Using Reflective Peer Assessment to Promote Students’ Conceptual Understanding through Asynchronous Discussions

Huann-shyang Lin1, Zuway-R Hong2, Hsin-Hui Wang2 and Sung-Tao Lee3

1Center for General Education, National Sun Yat-sen University, Kaohsiung, Taiwan // 2Institute of Education, National Sun Yat-sen University, Kaohsiung, Taiwan // 3Department of Applied Science, Naval Academy, Kaohsiung, Taiwan // huannlin@faculty.nsysu.edu.tw // a3803429@ms49.hinet.net // hsinhui5885@gmail.com // sungtao@mail.cna.edu.tw

ABSTRACT

This study explores the impact of using assessment items with competing theories to encourage students to practice evaluative reflection and collaborative argumentation in asynchronous discussions. Thirty undergraduate students from various departments worked in small groups and took turns collaboratively discussing the given item’s answer, reaching a consensus, and posting their consensual answer on the web. The remaining participants served as evaluators to reflect on the answer and provide comments. It was found that the students made significant progress in argumentation ability and conceptual understanding of related scientific content knowledge. In the beginning of the study, the group of students majoring in the sciences outperformed counterparts with non-science majors on the level of understanding of the assessment item’s scientific concepts. At the end of the semester, the differences diminished between the two groups both on conceptual understanding and in argumentation ability.

Keywords

Argumentation, Asynchronous, Conceptual understanding, Reflection

Introduction

Existing literature has indicated that online asynchronous discussion has the potential to generate the critical dimensions of learning found in traditional classrooms (Andresen, 2009); furthermore, it has the ability to enhance higher cognitive levels of knowledge construction (Schellens & Valcke, 2005). However, simply putting students in an asynchronous discussion environment does not necessarily bring about collaborative interactions and effective outcomes, because some students may be reluctant to disagree with others (Andriessen, 2006) or will vary in their level of involvement (Veerman, 2003). In order to promote the effectiveness of asynchronous discussions, pioneers in collaborative learning examined a variety of strategies. For example, assigning roles to students at the beginning of discussions resulted in a significant positive impact on students’ level of knowledge construction (De Wever, Van Keer, & Valcke, 2009); Elaborating on the meaning of discussion questions promoted balanced argumentation for all participants, especially for those with less knowledge of the question (Golanics & Nussbaum, 2008); and engaging students in reflective interactions such as explaining, justifying, and evaluating problem solutions has shown a productive learning outcome for physics modeling tasks (Baker & Lund, 1997). After considering the strategies recommended by existing literature, we designed an assessment instrument with competing theories for students to practice evaluative reflection through asynchronous discussions and examined its impact on key competencies of educational outcomes- students’ conceptual understanding and argumentation ability (Driver, Newton, & Osborne, 2000).

Argumentation in asynchronous discussion environment

Social constructivism theory furthers the idea that learning effect can be promoted through active interactions and communication among participants. For social constructivism, knowledge is created and legitimized by means of social interactions between and among individuals in a variety of community, societal, and cultural settings (Driver, Leach, Millar, & Scott, 1996; Staver, 1998). When this theory is appropriately applied in the context of education, students are encouraged to interact with others to construct individual understanding and knowledge. It also provides opportunities for students to reflect on other classmates’ comments, suggestions, presentations, and ways of learning.

