Computers & Education 55 (2010) 586–602

Contents lists available at ScienceDirect

Computers & Education

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

On-line synchronous scientific argumentation learning: Nurturing students' argumentation ability and conceptual change in science context

Kuan-Hue Yeh, Hsiao-Ching She*

Institute of Education, National Chiao Tung University, 1001 Ta-Hsueh Rd., Hsin Chu City 300, Taiwan, ROC

ARTICLE INFO

Article history:
Received 15 April 2009
Received in revised form
24 February 2010
Accepted 24 February 2010

Keywords:
Improving classroom teaching
Secondary education
Cooperative/collaborative learning

ABSTRACT

The purpose of this study is to examine the difference in effectiveness between two on-line scientific learning programs – one with an argumentation component and one without an argumentation component – on students' scientific argumentation ability and conceptual change. A quasi-experimental design was used in this study. Two classes of 8th grade students (the experimental group) received the on-line scientific argumentation learning program about chemical reaction, and the other two classes of 8th grade students (the control group) received the same on-line scientific learning program about chemical reaction, but without argumentation, for two weeks. All 140 students were administered the scientific conception test, conceptual change test, and argumentation test before, one week after, and eight weeks after learning. In addition, the experimental group students' on-line argumentation process was collected. Results showed that the students of the experimental group significantly outperformed the control group, regardless of scientific conceptions, conceptual change, and argumentation. Regression results indicated that hold of scientific conceptions is the best predicator for students' conceptual change, followed by argumentation ability. The quantity and quality of scientific arguments that students generated in a series of argumentation questions improved across the four topics. In addition, students also successfully changed their conceptions from pre- to post-driving questions across four topics. This clearly demonstrates that students' argumentation ability and conceptual change were both facilitated through receiving the on-line Synchronous Argumentation science learning program.

1. Introduction

Argumentation is considered to be a major component of science education by the national science standards (National Research Council, 1996) and by many science educators (Driver, Newton, & Osborne, 2000; Duschl & Osborne, 2002; Osborne, Erduran & Simon, 2004). Kuhn (2005) also proposed that the goal of science education is to promote a way of thinking in which argumentation serves as one of the central skills (Kuhn, 2005). Similarly, many studies advocate the need to educate students about how they know and why they believe in the scientific worldview, instead of merely dealing with what they know (Driver, Leach, Millar, & Scott, 1996; Millar & Osborne, 1998; Osborne et al., 2004). PISA 2006 Science Competencies for Tomorrow's World specifically pointed out that students should be able to analyze the sufficiency of the data, make judgments about the validity of conclusions, compare and differentiate among competing explanations by examining supporting evidence, formulate arguments by synthesizing evidence from multiple sources, and construct a logical argument for a hypothesis by using data from a number of sources (OECD, 2007, vol. 1, p. 101). Eemeren and Grootendorst (2004) defined argumentation as knowledge justification and persuasion. In summary, argumentation in science learning can be defined as a process of connecting claims and data through justification or through the evaluation of knowledge claims in light of empirical or theoretical evidence (Jimenex-Aleixandre & Erduran, 2008, p13). Kuhn also suggested parallel ideas that science education should be taught as promoting a way of thinking and not as a body of knowledge and fixed facts (Kuhn, 1993). Osborne et al. (2004) further pointed out that such a shift requires a new focus on how evidence is used in science for the construction of explanations, that is, on the arguments that form the links between data and the theories that science has constructed. More specifically, the construction of arguments is a core discursive activity of science (Osborne et al., 2004). Furthermore, Driver et al. (2000) and Newton, Driver, and Osborne (1999) suggested that the practices of

---

* Corresponding author. Tel.: +886 0 35731665; fax: +886 0 35738083.
E-mail address: hcshe@mail.nctu.edu.tw (H.-C. She).

0360-1315/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compedu.2010.02.020

argumentation in the discipline are critical for students' knowledge construction. Several studies support the notion that addressing argumentation skills as a distinct educational goal indeed fosters students' argumentation ability (Osborne et al., 2004; Zohar & Nemet, 2002); improves students' scientific knowledge (Aufschnaiter, Erduran, Osborne, & Simon, 2008; Zohar & Nemet, 2002); and makes students' scientific thinking and reasoning more visible (Bell & Linn, 2000; Chinn & Anderson, 1998).

From a constructivist perspective, learning is a process in which an individual is actively involved in linking new ideas with current ideas and experiences. It is basically the nature of conceptual change (Cobern, 1993). Conceptual change has been typically thought of as a rational process during students' reasoning from their initial conceptions with available evidence, given a particular conceptual ecology, to a new set of conceptions (Posner, Strike, Hewson, & Gertzog, 1982). As Strike and Posner (1985) pointed out, learning is primarily a task relating what one has encountered to one's current ideas. The students who learn something are the ones who understand a new idea, judge its truth value, judge its consistency with other ideas, and are willing to change their minds to accept it. West and Pines (1983) criticized Posner et al.'s (1982) conceptual change as being overly rational, and Pintrich, Marx, and Boyle (1993) consider it to be overly dependent on cognitive resources. Their comments reveal a call to introduce a social element into the process of conceptual change. Much research indeed has included social perspectives in their conceptual change models or approaches (She, 2004a; Driver & Oldham, 1986; Savinainen, Scott, & Viiri, 2005; Venville & Treagust, 1998).

