Fostering Collective Knowledge Advancement Through Idea-Friend Maps in a Large Knowledge Building Community

Xueqi Feng, The University of Hong Kong, fengxueqi@hotmail.com
Jan van Aalst, The University of Hong Kong, vanaalst@hku.hk
Carol K.K. Chan, The University of Hong Kong, cckchan@hku.hk
Yuqin Yang, Central China Normal University, yuqinyang0904@gmail.com

Abstract: This study designed and examined a knowledge-building environment that combined a learning analytics tool for creating “Idea-Friend Maps” with Knowledge Forum for Primary Five students. Students in the experimental class (n=53) used the Idea-Friend Map tool, whereas those in the comparison class (n=54) conducted the same knowledge-building activities without the use of analytics tools. Results showed that students in the experimental class showed better conceptual understanding, higher levels of social participation, and more advanced collective knowledge. Analyses were conducted on how students used analytics to support their knowledge building inquiry. These findings suggested that learning analytics can be used to promote knowledge building in large communities.

Introduction
China is committed to education reform and is open to Western theories that have been developed in the learning sciences (Ryan, 2013). Knowledge building is a major pedagogical approach in the learning sciences community that focuses on theory building and knowledge creation (Scardamalia, 2002). The online platform Knowledge Forum® (KF) was designed to promote collective knowledge advancement by creating an online environment where students can pose problems, propose explanations, test ideas, and conduct sustained pursuit of inquiry to facilitate collective knowledge advancement. Generally, most studies on knowledge-building practice involve smaller-sized classes (15-30 students) (Chen, Scardamalia, & Bereiter, 2015; Oshima et al., 2004; van Aalst & Truong, 2011; Zhang, Scardamalia, Reeve, & Messina, 2009); however, classes in China generally contain 50-60 students, and very few studies involve classes with more than 50 students. Thus, little is known about how to implement knowledge building in such settings, nor about issues of scale in online discourse. How to facilitate large knowledge-building communities engaging in productive discourse is a challenging issue.

Generally, when classes larger than about 20 students, teachers have adopted group-based instruction (van Aalst, 2009) in knowledge building environments. Zhang et al. (2009) examined different social configurations—i.e., fixed groups (students work in fixed groups), interactive groups (interacting groups with cross-group activities), and opportunistic groups (new groups forming and disbanding based on emergent goals)—and found opportunistic groups yielded the best learning outcomes. However, whether opportunistic interaction adapts to large class environments still needs to be verified. Furthermore, when such a community promotes flexible participation and emergent inquiries on KF, it is difficult for students to understand the changing status of collective knowledge at the community level (Kimmerle, Cress, & Hesse, 2007). Therefore, the most urgent challenge is how to design an environment to help large class students identify, from the large volume of online discussion notes, the cutting edge of community knowledge, conduct sustainable problem solving, and structure what the community should investigate. However, the large number of online notes generated by large classes may provide opportunities for educational data mining and learning analytics, as it involves large datasets (Baker & Inventado, 2014).

There is increasing interest in learning analytics that emphasizes informing and empowering learners and instructors to improve their learning processes (Nistor & Hernández-Garcia, 2018). However, much attention in this field involves evaluating what learners have done or predicting what they will do in the future (Siemens & Long, 2011); the area of how analytics are used as part of teaching and learning processes is relatively unexplored (Wise, Vytasek, Hauskncheht, & Zhao, 2016). Knowledge Building Discourse Explorer (KBDex) is a novel learning analytic tool designed to investigate student-, keyword-, and discourse-based social networks on KF (Oshima, Oshima, & Matsuzawa, 2012). However, many current studies use KBDex merely as a tool for evaluating learning processes (Ma, Matsuzawa, Chen, & Scardamalia, 2016; Matsuzawa, Oshima, Oshima, & Sakai, 2012; Resendes, Scardamalia, Bereiter, Chen, & Halewood, 2015). Our program aims to design representations of the learning analytics tool KBDex, integrated with relevant social configurations for Grade Five students in large knowledge-building communities, and to examine the designed environment’s impact on students’ learning outcomes and collective knowledge advancement. Our research questions are: (1) Did students in the experimental class have better learning outcomes, greater KF participation, and deeper KB discourse? (2)
How did students in the experimental class engage in collective knowledge advancement, with the support of learning analytics?

**Methods**

**Participants**
This study was carried out in two Grade Five classes in a primary school in China, with 53 students in the experimental class and 54 in the comparison class. The first author was also the science teacher of the class. These students have had one-year experience in learning using knowledge building with this teacher before this study.