In the process of online argumentation, students are encouraged to actively write, discuss, and debate online using text-based communication tools. Students make progress in argumentation by providing evidence-based conclusions, describing why they agree or disagree with the presented statements, and try to persuade others. In order to provide
quality arguments, students must explain their own positions, evaluate current arguments, summarize peer comments, and integrate related information or knowledge. All of these activities seem to be helpful for students' clarification of conceptual understanding and improvement of argumentation. More importantly, the time delays in text-based asynchronous discussions provide opportunities for students, especially for those who need more time, to reflect and scrutinize online information (Veerman, Andriessen, & Kanselaar, 2000). In collaborative learning, meaning is produced by examining the relationship between utterances through social interactions; meaning is examined and reconstructed as a direct result of conflict or argumentation in a social context (Jeong & Joung, 2007). For critical argumentation, students can use “counter-arguments” to “challenge” other students’ statement when they disagree with the statement (Veerman et al., 2000). The learning of argumentation is consistent with the development of scientific knowledge in most science communities. Vigorous discussions, debates, and peer-review procedures are exchanged among scientists. Scientists then provide counter-arguments or rebuttals to challenge the data, evidence, or assertions of other scientists with different theories. For example, in the eighteenth century, the phlogiston theory was almost universally accepted at the time and was the basis of the chemistry taught to college students then. The theory hypothesized that during combustion the substance of phlogiston was released and combined with air (Conant, 1957; Harre, 1981). However, French scientist Lavoisier provided empirical evidence of burning mercury and phosphorous to challenge the phlogiston theory. He found that the result of burning mercury or phosphorous did not decrease their weights as the theory predicted, instead, the final weights increased. This rebuttal and more empirical data from his follow-up experiments (e.g., an experiment collecting the gas formed after heating the red oxide of mercury) play key roles for the demise of the phlogiston theory. By reviewing the development of scientific knowledge and the history of science, we can find that argumentation plays a central role in the resolution of scientific controversies (Fuller, 1997).

Similar to the development of scientific knowledge, argumentation deserves a place in the pedagogy of science. In arguing the use of argumentation theory in education, Driver et al. (2000) concluded that helping students to construct coherent arguments and evaluate others are important skills, especially pertaining to topics reported in the media. This is even more so in our contemporary and democratic society, since there are many public policies relating to science and the public has a legitimate voice (e.g., use of bio-ethanol as fuel, restriction of genetically modified foods, and control of air quality) (Newton, Driver, & Osborne, 1999). Through the practices of posing and evaluating arguments, students become active participants in the learning community rather than just passive knowledge receivers. Another potential benefit of collaborative argumentation in stimulating students' conceptual understanding and belief revision has been examined (Ravenscroft, 2000). In the study, the learner adopts the role of an explainer while the computer system plays a facilitating role; participants collaborate to develop a shared explanatory model. Ravenscroft found that students revised their beliefs and improved their explanatory models in a delayed post-test.

Despite the emphasis of argumentation ability in science teaching, it is rarely adopted in typical classroom teaching. Major reasons ranged from teachers’ perception of the difficulty and challenge in managing group discussions to the time pressure imposed by the need to cover the national curriculum (Driver et al., 2000; Newton et al., 1999). The literature review reveals that there are limitations and constraints for teachers to implement collaborative argumentation or group discussion in typical classroom teaching practices. On the other hand, benefits of online asynchronous discussion may make students’ interactions and group discussions more effective than traditional face to face discussions. For example, it allows slow-paced students more time to construct arguments and contribute to the discussion (Veerman et al., 2000); the transcript of the discussion is always available for participants’ reference (Weinberger & Fischer, 2006); and discussions generally will not be interrupted by a particularly aggressive participant (Andriessen, 2006). Furthermore, in typical classroom face to face settings, most of the interactions are generally dominated by outspoken students. The learning opportunity of practicing communication or argumentation for students who are either shy or weaker speakers is unintentionally limited or deprived (Nussbaum & Jacobson, 2004). Therefore, the discourse of the science classrooms needs to be more deliberative or dialogic (Simon, Erduran, & Osborne, 2006). This situation of inequity in learning deserves attention from teachers and educators. Although there is reason to hypothesize that an asynchronous discussion could be one of the effective alternatives to conduct argumentation activities, more empirical studies focusing on inspiring the practice of reflection and promoting the highly expected educational outcome of aforementioned argumentation and conceptual understanding are needed. These studies should also aim to enable us to better understand what role computers can play in supporting the classroom teaching that has mostly failed to promote these higher level cognitive abilities (Webb, Jones, Barker, & van Schaik, 2004).
The importance of reflective ability