Engaging in argumentation can also promote conceptual change in the minds of scientists debating a new finding or theory (Dole & Sinatra, 1998; Driver et al., 2000) or in the minds of students asked to consider alternative points of view and evaluate alternative conceptions (Nussbaum & Sinatra, 2003). Argumentation is expected to not only allow students to consolidate existing scientific knowledge, but to also construct new knowledge for themselves based on the ideas of others (Brown & Campione, 1998). When students argue and are exposed to different ideas, they can reflect on their own ideas and ideas of others, aiding them in addressing alternative conceptions and developing better understandings (Cross, Taasoobshirazi, Hendricks, & Hickey, 2008). Driver (1995) indicated that whether or not an individual's ideas are affirmed and shared by others in classroom exchanges affects how the knowledge construction process is shaped. Basically, it is social constructivism by nature. Solomon argued that student ideas about nature stem “not from the logical processes of which science boasts, but from the ‘common sense” attitude that relies on being able to interchange perspectives and meanings with others” (1987, p. 66). The notion of “interchange” introduces a social element. Cobern (1993) also brings the sociology of science into learning, drawing attention to the importance of inter-conceptual issues in addition to intra-conceptual learning issues. All of these imply the need for, and potential of, social constructivism. From the aspect of sociology, examination of the nature of argumentation clearly indicates its potential for contributing to the collective development and judgment of scientific knowledge claims and the identification of reliable and consensual descriptions of nature (Kolsto & Ratcliffe, 2008, p. 117). Students will receive the opportunity to experience how their ideas are questioned, revised and changed as real scientists. Studies have indicated that engaging students in classroom communicative discourses around the use of computers in curriculum-related activities improves the quality of talk for construction of shared knowledge (Mercer, Littleton, & Wegerif, 2004; Wegerif & Mercer, 1996).

A number of studies show that engagement in argumentation improves conceptual understanding in science (Duschl, Ellenbogen, & Erduran, 1999; Zohar & Nemet, 2002). Though many studies have suggested that argumentation may promote students’ conceptual change, only a few studies have proven to promote conceptual change in university students (Nussbaum & Sinatra, 2003), or conceptual change involving physics conceptions of force and motion in secondary school students through the use of digital argumentative dialogue games (Ravenscroft, 2000, 2007; Ravenscroft & Pilkington, 2000). These studies shed the lights on the possibility of using argumentation skills to foster students' conceptual change. Obviously, argumentation has great potential for fostering students' communication skills in order to interchange perspectives and meanings. Assessing alternatives, weighing evidence, interpreting texts, and evaluating the potential validity of scientific claims are all seen as essential components in constructing scientific arguments. Unfortunately, science teaching and learning has paid too little attention to argumentation, or to enhancing conceptual change through argumentation. The major barrier to the development of students' scientific argumentation skills is the lack of opportunity to practice these types of activities within daily life or current classroom teaching. Therefore, integrating scientific argumentation into classroom science curriculum learning would be a great challenge. Therefore, this study seeks to empower students with the ability to argue scientifically and thereby change their alternative conceptions through argumentation.

2. On-line synchronous argumentation

Several studies suggest that on-line learning environments can provide excellent support for argumentation in the classroom (Andriessen, Erkens, Van de Laak, Peters, & Coirier, 2003; Clark, & Sampson, 2008; Sandoval & Reiser, 2004). Many studies have focused on developing on-line argumentation environments, and they include a broad range of specific instructional features to promote productive interactions between participants. Several studies have included the synchronous and asynchronous collaborative communication interfaces to support students' interaction. Asynchronous modes of communication allow learners to participate more equitably and to spend more time constructing well-conceived and complex arguments (Schellens & Valcke, 2006). Synchronous communication can deliver a higher degree of elaboration and construction of arguments as students work on a common shared artifact (De Vries, Lund, & Baker, 2002; Janssen, Erkens, & Kanselaar, 2006). Similar studies which included synchronous digital dialogue games indicated that this supports students in practicing higher-order thinking and discussion (Ravenscroft, Mcalister, & Baur, 2006), and also stimulates their thinking and conceptual development (Ravenscroft, Wegerif, & Hartely, 2007). Our study specifically designed a synchronous argumentation environment, centered on Web based learning, to provide students with the opportunity to argue with students in their group in real time.

