward writing. Because in this study knowledge-building discourse is realized through written communication in Knowledge Forum, students' writing apprehension may impact their performance in the discussions in Knowledge Forum. A one-way ANOVA revealed that students in the honours class were statistically less anxious about writing than students in the mainstream class, $F(1, 50) = 10.7, p < .005, \eta^2 = .18$. This result appeared to indicate a general dislike of public writing. Evidence for this claim can be found by examining the items with the largest between-class differences. For example, mean scores for "I like to write my ideas down," "I like seeing my thoughts on paper," and "I would enjoy giving my writing to magazines for evaluation and publication," all had between-class differences of approximately one standard deviation.

The portfolio scores are assumed to probe reflection and summarization, as explained in the methodology section. The results show that students in the honours class outperformed students in the mainstream class. A one-way ANOVA showed this effect was statistically significant, $F(1, 44) = 7.6, p < .01, \eta^2 = .15$.

To investigate the influence of the writing apprehension on the significance levels for between-class comparisons of the Analytic Toolkit indices, we conducted a MANOVA with the Analytic Toolkit indices as dependent variables and the writing apprehension score as covariate. There were two changes in the results. First, the difference between the classes for Note Revision was no longer significant at alpha = .01 ($p = .054$). Presumably, if students did not like to write, they were less likely to return to a note to revise it. In addition, the difference between the classes for Percentage of Notes with Links was no longer significant ($p = .03$). The between-class effect for portfolio scores was no longer significant ($p = .104$) with writing apprehension as a covariate.

Correlations among measures

To examine relationships among the variables, the Analytic Toolkit measures were aggregated to create two general Analytic Toolkit measures. The first was the average of the z scores for Notes Created and Percentage of Notes Read, calculated using data from both classes; the second was obtained similarly from the remaining indices (Percentage of Notes with Links, Note Revisions, Scaffold Uses). The first score (Analytic Toolkit Productivity) is a measure of productivity in Knowledge Forum, such as may occur in a wide variety of online discussions; the second score (Analytic Toolkit Knowledge Building) is a measure of actions that are more specific to knowledge building. Pearson correlations for these two measures and the scores from the Writing Apprehension Test and Portfolio Task are shown in Table 4; the upper entry in a given cell is the correlation for the mainstream class, and the lower entry the correlation for the honours class.

Table 4. Pearson Correlation Coefficients

	Writing Apprehension Test	Portfolio Task	Analytic Toolkit (Productivity)
Portfolio Task	.48*		
Analytic Toolkit (Productivity)	-.06	.31	
Analytic Toolkit (Knowledge Building)	-.13	.35	.72**
	.38	.24	.82**

Note: Upper entries are results for the mainstream class and lower entries results for the honours class.

* $p < .05$ (2-tailed).

Results show that in the mainstream class writing apprehension accounted for 23% of the variance in the portfolio scores (r^2); in the honours class, writing apprehension accounted for 20% of the variance in productivity. For both classes, there were very strong correlations between Analytic Toolkit Productivity and Analytic Toolkit Knowledge Building ($r = .72$ and $.82$ respectively).

Content Analysis

Means and standard deviations for the five principles are shown in Table 2. As the table shows, the scores in Implementation 2 were higher than in Implementation 1 for both classes. The scores for the mainstream class in Implementation 2 were similar to those for the honours course in Implementation 1. The total score for the honours class was 15.0 in Implementation 2 (75%), which provides relatively strong evidence for knowledge-building discourse; this result was obtained despite a dramatic decrease in some of the Analytic Toolkit indices. It is also worth noting that in both classes mean scores of at least 3.0 (75%) for two Knowledge-Building principles: *working at the cutting edge* and *progressive problem solving*. We attribute these improvements to the teacher's learning and changes to the instructional design suggested by the researchers. The researchers considered that these two influences could not be separated.