Reflective ability has long been regarded as one of the major goals for students’ learning outcome. Early in 1933, Dewey proposed reflection as a process of problem solving. Recently, the Organisation for Economic Co-Operation and Development (OECD, 2006) officially identified reflective ability as an important component of scientific, reading, and mathematical literacy. Based on the OECD’s definition, the competency of reflection is deemed to be essential for an improved future. Knowing the importance of reflective ability, researchers and educators have asserted that it can be taught and trained through proper design of activities (Boud, Keogh, & Walker, 1985; Lee & Hutchison, 1998). An investigation of teaching effectiveness by designing a variety of online activities has been conducted. For example, Chen et al. (2009) used “reflection prompts” to engage online learners in a reflective learning environment. They found that the reflection level of those students who had been provided with “high level prompts” was significantly higher than the level of those who were not shown these prompts. Review of the literature above reveals that providing reflective learning opportunities for students could be fruitful in promoting expected outcomes. Further examination of current practices of reflective activities found that most studies mainly focused on the assessment of reflective ability level (Chen et al., 2009; Gulberg & Pilkington, 2007; Yang, 2009). A limited number of studies have investigated how reflective activities influenced learners’ key competencies in educational outcomes (e.g., argumentation or science concept understanding). To this end, we designed a semester-long online reflective peer assessment and investigated its impact on students’ argumentation and conceptual understanding.

Typical assessment vs. Reflective peer assessment

The term of reflective peer assessment means in a collaborative learning environment, students critically assess each other’s feedback posted online, vigorously discuss various perspectives, and continuously reflect and elaborate on their own assertions (Veerman et al., 2000). For typical classroom teaching practices, although inquiry skills or group discussions can be used to promote student-teacher or student-student interactions to assess student conceptual understanding (Van Zee, Iwasyk, Kurose, Simpson, & Wild, 2001), time constraints or lack of professional ability may prevent instructors from employing these teaching strategies (Newton et al., 1999; Roth, 1996). In this study, we propose a reflective peer assessment in an asynchronous discussion learning environment to promote collaborative learning and critical argumentation. The details of reflective peer assessment will be explained with examples in the methodology section.

Although considerable research has been devoted to computer supported collaborative learning (CSCL) through the perspectives of investigating collaborative behaviors (Hsu, Chou, Hwang, & Chou, 2008), identifying strategies and factors related to better collaboration (De Smet, Van Keer, & Valeke, 2008; Onrubia & Engel, 2009), and analyzing the effect and the role of peer feedback (Lai & Lan, 2006; Tseng & Tsai, 2007), rather less attention has been paid to the systematic integration of key competencies of educational outcomes in the CSCL environment: reflection, argumentation, and conceptual understanding.

Purpose of the study

As mentioned earlier, interactions of student learning among the three variables of reflection, argumentation, and conceptual understanding have rarely been investigated simultaneously. A literature review reveals that asynchronous discussions provide opportunities for students to work collaboratively and have great potential in promoting learning outcomes, especially when students were encouraged to reflect on another team’s consensual answer and peer arguments and the reasons why they agree or disagree with the consensual answer. The purpose of this study is to explore the impact of using reflective peer assessment in asynchronous discussions on the development of students’ argumentation ability and conceptual understanding. Specifically, the research questions are as follows:

1. How students progress on the development of their argumentation ability and conceptual understanding through the six rounds of reflective peer assessment in asynchronous discussions?
2. Do both science major and non-science major students benefit from reflective peer assessment over the duration of the study?
Method

This study was conducted in the context of an undergraduate course—history of science. The reflective peer assessment instrument, participants, procedure, and data analysis are described as follows:

Reflective peer assessment instrument

The reflective peer assessment instrument contained six open-ended question items. We developed four items related to gas laws and buoyancy. The four items were validated by three science educators and two physical science teachers who were asked to judge each item by the following criteria:
1. The item examines the conceptual understanding and/or application of (Boyle’s law, Charles’s law, atmospheric pressure, or buoyancy).
2. The item is compatible to the topics that students have previously studied in high school physical science.
3. The item is clearly phrased.