**Pedagogical design**
The two classes investigated the topic of *Electricity* on KF (Figure 1) for 9 weeks. In Phase 1 (Weeks 1-3), students in both classes first created KF notes for inquiry, categorized their online notes into seven subtopics, formed learning groups, and chose subtopics according to interests. In Phase 2 (Weeks 4-6), students in both classes interacted with each other in different subtopics, did experiments, and conducted theory-building; however, students in the experimental class used the group-level Idea-friend Maps (IFM) (Figure 2, (a) and (b)), while those in the comparison group did not. In Phase 3 (Weeks 7-9), students in both classes worked in opportunistic groups (students formed/dischanded new groups based on emergent goals/common interests) to contribute to community views (included all the subtopics); however, students in the experimental class used the community-level IFM (Figure 2, (c) and (d)), while those in the comparison group did not.

![Figure 1.](image1)  
![Figure 1.](image2)  

**Figure 1.** The KF view of *Conductor and Insulator* (a); the experimental pedagogical design (b).

![Figure 2.](image3)  
![Figure 2.](image4)  

**Figure 2.** The intervention of IFM: (a) Group 9’s group-level IFM in Week 4; (b) Group 5’s group-level IFM in Week 5; (c) community-level IFM in Week 8; (d) community-level IFM in Week 9.
The learning analytic tool (Idea-Friend Map)
The IFM intervention was the word network exported from KBDex. We designed the group-level IFM to help interactive activities in Phase 2, and the community-level IFM to promote opportunistic knowledge advancement in Phase 3. Figure 2 (a) and (b) show group-level IFM of different groups in different weeks; the red and yellow balls indicate words that had or had not already been discussed by the group, respectively; yellow balls near red balls indicate idea friends. For instance, students in Group 9 had already discussed those red balls words as “wire”, “circuit” and “current” until Week 4 (see Figure 2 (a)), while those yellow balls might be discussed by other groups. Words “dry” and “wood” from Figure 2(b) were “plastic”’s idea friends because they were very close to “plastic”. Students in different groups first identified idea friends, then went to other subtopics that included those ideas friends, enabling them to identify connections among different science concepts, make learning plans, interact with different groups, which would result in the advancement of collective knowledge. Correspondingly, Figure 2 (c) and (d) show community-level IFM during the final weeks of the intervention; balls with different colors (except yellow) indicate key topics identified in the community view, allowing students in opportunistic groups to synthesize different topics, summarize a topic, and create new knowledge with the support of community-level IFM. For example, the green ball with “Franklin” from Figure 2 (c) represented the problem of “Why did Benjamin Franklin not die from his lightning strike?” (see Table 2).

Data analysis and findings

RQ1: Did students in the experimental class have better learning outcomes, greater KF participation, and deeper KB discourse?

Changes in science learning outcomes across classes
We analyzed students’ science domain tests to examine the effects of IFM on students’ learning outcomes. There was no significant difference in the pre-test scores of the two classes (F (1, 101) = 0.81, p= .372). A two-way (Intervention × Time) ANOVA showed a significant main effect of Time (F (1, 199) = 145.53, p < .001, Partial $\eta^2$ = .42). A significant Time × Intervention effect was obtained (F (1, 199) = 5.84, p = .017, Partial $\eta^2$ = .03), indicating that students in the experimental class improved more on science learning outcomes, compared to students in the comparison class.

Class differences in KF contribution and interaction
We assessed quantitative data about the number of students’ notes written, notes linked, notes read, scaffolds used, reference used, and views worked on KF, using the analytic toolkit (ATK) (Burtis, 1998). One-way ANOVA was performed to assess whether students’ contributions to the online discourse were different among the classes. There were significantly higher values of notes read (F (1, 105) = 10.10, p< .01), views worked (F (1, 105) = 10.54, p< .01), and references used (F (1, 105) = 13.11, p< .001) in the experimental class. In contrast, there was no significant difference in the notes written, notes linked, and scaffolds used between the two classes (p > .1). These results suggest that the IFM intervention helped students to engage in different views, read more notes, and reference useful ideas.

Class differences in student online knowledge building discourse
We analyzed 588 KF notes from the experimental class and 528 from the comparison class to characterize students’ contributions to knowledge-building discourse. We adopted a framework to code the notes in each interactive view and the community view. The coding schemes included three main categories—questioning, theorizing and community—and corresponding subcategories, and drew upon theoretical frameworks for progressive inquiry (Hakkarainen, 2003) and knowledge creation (Chuy, Zhang, Resendes, Scardamalia, & Bereiter, 2011; Fu, van Aalst, & Chan, 2016). Two raters independently coded about 50% of all the notes. The inter-rater reliability was .94 for questioning, .94 for theorizing, and .93 for community (Cohen’s kappas).