Lessons learned

Several lessons can be drawn from this analysis. First, the differences between mainstream and honours classes for Notes Created and Percentage of Notes with Links were large in Implementation 1 but were much smaller in Implementation 2. Second, students in the honours class outperformed students in the mainstream

class on the Writing Apprehension Test and Portfolio Task, although these effects were not large (η^2 around 20%). Some of the remaining differences for the Analytic Toolkit scores were no longer significant when the Writing Apprehension Test was used as a covariate (Percentage of Notes with Links and Note Revision). This finding is consistent with anecdotal evidence from teachers and previous research suggesting that many students in mainstream courses resist contributing their ideas, which they often view as inadequate, to public discussions (Slater & van Aalst, 2002). Third, there appeared to be substantial improvements from Implementation 1 to Implementation 2 in the total scores in the content analysis. This finding suggests that teacher learning and changes to the instructional design and teacher action in response to formative evaluations (i.e., the analysis of Implementation 1) can be important in compensating for individual differences that influence knowledge building. Fourth, the scores from the content analysis were similar to those obtained with grade twelve students of above average achievement as well as graduate students (van Aalst & Chan, 2007). Thus, our data from this implementation suggest that students of average ability can participate in knowledge-building discourse.

Conclusions and Implications for Teaching and Research

Beliefs that only the “best and brightest” students can participate in and benefit from learning approaches that depend on knowledge construction and ability to evaluate one’s own knowledge are common among teachers and researchers (Zohar & Dori, 2003; van Aalst & Hill, 2001). Despite empirical studies of conceptual change and metacognition which reach a contrary conclusion (e.g., White & Fredericksen, 1998), these beliefs have a negative impact on the perceived scalability of cognitively-based instructional approaches. In the case of knowledge building, the beliefs can be reinforced by apparent differences in the databases created by students in courses of different academic levels. In this study, we examined aspects of this issue by analyzing server-log data representing individual actions in the online environment and analyzing the discourses of collaborating groups as evidence for the qualitative, collective, and emergent features of knowledge-building discourse in mainstream and honours social studies courses at the same grade level. We analyzed online discourse from four relatively large classes (by Canadian standards) totaling more than 1600 notes.

Our findings suggest there is cause for optimism about the use of knowledge-building discourse across academic levels. Perhaps the most important finding is that there appears to be little relationship between very high levels of note-writing and note-linking and the evidence for knowledge building from content analysis. In Implementation 1 such high levels for the honours class were not accompanied by strong evidence for knowledge building in the content analysis, and though the productivity measures were lower for the second honours class than for the first, the scores from the content analysis were higher. Another important finding was that the evidence for knowledge building improved from Implementation 1 to Implementation 2. It is impossible to separate the influence of the researchers’ requests for changes to the procedures from changes resulting from the teacher’s learning about knowledge building. Nevertheless, our data suggests that *it is*

problematic to judge the scalability of an instructional approach on early outcomes. From the perspective of scalability, a better question would be how many iterations of design, teaching, and formative evaluations are needed to establish consistent evidence for knowledge building.

Of course, there were some important differences between the classes of different academic levels. The scores for the content analysis were higher for the honours courses in both implementations (22% in Implementation 1 and 13% in Implementation 2). There also were significant differences between the classes in Implementation 2 for writing apprehension and the portfolio task, with effect sizes of 0.8 to 0.9 standard deviation favouring the honours class. Such differences raise an important question for classroom research: how do teachers deal with writing apprehension in facilitating knowledge building. With respect to the portfolio task, there is much research on cognitive strategy instruction teachers can use to attempt to close the gap (Bransford, Brown, & Cocking, 1999). Results also suggest that once knowledge building is integral to classroom processes, all students become engaged productively (Zhang et al., 2007). However, the question of scalability does not hinge on the *existence* of differences but on whether *in spite of them* students can participate in knowledge building and benefit from it. In this respect, we think the evidence for participation in knowledge-building discourse from the content analysis was relatively strong for both classes in Implementation 2. In this study we were unfortunately not able to examine growth in domain knowledge directly (i.e., the *outcomes* of the knowledge-building process).