On-line argumentation provides the advantage of allowing students to see arguments and counterarguments on the screen, which supports them in refining their argumentation (Kirschner, Buckingham Shum, & Carr, 2003). Learners can construct arguments to justify their position, which facilitates self-explanation of the learning material (Baker, 2003). Their learning partners construct counterarguments to challenge and reconsider these positions. Counterarguments assist learners in rethinking their initial argument (Leitao, 2000), constructing new conceptions, and refining their initial position. Constructing a good argument is not a simple task and students need guidance and support in order to scaffold and build their sense of an effective argument (Osborne et al., 2004). Some studies indicated that providing students with scaffolding tools, such as interactive digital dialogue game tools, can successfully facilitate their development of argumentative dialogue (Ravenscroft, 2007; Ravenscroft et al., 2007). Wray and Lewis (1997) have shown that when such genres of writing or expression are not familiar "writing frames" that support the process of writing can provide vital support and clues as to what is needed. Therefore, a set of stems such as "I think/believe…because…; the reason I agree with…argument, is because the evidence of…; I do not agree with…; my reason is…" were provided under each component of data, claim, warrant, backing and rebuttal in the study. Osborne et al. (2004) indicated that stems provide students with prompts to construct their argument in a coherent manner and writing frames can then be used as a structure for producing a written argument.

Cho and Jonassen (2002) found that undergraduate students using the software argumentation template provided increased claims about how to solve the problems, which indicated that the scaffold actually enhanced the students' argumentation skills. Nussbaum's study also indicated that an argumentation template based on Toulmin's (1958) argumentation theory helped sixth-grade students generate more complete and explicit arguments in social studies classrooms after some practice in using the scaffolds. Another study indicated that a six-page software argumentation template was helpful for critical reasoning about multimedia-supported history problem-solving among 11th grade students (Saye & Brush, 2002). Li and Lim (2008) showed that students' historical inquiry skills can be enhanced through scaffolding students with an argumentation template. Therefore, we specifically programmed our learning environment to provide students with writing frames of five argumentation components in order to scaffold their arguments in science learning.

The Toulmin's (1958) model of argumentation was used in our study for constructing our on-line argumentation template. Toulmin's model contains five essential components which are data, claim, warrant, backing, and rebuttal. Data is the observation or fact that students had in their statements. Data describes the fact or phenomena or experiment that the student already knows or sees. Claims are arguments to explain why the student thinks or believes such an argument, which is based upon what the student already knows or sees about the fact or phenomena or experiment. Warrants provide a theory or principle or reason to support or justify the connection between a particular data and claim. Warrants show that the move from data to claim is valid. Backing provides evidence, experience or an experiment the student has known/believed/conducted, together with an explanation of why it would support their or others' arguments. Backing for warrants may be expressed as statements of fact or data which are appealed to in direct support of a claim. Rebuttals provide evidence, experience, an experiment or a theory that the student has known or believed or conducted, together with an explanation of why it would not support their or others' arguments. In short, rebuttals provide evidence to contradict or nullify other evidence that has been presented.

3. Analytical framework of argumentation discourse

Osborne and his colleagues have developed an analytic framework to assess the nature of argumentation, developed from the Toulmin's argumentation pattern (Osborne et al., 2004). It became one of the major existing analytic frameworks for science educators to use in their studies. Their analytic framework contains a set of five levels of argumentation. We modified their analytic frameworks in order to assess the quality and quantity of argumentation for each statement produced by each individual student. Therefore, each statement produced by each individual would be assessed by their levels and types of argumentation. For the quality of argumentation, each statement would be classified into two different levels of claim, warrant, backing and rebuttal, respectively. Data is considered to be non-argumentative statements. There are two levels of claim: A level 1 claim is an argument consisting of a claim without any data or fact. A level 2 claim is an argument consisting of a claim with data or fact. There are two levels of warrant: A level 1 warrant is an argument consisting of a theory or principle without connection to the claim, or one which does not clearly describe the theory. A level 2 warrant is an argument consisting of a claim with a clearly described theory or principle. There are two levels of backing: a level 1 backing is an argument consisting of a backing without any connection to claim/warrant, or one which does not clearly describe the connection among them. A level 2 backing is an argument consisting of a claim with backing, or with data or warrant. There are two levels of rebuttal: a level 1 rebuttal is an argument consisting of a weak rebuttal or one without a clear explanation. A level 2 rebuttal is an argument consisting of a claim with a clearly identifiable rebuttal. We hoped that each student would be able to use clear statements to provide a concise argument consisting of complete information. For example, regarding claim, they needed to provide a claim with data or fact or evidence within a short statement. In short, each statement would be classified as a non-argumentative statement, level 1 or level 2 claim, warrant, backing, or rebuttal, respectively. The analytical framework we developed above was intended to evaluate the nature of discourse and its quality and quantity of arguments (Table 1).

The studies presented earlier indicated that argumentation seems very promising in promoting students' conceptual understanding in science and argumentation skills. Other on-line argumentation research shows positive results concerning enhancing students' argumentation skills in disciplines other than science. We are very much interested in whether or not students' conceptual change and argumentation skills would improve significantly while receiving our on-line argumentation as compared to the group receiving no argumentation. In addition, the studies presented earlier in this paper also indicated that providing both synchronous and argumentation templates has a great potential for enhancing students' ability to construct more complete and coherent arguments, and that it promotes more construction of arguments with their group in real time. Therefore, we specifically examined, throughout the learning process, the quality and quantity of argumentation on students who received the on-line argumentation program, in order to determine its impact. A significant deficit in the literature remains in empirical research on the relationships between argumentation skills and conceptual change, thus that relationship was examined in this study. In short, the purpose of our study is to explore whether or not students' conceptual change and their argumentation ability would improve over time through the On-line Scientific Argumentation program.