Two more items were derived from the released item bank of Programme for International Student Assessment (PISA) 2006 (OECD, 2007) requiring students to provide evidence-based conclusions. The six-item instrument was pilot tested for one year prior to the study and found to be reasonably reliable (Cronbach $\alpha = 0.72$). In order to check for students’ conceptual development, two items with a similar difficulty level related to gas laws were selected for comparison—one was randomly assigned and assessed at the beginning of the semester while the other item was measured at the end of the course.

In order to promote students’ argumentation discussions, we followed the suggestion of Osborne, Erduran, and Simon (2004) to integrate competing theories into the assessment items. For each item, students are asked to make an assertion or take a position and are encouraged to provide persuasive arguments with appropriate theory, principle, reasonable data or evidence, and supportive warrant or backing. Whenever there are disagreements or different positions of the posted statements, students are encouraged to use rebuttals to challenge the existing statements. One sample item can be seen in the Appendix. For this item, the scientific claim for the first question should be “the mercury moves to the left-hand side”. For the second question of the item, reasonable arguments should be similar to the following explanation containing sound conceptual understanding, basic comprehension of data (relating variables of volume, temperature, and pressure), and the backing (with theories or principles) of an argument: “The air inside the flask enclosed by the mercury is a closed system. In the beginning status of 25° C, its pressure is equal to the surrounding atmospheric pressure. When the set of the flask is moved to the outdoor 5°C environment, the surrounding pressure (i.e., atmospheric pressure) stays the same. Therefore, the volume inside the closed system decreases as the temperature decreases from 25°C to 5°C. This is based on Charles’s Law which states that when pressure is kept constant, a certain amount of gas volume (V) is proportional to the temperature (T) (i.e., $V_1/T_1 = V_2/T_2$). For this question, high quality or advanced arguments even provide rebuttals or counter-arguments to challenge the statements with different claims or predictions.

Participants

The 30 participants (21 males and 9 females) of the study were undergraduate students from colleges of art and humanities, sciences, engineering, management, and social science. They were enrolled in the course entitled History of Science and invited to participate in the online asynchronous discussions. Their ages ranged from 20 to 24 years old. It should be noted that the participants were assumed to have learned all of the relevant content (i.e., Boyle’s law, Charles’s law, atmospheric pressure, greenhouse effect, and buoyancy) related to the six test items in their high school physical science course and no further scientific concepts were instructed during the study. For the purpose of promoting collaborative interactions in solving a given item, the students were randomly divided into six groups. One of the six groups was asked to collaboratively discuss the questions of the given item, reach a consensus, and post their consensual answer on the web. A screenshot of the discussion system is shown in Figure 1. The rest of the students were asked to serve as evaluators for the statement posted by the group.
The role of the evaluator is to reflect on the posted statement, analyze the level of statement using the model developed by Osborne and his colleagues (2004), check the correctness of the statement, assign an appropriate score from 1 to 5 where 1 is the lowest quality while 5 stands for the highest quality of argumentation level for the posted statement, and explain the details of evaluation, reflection, or rebuttals.

Procedure

In the beginning weeks of the course, the instructor explained the difference between scientific arguments and personal opinions, described the elements of an argument with examples, introduced the role of online discussions, and assigned the sequence and date for each group of students to work either as poster (of an item’s answer) or as evaluators. We use Toulmin’s argumentation model (Toulmin, 1958) to introduce the essential elements—data, warrant, backing, and rebuttal and follow the recommendations of Osborne’s research team on the evaluation of arguments (Osborne, Erduran, & Simon, 2004b). In the fourth week, the first group was asked to respond to item 1 and post their consensual answer on the web within one week. During the fifth week, the rest of the students were asked to evaluate the posted answer individually. The evaluators were free to revise their comments according other evaluators’ feedback—allowing for a dynamic process of reflection. In the same time, the group members who posted the answer continued to reflect on the evaluators’ feedbacks and arguments either support or against their answer. At the end of the fifth week, they were asked to present their final answer and justification of their position in a class meeting. For each of the rest of the 5 argumentation items, the procedure was similar to item 1; that is, in the first week the assigned group was responsible for answering the item, while the rest of the students served as reflective evaluators responsible for posting personal comments within the following week. In total, the asynchronous discussions of reflective peer assessment lasted for 12 weeks for the six items.