As Table 1 shows, the experimental class created more theorizing and community notes, while the comparison group tended to raise questions. One-way ANOVA was adopted to further examine the differences between the two classes and found significantly higher values of improving an explanation (F (1, 14) = 5.88, p< .05), bridging knowledge (F (1, 14) = 8.30, p< .05), and synthesis (F (1, 14) = 5.15, p< .05) in the experimental class. We also observed significant higher values for simple claim (F (1, 14) = 5.88, p< .05) created in the comparison class. The differences in other codes were not significant between the two classes (p > .05). These results indicate that, when confronted with new problems, students in the experimental class tended to bridge knowledge through the website, engage in deeper theory-building processes, and synthesize useful ideas, while the comparison group tended to ask more fact-oriented questions and respond with simple claims, rather than seek relative notes or authoritative materials.
Table 1: Number of different categories of questioning, theorizing, and community in interactive views (experimental class/ comparison class)

<table>
<thead>
<tr>
<th>View</th>
<th>Questioning</th>
<th>Theorizing</th>
<th>Community</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fact seeking</td>
<td>Explanati on seeking</td>
<td>Sustained inquiry</td>
</tr>
<tr>
<td>#1</td>
<td>7/7</td>
<td>15/10</td>
<td>8/0</td>
</tr>
<tr>
<td>#2</td>
<td>8/9</td>
<td>9/12</td>
<td>6/3</td>
</tr>
<tr>
<td>#3</td>
<td>0/5</td>
<td>3/7</td>
<td>1/2</td>
</tr>
<tr>
<td>#4</td>
<td>3/16</td>
<td>8/9</td>
<td>6/6</td>
</tr>
<tr>
<td>#5</td>
<td>8/13</td>
<td>7/7</td>
<td>3/3</td>
</tr>
<tr>
<td>#6</td>
<td>15/19</td>
<td>11/12</td>
<td>6/7</td>
</tr>
<tr>
<td>#7</td>
<td>2/10</td>
<td>3/14</td>
<td>4/1</td>
</tr>
<tr>
<td>#8</td>
<td>2/7</td>
<td>9/11</td>
<td>19/13</td>
</tr>
<tr>
<td>Total</td>
<td>45/86</td>
<td>65/82</td>
<td>53/35</td>
</tr>
<tr>
<td>Mean</td>
<td>5.63/10.75</td>
<td>8.13/10.25</td>
<td>6.63/4.38</td>
</tr>
<tr>
<td>SD</td>
<td>4.87/4.86</td>
<td>3.98/2.49</td>
<td>5.45/4.21</td>
</tr>
</tbody>
</table>

Notes: Views #1 to #7 mean subtopics that interactive groups worked on, while view #8 means the community view. #1=Electricity generation; #2=Franklin and electricity; #3=Lightning; #4=Static electricity; #5=Circuit; #6=Conductor and insulator; #7=Electricity and magnet

Table 2: Categories of problems in the community views for both classes

<table>
<thead>
<tr>
<th>Static electricity</th>
<th>Circuit</th>
<th>Conductivity</th>
<th>Franklin and electricity</th>
<th>Electricity and magnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental class</td>
<td>#1 How does lightning form? #2 What is the object can create static electricity? #3 Why is static electricity formed from friction? #4 Can I see current when static electricity is transmitted?</td>
<td>#1 Why does circuit need a closed loop? #2 How does the body make electricity? #3 Why does people might die from an electric shock?</td>
<td>#1 Why is mineral water/beverage/wet wood/metal conductive?</td>
<td>#1 Why did Benjamin Franklin not die from his lightning strike? #2 What is the relationship between a lightning rod and insulator?</td>
</tr>
<tr>
<td>Control class</td>
<td>#1 How does lightning form? #2 Why is static electricity formed from friction? #3 How to remove static electricity?</td>
<td>#1 Why does circuit need a closed loop? #2 How is electricity produced? #3 Will a short circuit cause an explosion?</td>
<td>#1 Why can't the insulator be conductive? #2 Why is metal/mineral conductive?</td>
<td>#1 Why did Benjamin Franklin not die from his lightning strike?</td>
</tr>
</tbody>
</table>
RQ2: How did students in the experimental class engage in collective knowledge advancement with the support of the learning analytics?