4. Purpose

One of the aims of this study is to report the development of an on-line synchronous scientific argumentation learning environment with argumentation templates. A second aim is to replace regular physical science learning with on-line synchronous scientific argumentation learning for about two weeks. A third aim is to explore the effectiveness between an on-line scientific learning program with and without argumentation component on the dimension of conceptual change, argumentation and academic achievement. A fourth aim is to examine, through mixed methods, the effects on two classes of 8th grade students (the experimental group), who received the on-line argumentation learning program about chemical reaction, as compared to two other classes of 8th grade students (the control group), who received the

Table 1
Analytical framework used for determining the quality of argumentation.

Components	Levels	Definition	Examples
Claim	Level 1	An argument consists of a claim without any data or fact.	The greater the concentration, the faster the reaction.
	Level 2	An argument consists of a claim with data or fact.	I saw that the greater the concentration of HCl, the faster the reaction with marble. Thus I think that the greater the concentration, the faster the reaction.
Warrant	Level 1	An argument consists of a theory or principle without connection to the claim, or one which does not clearly describe the theory.	The more molecules, the greater the opportunity for collision.
	Level 2	An argument consists of a claim with theory or principle.	The greater the concentration, the faster the reaction. It is because the more molecules, the greater the opportunity for collision.
Backing	Level 1	An argument consists of a backing without any connection to claim/warrant, or one which does not clearly describe the connection among them.	I agree with David's idea, because I had a similar experience that producing oxygen experiment with high concentration of hydrogen peroxide.
	Level 2	An argument consists of a claim with backing, and or with data or warrant.	I support Ann's idea, because I have done the concentration experiment (HCl react with marble), which proves that the greater the concentration, the faster the reaction. So there is greater intensity of the molecular collisions.
Rebuttal	Level 1	An argument consists of a weak rebuttal and without clear explanation.	I do not agree with Thomas's idea, because some people who drink high concentration wine do not get drunk at all.
	Level 2	An argument consists of a claim with a clearly identifiable rebuttal.	I disagree with Jim's idea that the lower the concentration, the faster the reaction. The lower the concentration, the smaller the amount of molecules, thus the lower the opportunity for collision.

same on-line learning program without argumentation. A fifth aim was to examine the effectiveness between the learning outcomes of the two groups and relationships among argumentation ability, conceptual change, and scientific conceptions, therefore, three quantitative instruments of Chemical Reaction Achievement Test (CRAT), two-tier Chemical Reaction Conceptual Change Test (CRCCT) and Chemical Reaction Dependent Argumentation Test (CRDAT) were developed. Finally, a sixth aim was to collect the experimental group's students' on-line scientific argumentations, in order to determine the quality and quantity of students' argumentation and conceptual change in the web learning. This allowed us to see whether or not students' argumentation and conceptual change would be improved as time went on.

5. Development of the on-line synchronous scientific argumentation program

The On-line synchronous scientific argumentation program was developed based on the theories of argumentation and conceptual change. A panel of six physical science teachers and science educators were involved in developing the chemical reaction on-line argumentation learning materials. One group of four computer science graduates met with science educators every two weeks for about one year to finish the on-line synchronous scientific argumentation prototype, followed by a year of testing and modification.

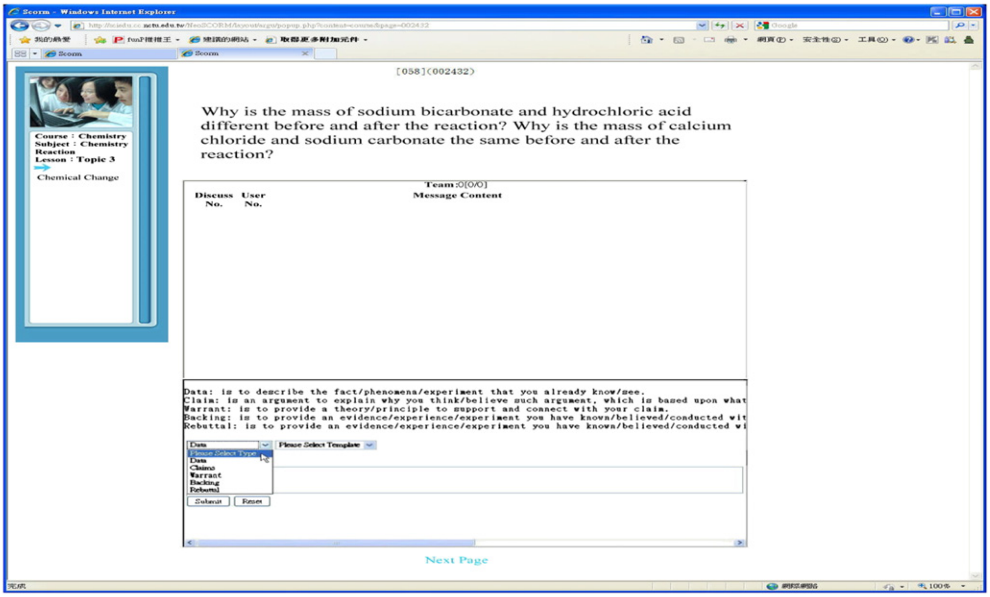
5.1. Characteristics of the on-line synchronous scientific argumentation program