In order to promote students’ involvement in argumentation, all six items asked students justify their position or explain their reason. In the beginning of the semester, examples of high level arguments were explained to the students for the purpose of scaffolding their ability in argumentation. Through analyzing their arguments in six
rounds of reflective peer assessment, we hope to gain insights that inform subsequent initiatives aimed at a wider application of asynchronous discussion in the development of high level cognitive abilities.

Data analyses

Students’ conceptual understandings were examined and analyzed verbatim based on their reasoning arguments, explanations, and comments. The number of alternative conceptions, the level of conceptual understanding, and the quality level of students’ arguments were used as quantitative indicators.

The scoring scheme for students’ level of conceptual understanding is based on our previous studies (Lin, Cheng, & Lawrenz, 2000; Lin, Hung, & Hung, 2002). The scheme gives 3 points to the answers with correct statements and use of target scientific concepts (e.g., for the sample item, appropriately use the key concepts and identify that in a closed system with constant pressure the gas volume inside the system would change proportionally with the surrounding temperature); 2 points for those answers with sound explanations but minor mistakes (e.g., unable to identify any one of the above key concepts. For instance, fail to explain that the air inside the flask enclosed by the mercury is a closed system); 1 point for the statements showing partial misconceptions but indicating some degree of relevance toward the target concept (e.g., refer to the test item is related to Boyle’s law or Charles’ law but fail to explain how the laws can be used in the item); 0 points for no explanation or explanations with irrelevant statements or misconceptions.

For the scoring scheme on the level of students’ arguments, we follow the method developed by Osborne, Erduran, and Simon (2004), in which, level 1 arguments are arguments with a simple claim without containing any other elements (e.g., I thought that they explained it very well. Without reviewing the rebuttals posted by others, it is difficult for me to identify the weakness of the answer.); level 2 arguments consist of claims with either data, warrants, or backing but do not contain any rebuttals (e.g., this group appropriately used Boyle’s law to explain the mercury movement. However, the lack of real life examples makes it not persuasive); level 3 arguments consist of a series of claims or counter-claims with either data, warrants or backings with the occasional weak rebuttal (e.g., I think there is little difference between the indoor and outdoor air pressure. The major influence for the mercury’s movement should be from the volume change of the enclosed gas.); level 4 arguments consist of a claim with a clearly identifiable rebuttal (e.g., Moving the whole set from 25°C indoors to 5°C outdoors at a constant pressure of 1 atm, the mercury would not move to the right, because Charles’ law tells us that at constant pressure, gas volume would shrink instead of expanding); level 5 arguments contain claims supported by data and warrants with more than one rebuttal (e.g., using theoretical backgrounds or evidences to refute the prediction of mercury moving to right).

Two science educators with the domain knowledge of chemistry and physics evaluated the participants’ argumentation content and statement based on the above scoring scheme. Since item 1 is used as a practice item for students, answers and comments of this item served as examples for the two evaluators to discuss the detail of the scoring procedure. After the discussion, the two evaluators scored items independently. At the end of the scoring procedure, the two evaluators discussed the statements which had discrepancies between their scoring until a consensus was reached. The mean of the two evaluators’ scores was used as the final score for each individual student’s score of the target item.

Result

The first research question of the study was intended to investigate student argumentation ability and conceptual understanding in an asynchronous discussion environment. Meanwhile, the second research question attempted to check the “equity” of the educational opportunity and learning environment provided by the study’s treatment.