Class differences in interactive views
We compared key interactive views reflecting how student advanced collective knowledge with the support of the group-level IFM. We first quantitatively compared the Total Degree of Centrality (TDC) of the KBDEX word network to assess the collective knowledge advancement (Oshima et al., 2012) of the interactive views Conductor and Insulator (Figure 3(a)) and Circuit (Figure 3(b)). We focused on these two views because they are the most fundamental parts when students discussed Electricity. A higher degree centrality means a denser social network. Figure 3 shows the word network’s increasing TDC over time, indicating how students collectively worked on key science ideas. In Phase 1, there were no significant differences in the collective knowledge advancement of the two classes in either view (p> .5), when they shared the same instruction. However, there were significantly higher TDC values in Conductor and Insulator (F (1, 4) = 101.86, p= .001) and Circuit (F (1, 4) = 112.16, p< .001) for the experimental class compared to the comparison class when students used IFM, in Phase 2. We also qualitatively analyzed students’ online notes, prompt sheets and experiment reports when they engaged in the interactive views, Conductor and Insulator and Circuit.

Figure 3. TDC of word network for the Conductor and Insulator view (a) and the Circuit view (b).

View: Conductor and insulator
We focused on the students from Group 9 in the experimental class who took responsibility for the Conductor and Insulator view to clarify how they engaged in collective knowledge advancement with the support of group-level IFM. After posting questions, proposing explanations, and searching authoritative information in the previous three weeks of Phase 1, Group 9 students had a general understanding of conductors and insulators, then they looked at their group IFM in Week 4, they found the keywords related to their topics, such as “conductivity”, occupied the right half of the entire IFM (Figure 2 (a)), while some keywords, such as “battery” and “voltage”, belonged to the Circuit group. Therefore, they speculated that, if they wanted to further explore their own topic, they must first understand the basic knowledge of “circuit”. Then they went to the Circuit view and managed to design an experiment to verify the elements of a circuit by lighting up a light bulb (Figure 4 (a)) by themselves. Later, they successfully revised their circuit experiment to explore whether water/wood was conductive (Figure 4 (b)). In addition, the students could also spontaneously review their thinking processes and try to explain their experimental phenomena in light of another's perspective. Below is a summary note tracing how they explored the relationship between Circuit and their own topic, Conductor and Insulator.

[My idea] What is the relationship between circuit and conduction? In fact, the most basic thing about conduction is the circuit… To make an object conductive, we must first do a circuit experiment.

(“Group summary of experiment” by Group 9) We conducted two experiments: the first was to make the bulb light up. This experiment was a pilot for conductor and insulator… The second experiment, of water’s conductivity, was the experiment we originally wanted to explore… Because pure water has no impurities, most of it is not electrically conductive. In mineral water there are impurities, so the light bulb is illuminated when the wire is placed in it, so the two are electrically conductive.

(“Reply” by student c726) c726 also said that water can conduct electricity because there are cells inside it. You can understand in this way.
But the Conductor and Insulator group in the comparison class tended to ask fact-oriented questions, simply look up information, and copy/paste difficult conceptions from websites. The following is a cluster from the comparison group discussing the same topic (water’s conductivity).

[I want to know] What object can be conductive?
[My idea] Water!
[I want to know] Why is water conductive?
[My idea] If all the impurities in the water are removed, the water resistivity will increase up to 18.2 megaohms x cm…
[My idea] Please provide information that we can understand.

View: Circuit
Then we focused on Group 5, who worked on Circuit in the experimental class. After they acquired the basics of circuit and successfully connected circuit to light up a light bulb, they began to explore deeper knowledge of electricity from other views. Therefore, they first identified the keywords “wood” and “dry” as idea friends from their group-level IFM in Week 5 (Figure 2 (b)), went to the Conductor and Insulator view to get the information they wanted, made a group plan for exploration, and did experiments to explore why wet wood can conduct electricity while dry wood cannot (Figure 4, (c) and (d)). Finally, Group 5 and Group 9 students collaborated to explain their common experimental phenomena and created a theory-building note (below) to explain the conductivity of different waters and wood.

[Putting our theories together] 1. Mineral water/beverage is conductive because there are ions that can move freely; pure water can’t conduct electricity because there are no ions that can move freely. 2. Why wet wood is conductive: wet wood contains water, water is conductive.