5.1.1. Facilitate students' argumentation ability

In order to promote students' argumentation ability, this on-line synchronous scientific argumentation learning environment has tools specifically designed for students to use while they are participating in argumentation. In order to facilitate students' ability to produce a good written argument, our interface specifically designed two layers of templates for students to use. The first layer provides the definition and choices of five components of argumentation: data, claim, warrant, backing, and rebuttal (Fig. 1); the second layer provides 3–4 writing frames for each component of argumentation (Fig. 2). The writing frame provides guidance and support to help students construct a good argument. The following stems were provided: “I think/believe…, because…; the reason why I agree with…argument, is because the evidence of…; I do not agree with…, my reason is…” Students need to choose one of the components of argumentation first and then choose one from 3 to 4 writing frames that they feel appropriate to share their argument. The learning environment provides the advantage of real time argumentation, so students can receive prompt rebuttals to their arguments. This better retains their interest and makes learning more effective. Moreover, students are able to see arguments and counterarguments on the screen, which supports them in refining their argumentation in real time. Students can be allocated into heterogeneous or homogenous groups, and into different sized groups depending on the learning purposes. This study mixes students into both heterogeneous and gender mixed groups.

5.1.2. Facilitate students' conceptual change

To facilitate students' conceptual change, the design of the on-line synchronous scientific argumentation learning environment requires students to first provide a personal answer to the driving question. The driving question is to help stimulate learners to synthesize different types of information and activities they have learned and known. This is then followed by group argumentation. The purpose of argumentation is to create an opportunity for students to face the conflict. During this stage, students argue with their peers and many alternative ideas and disagreements appear. After argumentation, the same driving question was asked in order to see whether each individual student had reconstructed and changed their alternative conceptions throughout the argumentation. After the driving question, a series of well-designed multimedia learning materials, such as demonstrations, videos, simulation and animation, were presented. This provided students with the opportunity to actually see what happens and to directly encounter the dissonances. The intention is to provide a possible and plausible mental structure as a platform for them to reconstruct their conceptions. This process is intended to encourage students to reconstruct their knowledge through a series consisting of the driving question, the argumentation question, and active learning materials.



[058](002432)

Why is the mass of sodium bicarbonate and hydrochloric acid different before and after the reaction? Why is the mass of calcium chloride and sodium carbonate the same before and after the reaction?

Team:0[0/0]

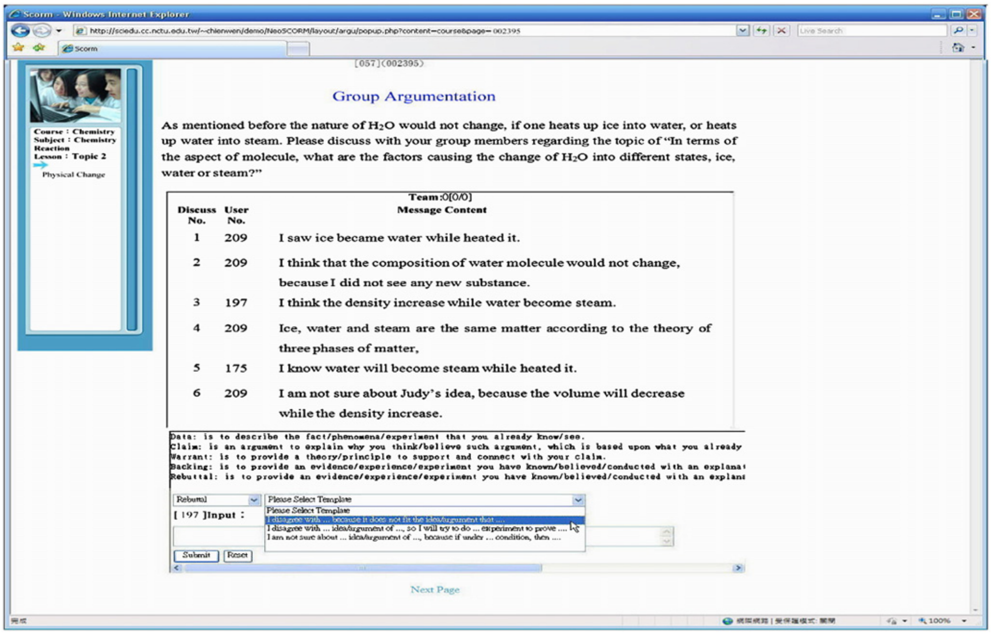
Discuss No.	User No.	Message Content

```
Data: is to describe the fact/phenomena/experiment that you already know/see.
Claim: is an argument to explain why you think/believe such argument, which is based upon what
Warrant: is to provide a theory/principle to support and connect with your claim.
Backing: is to provide an evidence/experience/experiment you have known/believed/conducted wit
Rebuttal: is to provide an evidence/experience/experiment you have known/believed/conducted wi
```

Fig. 1. Web page of argumentation templates.