Students’ performance of argumentation and conceptual understanding in an asynchronous discussion environment

Table 1 presents the means and standard deviations of the participants’ performance on the five assessment items. Since item 1 was used as practice item for students to be familiar with the procedure of asynchronous discussions, it
was not evaluated for data analysis. It can be seen from Table 1 that students made gradual progress on their argumentation ability from a mean score of 2.58 for item 2 to a mean score of 3.86 for item 6. It should also be noted that the standard deviations decreased from 1.04 to 0.33, which suggests that in the beginning, students’ argumentation ability was much more diversified than it was at the end of the study. In other words, the participants’ argumentation ability had gradually become more homogeneous. The pattern of students’ progress on their conceptual understanding is similar to the progress pattern of argumentation. The mean score improved from 1.98 for item 2 to 2.83 for item 6, while the standard deviation decreased from 0.67 to 0.34.

Table 1. Mean and sd of student performance

<table>
<thead>
<tr>
<th>Item#</th>
<th>mean of argument</th>
<th>sd</th>
<th>mean of conception</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.58</td>
<td>1.04</td>
<td>1.98</td>
<td>0.67</td>
</tr>
<tr>
<td>3</td>
<td>2.87</td>
<td>0.97</td>
<td>2.40</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>3.69</td>
<td>0.47</td>
<td>2.91</td>
<td>0.29</td>
</tr>
<tr>
<td>5</td>
<td>3.36</td>
<td>0.49</td>
<td>2.81</td>
<td>0.40</td>
</tr>
<tr>
<td>6</td>
<td>3.86</td>
<td>0.33</td>
<td>2.83</td>
<td>0.34</td>
</tr>
</tbody>
</table>

a: Since item 1 was used as a practice item for students, it is not used for statistical analysis.

Readers may doubt that the difference of each item’s difficulty level and the variability of content knowledge could affect students’ performance. In order to avoid ambiguity, an attempt was made in designing the study to control the similarity of difficulty level and content knowledge of the assessment items. Three of the five items with similar difficulty level relating to gas laws are randomly assigned as items 2, 3, and 6. Item 2 was assessed in the beginning while item 6 was measured at the end of the study. The statistical analysis of pair-wised t test was conducted to compare students’ performance on the same content knowledge of gas laws. It can be seen from Table 2 that both of students’ argumentation ability and conceptual understanding significantly improved (p< .001).

Table 2. Pair-wised t test result

<table>
<thead>
<tr>
<th>Assessment</th>
<th>pre-test mean (sd)</th>
<th>post-test mean (sd)</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argumentation</td>
<td>2.50(1.15)</td>
<td>3.88(0.33)</td>
<td>5.22***</td>
</tr>
<tr>
<td>Conceptual understanding</td>
<td>1.94(0.68)</td>
<td>2.82(0.35)</td>
<td>5.00***</td>
</tr>
</tbody>
</table>

***: p<.001.

Table 3. Comparisons between science and non-science major students

<table>
<thead>
<tr>
<th>Variable</th>
<th>group</th>
<th>mean(sd)</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argument 2</td>
<td>science major</td>
<td>2.93(0.86)</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>2.23(1.12)</td>
<td></td>
</tr>
<tr>
<td>Argument 3</td>
<td>science major</td>
<td>3.06(1.00)</td>
<td>1.19</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>2.64(0.93)</td>
<td></td>
</tr>
<tr>
<td>Argument 4</td>
<td>science major</td>
<td>3.80(0.41)</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>3.55(0.52)</td>
<td></td>
</tr>
<tr>
<td>Argument 5</td>
<td>science major</td>
<td>3.46(0.52)</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>3.27(0.47)</td>
<td></td>
</tr>
<tr>
<td>Argument 6</td>
<td>science major</td>
<td>3.85(0.34)</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>3.88(0.35)</td>
<td></td>
</tr>
<tr>
<td>concept 2</td>
<td>science major</td>
<td>2.27(0.62)</td>
<td>2.57*</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>1.68(0.61)</td>
<td></td>
</tr>
<tr>
<td>concept 3</td>
<td>science major</td>
<td>2.60(0.51)</td>
<td>3.04**</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>2.10(0.32)</td>
<td></td>
</tr>
<tr>
<td>concept 4</td>
<td>science major</td>
<td>3.00(0.00)</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>2.78(0.44)</td>
<td></td>
</tr>
<tr>
<td>concept 5</td>
<td>science major</td>
<td>2.88(0.35)</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>2.75(0.46)</td>
<td></td>
</tr>
<tr>
<td>concept 6</td>
<td>science major</td>
<td>2.90(0.21)</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>non-science major</td>
<td>2.75(0.46)</td>
<td></td>
</tr>
</tbody>
</table>