It is important to understand why there was not a strong relationship between the server-log indices and the results of the content analysis of knowledge-building discourse in this study. The Analytic Toolkit is conceived by its developers as a tool that students and teachers can use to examine their own knowledge-building discourse with a view to improving it, and some prior studies did detect a relationship between Analytic Toolkit indices and evidence for knowledge building. Lee et al. (2006) conducted a study of four ninth grade geography classes in which students worked on Knowledge Forum throughout a semester; their study setting is perhaps the most similar of all studies with which we are familiar to the mainstream classes in the present study in terms of grade level, academic achievement, organization of online work in groups, available class time for online work, and frequencies of scaffold use and note revisions relative to the number of notes created. Lee et al also conducted protocol analysis of all the questions and explanations posted to the database. They found strong positive correlations among the number of high-level explanations and a measure derived from Analytic Toolkit indices (notes created, notes read, scaffold use, and note revisions), scores based on knowledge building, and a measure of conceptual understanding. Van Aalst and Chan (2007) obtained similar results with older and academically above-average students; in two of three implementations they report strong positive correlations between scaffold use and evidence for knowledge building based on the ACL Principles. These studies suggest that while productivity can contribute to knowledge building it needs to be productivity aimed at constructing explanations (or “theories”) that reveal understanding of the domain. There is evidence that

explanation-seeking discourse can enhance conceptual change (Chan, Burtis, & Bereiter, 1997; Hakkarainen, 2003). Scaffolds are a Knowledge Forum feature designed to focus student work on a knowledge-constructing discourse.

In summary, we propose that to evaluate the promise of knowledge-building discourse for students of wide-ranging academic achievement, teachers need to examine evidence that students are individually developing high-level explanations. Having a “lively” database with many notes and links between notes is neither sufficient nor necessarily helpful. It seems helpful to part with the notion of “online discussions” that supplement classroom activities and to conceptualize online work, as indicated earlier in this article, as collaborative and iterative work to build new explanations and ideas, in an effort to advance the state of knowledge in the community (Scardamalia & Bereiter, 2007; van Aalst, 2006). In this, the online environment is not a “discussion environment” but a “knowledge-building environment” with tools designed to support working with ideas after they have been entered.

In closing, it may be useful to point out some limitations of the study and its implications for further research. First, there were methodological difficulties that are common to research in classrooms; the study could be strengthened by additional instruments that could be used to measure relevant psychological variables and the growth of domain knowledge, and the inquiry unit also was brief. Second, with only four groups per course it was impossible to analyze the relationship between individual actions in Knowledge Forum and evidence for knowledge-building discourse statistically. It may be useful to conduct a study of much larger scale to examine this relationship. Finally, this study examined only one aspect of knowledge-building discourse—work in an online environment. Further studies would be useful for examining the scalability of the deep integration of this aspect with classroom activities across academic levels. Despite these limitations, this study provides empirical support for arguments for rethinking the nature and purpose of students’ work in online environments and suggests that students in different academic levels but at the same grade level can engage in knowledge-building discourse.

Reference List

Abelson, R. P. (1963). Computer simulation of “hot cognitions”. In S. Tomkins & S. Messick (Eds.), *Computer simulation and personality: Frontier of psychological theory* (pp. 277-298). New York: Wiley.

Baker, M. (2003). Computer-mediated argumentative interactions for the co-elaboration of scientific notions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 47-78). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Bell, P. (2004). Promoting students’ argument construction and collaborative debate in the science classroom. In M.C. Linn, E.A. Davis, & P. Bell (Eds.), *Internet*

environments for science education (pp. 115-143). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Bereiter, C., & Scardamalia, M. (1993). *Surpassing ourselves: An inquiry into the nature and implications of expertise*. Chicago, IL: Open Court.

Bereiter, C., Scardamalia, M., Cassells, C., & Hewitt, J. (1997). Postmodernism, knowledge-building, and elementary science. *Elementary School Journal*, 97(4), 329-340.

Bloom, L. Z. (1980, March). *The composing processes of anxious and non-anxious writers: A naturalistic study*. Paper presented at the annual meeting of the Conference on College Composition and Communication, Washington, DC.

Brand, A.,G.(1986). Hot Cognition: Emotions and Writing Behavior. *Journal of Advanced Composition*, 6, 5-15.