5.2. *Design of the on-line synchronous scientific argumentation learning content: Unit of chemical reaction*

The on-line synchronous scientific argumentation program was designed to replace the regular physical science course in middle school. Therefore, the topics concerning chemical reaction were chosen from the current middle school physical science mandatory content and standard. These topics consist of the changes of matter, physical change, chemical change, and chemical formula. Each topic is specifically designed to initiate a driving question for an individual student to answer at the beginning. The driving question is to help stimulate learners to synthesize different types of information and activities they have learned and known. This evaluates their initial state of knowledge. This is followed by group argumentation, during which students get into groups to focus on the argumentation question and to develop their arguments. After argumentation, the same driving question is asked in order to see whether each individual student reconstructed and changed their alternative conceptions. Then a series of well-designed multimedia learning materials, such as hands-on demonstrations, videos, simulation and animation was provided. This was purposely designed to help students reconstruct and change their alternative conceptions through active engaging in learning activities. The learning content and sequence for the on-line learning group is the same as the on-line argumentation learning group except that it does not include argumentation nor does it include the post-driving question. Both groups' students were required to answer the driving question before learning.



[057](002395)

Group Argumentation

As mentioned before the nature of H_2O would not change, if one heats up ice into water, or heats up water into steam. Please discuss with your group members regarding the topic of "In terms of the aspect of molecule, what are the factors causing the change of H_2O into different states, ice, water or steam?"

Team:0[0/0]

Discuss No.	User No.	Message Content
1	209	I saw ice became water while heated it.
2	209	I think that the composition of water molecule would not change, because I did not see any new substance.
3	197	I think the density increase while water become steam.
4	209	Ice, water and steam are the same matter according to the theory of three phases of matter,
5	175	I know water will become steam while heated it.
6	209	I am not sure about Judy's idea, because the volume will decrease while the density increase.

```
Data: is to describe the fact/phenomena/experiment that you already know/see.
Claim: is an argument to explain why you think/believe such argument, which is based upon what you already
Warrant: is to provide a theory/principle to support and connect with your claim.
Backing: is to provide an evidence/experience/experiment you have known/believed/conducted with an explana
Rebuttal: is to provide an evidence/experience/experiment you have known/believed/conducted with an explan
```

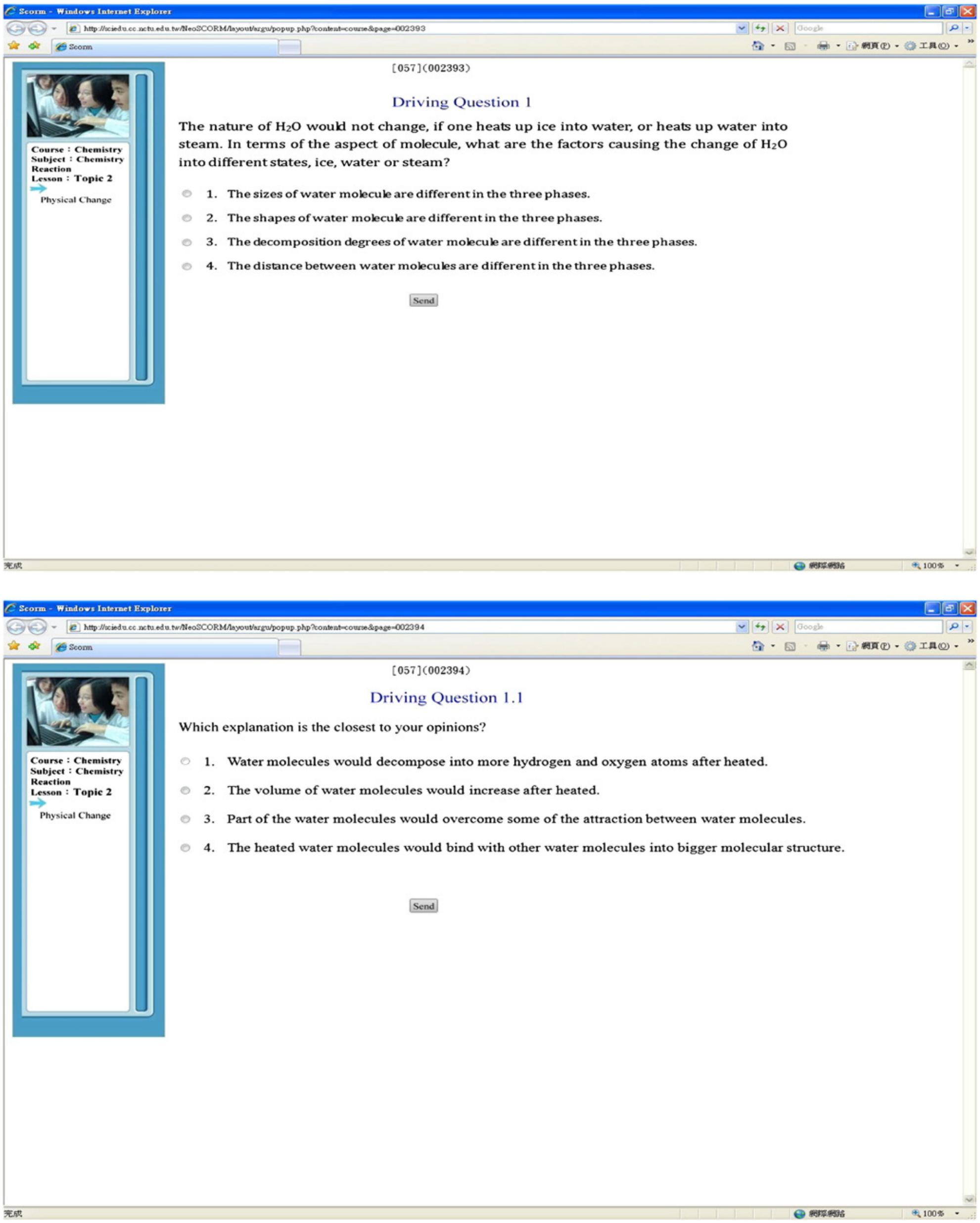
[197]Input :