*: p< .05.
**: p<.01.
Did both science major and non-science major students benefit from the reflective peer assessment over the duration of the study?

In order to check the equity issue of educational opportunity and learning environment, we investigated the progress pattern and gap difference between science major and non-science majors performance of argumentation and conceptual understanding. The science major students (n=15) were from colleges of science, engineering, and marine science while the non-science major students (n=15) were from colleges of humanity and art, management, and social science. It can be seen from Table 3 that both groups made progress not only on argumentation ability but also on conceptual understanding. The gap of argumentation ability between the two groups starts from a mean difference of 0.70 (2.93 vs. 2.23) in item 2 to a mean difference of 0.03 (3.85 vs. 3.88) in item 6. On the other hand, at the beginning of the asynchronous discussion the conceptual understanding mean scores of the two groups (2.27 vs. 1.68) were significantly different (p<.05) for item 2 and for item 3 (p<.01), but the difference gets smaller and gradually disappears week by week. There is no significant difference between the two group students’ conceptual understanding starting from item 4 to item 6.

Discussion and implication for education

Previous studies found that problem-based learning in an asynchronous discussion environment were not successful for courses in physics (Kortemeyer, 2006) and in statistics (Hong, Lai, & Holton, 2003). Contrarily, in this study, we found that the participants made significant progress (p < 0.001) on their conceptual understanding and application in open-ended problem solving items. Possible reasons for the progress may be attributed to the learning opportunities gained by analyzing the quality level of the online answer statements and reflecting on other students’ posted comments. Through the constant practices of these high level cognitive processes (e.g., analyzing and reflecting) and the continuous involvement in a specific learning topic such as gas laws, students were exposed to a learning environment that enabled them to formulate their own distinct opinion when they receive other students’ rebuttals and criticism. Students integrated plausible content knowledge while contemplating different explanations of others’ and finally constructing their own conceptual understanding and ways of applying these concepts in solving the next test item containing similar content knowledge but in a different context. In typical face-to-face classroom teaching, students are rarely afforded a long time to check and reflect on their own conceptual understanding and application of knowledge when solving problems. It is even harder to compare their own ideas with the ideas of others. The above finding led us to propose the approach of using reflective peer assessment in an online asynchronous environment for students to explore their misconceptions or misunderstandings and further construct understanding of scientific concepts. It is no surprise to us to find that the students made progress on their argumentation ability, since they were taught how to construct persuasive arguments. However, we are impressed by the improvement of the students’ conceptual understanding and problem solving ability in the test items relating to gas laws—considering that they were not given further tutorial on the concepts related to gas laws in the study other then their high school physical science course.

Although additional studies are needed to confirm the practical utility, the initial finding of the diminished difference of conceptual understanding between science majors and non-science majors provides an indicator that the approach of reflective peer assessment has the potential to support student learning, particularly for those who have greater room for improvement (e.g., low achieving or non-science major students). However, this is not to say that this approach is not beneficial to high achieving or capable students. This study also found that students’ majoring in science made progress in their conceptual understanding, as illustrated in Table 3. In total, the reflective peer assessment approach is likely to make positive contribution toward equitable distribution in learning outcome, and not at the expense of capable students. We suspect that the use of reflective peer assessment in asynchronous discussion provides opportunities for students to identify and discuss their alternative conceptions explicitly and publicly, which is helpful and constructive for students’ conceptual understanding (Eryilmaz, 2002). This learning opportunity is rarely seen in traditional classroom teaching. In addition, when students work together collaboratively as a small group to answer the test items, it provides a working environment for them with less pressure than individual written tests. As Frenzel et al. (2007) indicated from their study of 1623 students, there was a close relationship between environmental variables and students’ emotional experiences. Furthermore, higher learning achievement was related to higher enjoyment and lower anxiety. If the conclusion of the above literature is persuasive, then educators and teachers are strongly encouraged to provide a learning environment that allows
students to publicly and explicitly discuss their understanding in their own words without any unnecessary pressure. Meanwhile, opportunities should be provided to students to work in interactive, cooperative, and collaborative teams.