Bransford, J., Brown, A., & Cocking, R. (Eds.) (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

British Columbia Ministry of Education (n. d.). Grade 10 - Applications of social studies. Retrieved July 11, 2005, from <http://www.bced.gov.bc.ca/irp/ss810/ass10.htm>

Burtis, J. (1998). *The analytic toolkit*. The Ontario Institute for Studies in Education, The University of Toronto: Knowledge Building Research Team.

Chan, C., & van Aalst, J. (2003). Assessing and scaffolding knowledge building: Pedagogical knowledge building principles and electronic portfolios. In B. Wasson, S. Ludvigsen & U. Hoppe (Eds.), *Designing for change in networked learning environments. Proceedings of the international conference on computer support for collaborative learning* (pp. 21-30). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Chan, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, 15, 1-40.

Conley, A., Pintrich, P., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29(2), 129-163.

Daly, J. (1978). Writing apprehension and writing competency. *Journal of Educational Research*, 2, 10-14.

Dillenbourg, P. (1999). Introduction. What do you mean by 'collaborative learning'? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-19). Amsterdam, the Netherlands: Pergamon, Elsevier Science.

Edelson, D. C., Pea, R. D., & Gomez, L. M. (1996). The Collaboratory Notebook. *Communications of the Association for Computing Machinery*, 39(4), 32-33.

Faigley, L., Daly, J. A., & Witte, S. P. (1981). The role of writing apprehension in writing performance and competence. *Journal of Educational Research* 75, 16-21.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-40.

Fishman, B., Marx, R.W., Blumenfeld, P., Krajcik, J., & Soloway, E. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences*, 13(2), 43-76.

Fjermestad, J., Hiltz, S. R., & Zhang, Y. (2005). Effectiveness for students: Comparisons of "in-seat" and ALN courses. In S. R. Hiltz & R. Goldman (Eds.), *Learning together online: Research on asynchronous learning networks* (pp. 39-80). Mahwah, NJ: Lawrence Erlbaum Associates.

Guzdial, M., & Turns, J. (2000). Effective discussion through a computer-mediated anchored forum. *Journal of the Learning Sciences*, 9, 437-470.

Hakkarainen, K. (2003). Emergence of progressive-inquiry culture in computer-supported collaborative learning. *Learning Environments Research*, 6, 199-220.

Hakkarainen, K., Lipponen, L., & Järvelä, S. (2002). Epistemology of inquiry and computer-supported collaborative learning. In T. Koschmann, R. Hall, & N. Miyake (Eds.), *CSCLE 2: Carrying forward the conversation* (pp. 11-41). Mahwah, NJ: Lawrence Erlbaum Associates.

Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion threads. *Journal of Educational Computing Research*, 28, 31-45.

Hewitt, J. (2005). Toward an understanding of how threads die in asynchronous computer conferences. *Journal of the Learning Sciences*, 14(4), 567-589.

Hsi, S. (1997). Facilitating knowledge integration in science through electronic discussion: The Multimedia Forum Kiosk. Unpublished doctoral dissertation, University of California, Berkeley, CA.

Kolodner, J.L., Camp, P.J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., et al. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting Learning by Design™ into practice. *The Journal of the*

Learning Sciences, 12, 495-547.

Koretz, D., Stecher, B., Klein, S., & McCaffrey, D. (1994). The Vermont portfolio assessment program: Findings and implications. *Educational Measurement: Issues and Practice*, 13(3), 5-16.

Lamon, M., Secules, Petrosino, A. J., Hackett, R., Bransford, J. D., & Goldman, S. R. (1996). Schools for thought: overview of the project and lessons learned from one of the sites. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New education environments* (pp. 243-288). Mahwah, NJ: Lawrence Erlbaum Associates.

Law, N., & Wong, E. (2003). Developmental trajectory in knowledge building: An investigation. In B. Wasson, S. Ludvigsen, & U. Hoppe (Eds.), *Designing for changes* (pp. 57-66). Dordrecht, Netherlands: Kluwer Academic Publishers.