Please Select Template
I disagree with ... because it does not fit the idea/argument that ....
I disagree with ... idea/argument of ..., so I will try to do ... experiment to prove ....
I am not sure about ... idea/argument of ..., because if under ... condition, then ....

Fig. 2. Web page of students' argumentation.

5.2.1. Example of argumentation learning content

The following example activity in the physical change topic was provided to show how conceptual change and argumentation theories were implemented in the design of on-line synchronous scientific argumentation learning content involving the microscopic viewpoint of changing the states of water. First, the driving question was initiated for individual students to choose their answer and reasons: The nature of H_2O does not change, if one heats ice into water, or heats water into steam. In terms of the aspect of molecules, what are the factors causing the change of H_2O into the different states of ice, water and steam (Fig. 3)? A similar question was given for group argumentation, during which students formed groups to focus on the argumentation question and to develop their arguments (Fig. 2). After argumentation, the same driving question was asked of individual student in order to see whether or not they had reconstructed and changed their alternative conceptions. Immediately after, feedback in the form of the correct answer was provided. After this, a series of activities was provided for students to construct a more scientific viewpoint of the conceptions: (1). Connecting a macroscopic representation of ice with a microscopic representation through animation. A microscopic representation of ice animation was presented to show students that the arrangement of the water molecules of the ice is ordered. The attractions between molecules are strong, so the water molecules cannot move easily. The water molecules of the ice can only vibrate in place, therefore, the shape and volume of the solid phase of water (ice) are fixed. (Fig. 4) (2). An animated microscopic representation of H_2O changing state from ice to water was presented to show students the process and change of the arrangement and movement from the ice molecule to the water molecule. The water molecules in liquid move fast enough to overcome some of the attractions between them during the process of ice melting into water. The particles are able to slide past one another. (3) Connecting a macroscopic representation and a microscopic representation of heating water to steam through animation.



[057](002393)

Driving Question 1

The nature of H_2O would not change, if one heats up ice into water, or heats up water into steam. In terms of the aspect of molecule, what are the factors causing the change of H_2O into different states, ice, water or steam?

○ 1. The sizes of water molecule are different in the three phases.
○ 2. The shapes of water molecule are different in the three phases.
○ 3. The decomposition degrees of water molecule are different in the three phases.
○ 4. The distance between water molecules are different in the three phases.