The Circuit topic is the foundation of Electricity, so Group 5 students began to explore other important topics from Week 5, at which point the TDC value began to show a slight downward trend, as shown in Figure 3 (b); however, the TDC value in the experiment class remained higher than that of the comparison class. Students in the comparison class were concerned with fact-seeking questions—e.g., “What is voltage?”; “What is the unit of voltage?”; “What contribution does Ampere make to electricity?”; and “Where can I see the circuit in my life?”

Class differences in the opportunistic community view
To characterize how community-level IFM helped students advance collective knowledge, we analyzed students’ activities in the community view. First, we compared five categories of big problems identified in each class (Table 2). Results show that the experimental class identified more and deeper problems, while the comparison class tended to raise fact-seeking questions. This phenomenon is consistent with the notes coding results in RQ1.

In addition, students in both classes were invited to explore relationships, summarize, and create new knowledge according to the identified problems in Table 2. Students in the experimental class could conduct those activities smoothly with the support of community-level IFM (Figure 2, (c) and (d)), while students in the comparison class could only build on their original notes. We selected some community notes from the experimental class to show how those students synthesized, summarized, and created new knowledge through IFM.
Theme 1: Synthesizing different topics
Student c705 from Group 5 found the keywords of the two topics “Franklin” and “metal” on IFM (Figure 2 (c)) were particularly close, then they tried to explain “why Franklin did not die” from the perspective of “why metals can conduct electricity”.

[I want to discuss this topic] Why did Benjamin Franklin not die from his lightning strike?
[I want to explore relationships among those keywords] metal, Franklin and copper key
[My theory] Thunderbolt voltages range from hundreds of millions to billions of volts, but Franklin was not electrocuted. We can conclude that he did not directly touch the copper key. Even if using Leiden bottles and ribbons, people will die if they touch a few hundred volts (unless the voltage is very, very small).

Theme 2: Summarizing a key theme
Students c718, c726 and c711 from Group 6 found there were many keywords around “Franklin” in Figure 2 (c) and speculated those might be factors in protecting Franklin from the lightning strike. Then, they wrote down those keywords and selected relative information from the community.

[I want to discuss this topic] Why did Benjamin Franklin not die from his lightning strike?
[I want to summarize those keywords] Voltage, Leiden bottle, capacitor, wire, copper key, lightning rod, wind, ribbon, current, and metal.
[Putting our theories together] 1. Voltage (Low voltage by student c751); 2. Leiden bottle (Low voltage by student c722); 3. Iron wire (Iron wire by student c723); 4. Copper key (Synthesizing: Franklin and Metal by student c705); 5. Lightening rod (Lightening rod by student c713); 6. Ribbon (Ribbon by student c722); 7. Current (Current by student c723).

Theme 3: Creating new knowledge
Students c733 and c734 from Group 4 took responsibility for Static Electricity and found that the keywords “cloud” and “insulator” were very close in Figure 2 (c). Then they searched for information, identified related notes, and finally came up with new “knowledge”: “the cloud is an insulator.”

[I want to discuss this topic] What is the object can create static electricity?
[I want to create new knowledge on these keywords] cloud and friction.
[My theory] Our new knowledge is “the cloud is an insulator.” Evidence: 1. Electrostatics are all insulators (from Baidu); 2. Lightning is caused by frictions among clouds (from notes)

Students c708 from Group 5 and c752 from Group 2 provided more evidence to support the new knowledge—i.e., water evaporates into pure water vapour, which in turn forms clouds. Experiments proved pure water is an insulator, so the resulting cloud is also an insulator. Finally, we can see the new knowledge of “evaporation” in the next-week community-level IFM (Figure 2(d)). These data suggest that these young children, even in their simple writing, had strong knowledge-creation abilities, with the aid of IFM.

Discussion
This study investigated how learning analytics can be used to promote knowledge building in a Grade 5 science classroom. First, we compared differences in learning outcomes, KF participation, and depth of knowledge building discourse in both experimental and comparison classes. Students in the experimental class showed more improvement in science concept understanding, higher social dynamics of reading, referencing and interacting across different subtopics, and a deeper discourse on theory building and synthesizing. Analysis of students’ group plans, experiment reports, and summary notes revealed how students used both group- and community-level IFM to identify learning gaps, make plans, co-construct, collect evidence, synthesize key ideas, summarize, and create new knowledge. Even though the IFMs are not concurrent, representations of learning analytics could be helpful suggesting some practical use in classrooms. Furthermore, we integrated learning analytics with social configurations in large classes—i.e., interactive groups with group-level IFM, and opportunistic interactions with community-level IFM—provide a pedagogical design for the less explored field of knowledge building in large communities (Scardamalia & Bereiter, 1996).
References