Lee, E., Chan, C., & van Aalst, J. (2006). Students assessing their own knowledge advances in a knowledge building environment. *International Journal of Computer-Supported Collaborative Learning*, 1, 277-307.

Linn, M., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Mahwah, NJ: Lawrence Erlbaum Associates.

National Research Council (NRC) (1996). *National science education standards*. Washington, DC: National Academy Press.

Niu, H. (2006). *Exploring participation in knowledge building: An analysis of online discussions in mainstream and honours social studies courses*. Unpublished master's thesis, Simon Fraser University.

Reed, M., Burton, J., & Vandett, N. (1988). Daly and Miller's writing apprehension test and Hunt's T-unit analyses: Two measurement precautions in writing research. *Journal of Research and Development in Education*, 21(2), 1-8.

Roth, W., & Tobin, K. (2002). *At the elbows of another: Learning to teach through coteaching*. New York: Peter Lang.

Sawyer, R. K. (2006). Analyzing collaborative discourse. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 187-204). New York, NY: Cambridge University Press.

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal education in a knowledge society* (pp. 67-98). Chicago, IL: Open Court.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning*

sciences (pp. 97-115). New York, NY: Cambridge University Press.

Scardamalia, M., & Bereiter, C. (2007). "Fostering communities of learners" and "knowledge building": An interrupted dialogue. In J. C. Campione, K. E. Metz, & A. S. Palincsar (Eds.), *Children's learning in the laboratory and in the classroom: Essays in honor of Ann Brown* (pp. 197-212). Mahwah, NJ: Lawrence Erlbaum Associates.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into World 3. In K. McGilley (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 201-228). Cambridge, MA: MIT Press.

Slater, A., & van Aalst, J. (2002). An exploration of the role of sociocultural factors in students participation in knowledge-building communities. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of the Computer-supported Collaborative Learning 2002 Conference* (pp. 617-618). Hillsdale, NJ: Erlbaum.

Stahl, G. (2002). Rediscovering CSCL. In T. Koschmann, R. Hall, & N. Miyake (Eds.), *CSCL 2: Carrying forward the conversation* (pp. 169-181). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

van Aalst, J. (2006). Rethinking the nature of online work in asynchronous learning networks. *British Journal of Educational Technology*, 37, 279-288.

van Aalst, J., & Chan, C.K.K. (2001, March). *Beyond "sitting next to each other": A design experiment on knowledge building in teacher education*. In P. Dillenbourg, A. Eurelings, and K. Hakkarainen (Eds.), *European perspectives on computer-supported collaborative learning: Proceedings of the first European conference on computer-supported collaborative learning* (pp. 20-28). Maastricht: University of Maastricht.

van Aalst, J., & Chan, C.K.K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *The Journal of the Learning Sciences*, 16, 175-220.

van Aalst, J., & Hill, C. (2001, June). *Experienced teachers as novice knowledge builders: Informing professional development*. Paper presented at EdMedia 2001, Tampere, Finland.

Wells, G. (1999). *Dialogic inquiry: Toward a sociocultural practice and theory of education*. New York, NY: Cambridge University Press.

WISE. Web-based Inquiry Science Environment. Retrieved from <http://wise.berkeley.edu/>

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling and metacognition: Making science accessible to all learners. *Cognition and Instruction*, 16(1), 3-118.

Zhang, J., Scardamalia, M, Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. *Education Technology Research Development*, 55, 117-145.

Zohar, A. & Dori, Y. J. (2003). Higher order thinking skills and low achieving students—are they mutually exclusive? *Journal of the Learning Sciences*, 12(2), 145-182.

Acknowledgments

This research was completed when both authors were at Simon Fraser University, and was supported by a collaborative research network grant from the Social Sciences and Humanities Council of Canada to M. Scardamalia. We thank all collaborating teachers at the school, especially the classroom teacher in this study, Bruce Cunnings. We are also grateful to Carol Chan, Laura D'Amico, Stephen Campbell, and the editors of this special issue for useful comments on the study.

Appendix

Working at the cutting edge.