The initial findings of the study shed additional light on the potential benefit of asynchronous discussions in promoting the development of high level cognitive abilities. In using the term “reflective peer assessment,” we intend to highlight the importance of providing opportunities for students to practice “reflective evaluation” and “evaluative reflection” in which the instructor assigns the roles as De Wever et al. (2008) recommends to students who take turns to serve as “answer provider” or “reflective evaluator.” The answer providers are encouraged to work collaboratively within their groups to reach a consensual conclusion in an assessment item with competing theories. In order to construct a persuasive conclusion, the students have to provide data and evidence and use warrants and backings to support their conclusions based on Toulmin’s model (Toulmin, 1958) that was introduced to them in the beginning of the class. In addition, each evaluator is responsible to exercise reflective evaluation, assign a score, and provide personal comments to the posted answer using the scoring scheme of Osborne et al. (2004). With the practice of assigning a score and writing comments, each student is exposed to the learning environment of practicing “reflective evaluation” (i.e., reflecting on the answer and then executing the evaluation). They are expected to learn how to provide counterarguments or rebuttals that disagree with existing arguments, or learn how to explain why they support a certain conclusion. Meanwhile, the asynchronous learning environment provides opportunities for students to review and reflect on other evaluators’ comments. Being evaluators, they are allowed to revise their comments and the original scores they assigned to the answer provider. In this stage, they are encouraged to practice “evaluative reflection” (i.e., evaluating other students’ comments and reflecting on their own comments). The evaluative reflection encourages students to observe others’ comments and critics. Based on McKendree’s (1998) assertion, observing a dialogue is beneficial for learning, especially when it is combined with reflection.

Despite the fact that the initial finding of the study is impressive and encouraging, readers are reminded that the sample size is relatively small. Therefore, caution should be taken in inferences of its quantitative results. In addition, care must be taken by making inferences from the research design of one group pretest and post-test which is not a true experimental design. In this study, since the participants have learned the content of gas laws in their high school years, no further gas law concepts were taught. The treatment was mainly used to help students clarify alternative conceptions and apply appropriate scientific concepts in contextual problem-solving situations through reflective analysis and criticism of peer answers and argument. During this period of time, the participants were not likely to have other learning resources in the specific topic of gas laws except the treatment. Therefore, the major potential threats of internal and external validity of this design can be reasonably avoided by selecting a learning topic where students (even some science teachers) have deep-rooted alternative conceptions (i.e., resistant to conceptual change) (Authors, 2000). Additionally, a longer period of interactive and collaborative dialogical reflections and argumentation would allow students to explicitly discuss and find the conflict between their own alternative conceptions and scientific argument. Further research studies focusing on different topics or subject matters of content knowledge with bigger sample sizes are strongly recommended.

References


Appendix: Sample item of the assessment instrument

As shown in the following figure, an empty flask is sealed with a rubber stopper which includes a glass tube. At the end of the glass tube, there is a drop of mercury. When the flask is immersed in a beaker filled with water of 3°C, the mercury will move to the left. On the other hand, when the flask is immersed into a beaker filled with water of 80°C, the mercury moves to the right. If the whole set in the figure (not including the beaker) is moved from an indoor temperature of 25°C to an outdoor one of 5°C, can you predict and explain the movement of the mercury?