The following excerpt of 14 notes from a view on air quality involved all eight students in a group from the honours class in Implementation 1; it focused on causes of air pollutions in big cities and received a score of 3. This excerpt shows how points made earlier by group members became part of a coherent argument for a new problem. Student A identified a gap in the ideas that were under discussion:

Another cause which I have not seen posted yet is wind speed.” He then explained the importance of introducing this idea to the group by explaining two sides of the issue: Higher wind speeds result in less pollution. This is because the stronger the wind is, the more dispersed the pollution will be. The opposite ... is true for places with very low wind speeds where the pollution will be much worse.

Student B lent support to student A's idea with an example and explanations: “I agree that wind is an important factor in air pollution. An example of this is seen in Mexico City.” This contribution made a link to previous contributions that examined pollution in Mexico City, including the city's location “in the crater of an extinct volcano” and incomplete combustion leading higher emissions of carbon monoxide and other substances. However, this discussion had not raised the influence of wind speed. In sum, in this discussion students revisited and integrated earlier

ideas in order to articulate an emerging problem, warranting a rating higher than a 2. However, the problem still lacked the widespread commitment to its pursuit needed for a 4.

Progressive problem solving.

A sequence of 43 notes by all students in a group in the mainstream class in Implementation 1 received a rating of 4. Student E stated that “*without this Glycerol, [pine beetles] will die due to the cold temperature.*” This information stimulated an ongoing discussion within the group. First, student F asked whether the pine beetles produced the anti-freeze (i.e., glycerol) throughout the year. Student E then explained in more detail:

I think I did not explain it very well. The GLYCEROL they produce just before winter is the anti-freeze they need to protect themselves in winter. So, if the temperature drops lower when they are still making this anti-freeze then they cannot stand the coldness because they haven't got the anti-freeze they need.

Student F then explained this idea to other students who were still confused by the anti-freeze idea. For example, when student G stated “*not making sense,*” student E explained another way why pine beetles are killed when it starts freezing before they have made sufficient glycerol. Student E also constantly asked group members whether they understood the idea, from: “[2 students] *still don't get it ... so I'm asking them which part they don't understand,*” to the final note of the thread “*so u guys actually understand it??*”. This episode contained continual effort and sustained inquiry to work on problems wherein one problem led to another and there was growth; this process was realized through evolution of ideas, identifying and solving problems, as well as raising further questions from original thoughts.

Collaborative effort.

The following note was part of a 32-note discussion on reducing pollutions by automobiles by a group from the honours class in Implementation 1; it received a score of 3. The note made use of three scaffolds, shown in parentheses:

(Opinion:) I agree that people should start getting into a habit of walking more. However, don't forget that we're not only focusing on Mexico City. Also, (My theory:) I don't think that just having most people walking would solve the pollution problem, seeing as some people would probably still have to use cars on certain occasions, and some old cars are very bad for the environment. (Example:) During spring break, I saw some old car spewing pure black smoke out of its exhaust pipe. Certain old cars like that need to be changed.

The use of scaffolding in this note reveals an effort to help the reader understand, beginning from an *opinion* and then elaborating this as a *theory* and providing an example to illustrate the theory. While this note was designed to help others understand an issue, it did not provide evidence for the integration of multiple

perspectives needed for a 4.

Identifying high points.

The following note was part of a 16-note discussion of waste management by a group from the honours class in Implementation 2; it received a rating of 3. Student K stated that *“Looking back on my note, I realized that taxes are not going to be the best solution...in any solution we are trying to come up with, we should be considering the people’s feelings.”* This note reveals that the student had some insight into his learning (that his previous idea had limitations) and proposed a new strategy (to consider other people’s feelings).

Constructive use of authoritative sources.

The following note, part of 13 notes on rainforests by a group in the honours class in Implementation 1 received a rating of 2. Student L introduced a research result: *“A recent study by Professor James Alcock has shown that current logging rates are reducing the Amazon rainforest by 1% a year. That may seem like a small number, but it actually is devastating the ecosystem of the forest.”* This note was not considered sufficient for a 3 as it did not provide bibliographical details or link to the research report, and did not relate the finding to the discussion or raise questions about it. Though not a strong example of the principle, the note indicates the student consulted an external source.