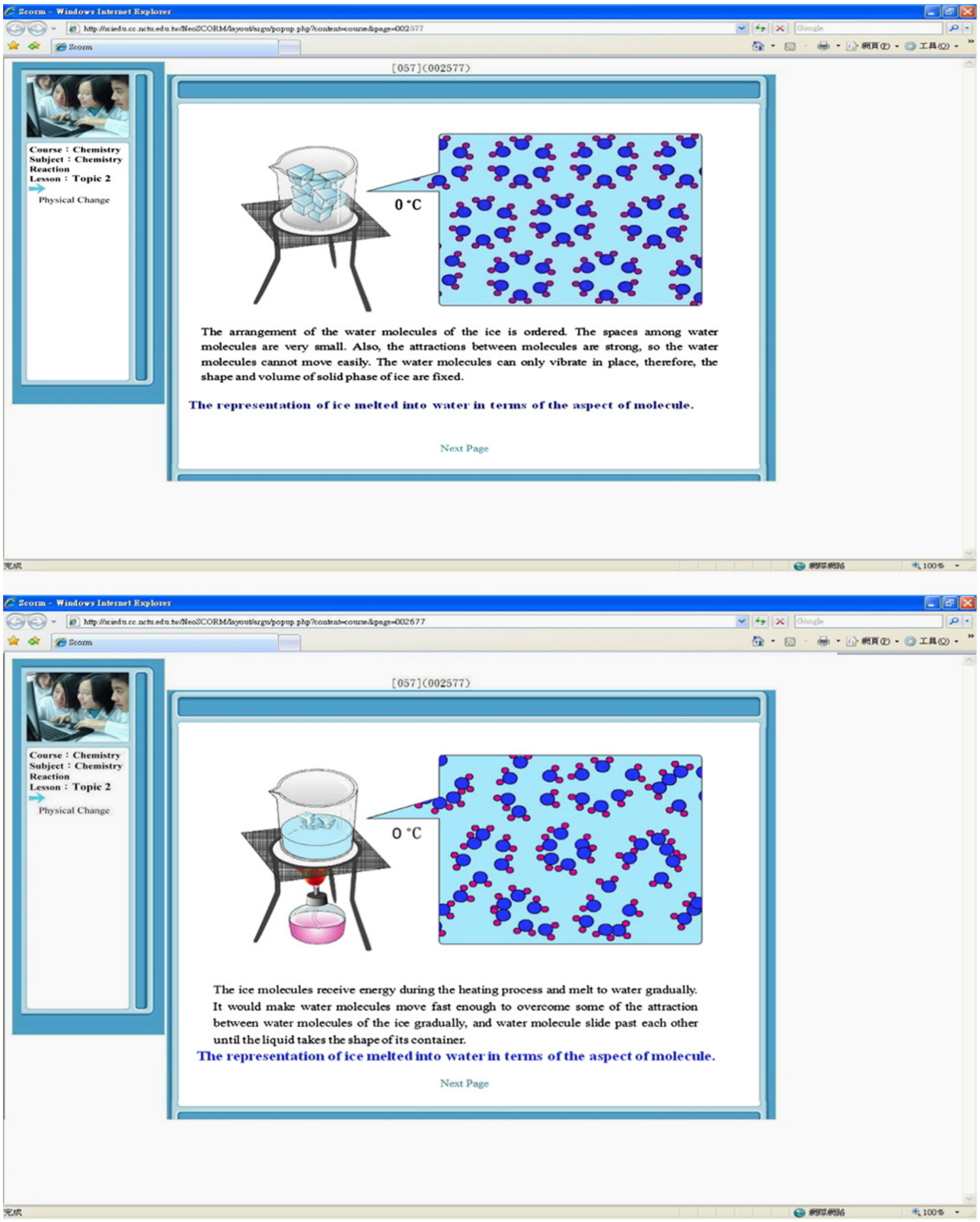
[057](002394)

Driving Question 1.1

Which explanation is the closest to your opinions?

○ 1. Water molecules would decompose into more hydrogen and oxygen atoms after heated.
○ 2. The volume of water molecules would increase after heated.
○ 3. Part of the water molecules would overcome some of the attraction between water molecules.
○ 4. The heated water molecules would bind with other water molecules into bigger molecular structure.

Fig. 3. Web page of driving question.



The arrangement of the water molecules of the ice is ordered. The spaces among water molecules are very small. Also, the attractions between molecules are strong, so the water molecules cannot move easily. The water molecules can only vibrate in place, therefore, the shape and volume of solid phase of ice are fixed.

The representation of ice melted into water in terms of the aspect of molecule.

The ice molecules receive energy during the heating process and melt to water gradually. It would make water molecules move fast enough to overcome some of the attraction between water molecules of the ice gradually, and water molecule slide past each other until the liquid takes the shape of its container.

The representation of ice melted into water in terms of the aspect of molecule.

Fig. 4. Animation of ice melting into water.

The detailed, animated, microscopic representation of H_2O changing state from water to steam was presented to show students that the boiling water would receive energy during the heating process, which would make the water molecules move fast enough to overcome some of the attraction between the water molecules gradually, allowing the water molecule to free itself and becomes steam. The representation was purposely designed to help students reconstruct and change their alternative conceptions through active involvement in learning activities.

6. Method

6.1. Subjects and procedures

A total of 140 eighth grade students, recruited from four classes of a middle school, participated in this study. Students were randomly assigned into an experimental group (71 students) and a control group (69 students). The experimental group's students were further divided into 14 groups, with an average of five students assigned to each group. The experimental group received on-line argumentation learning lessons about chemical reaction for two weeks – about ten classes (each class-period is about 45 min).

The control group received the same on-line content without argumentation or the post-driving question. All students were administered the two-tier Chemical Reaction Conceptual Change Test (CRCCT) and the Chemical Reaction Dependent Argumentation Test (CRDAT) before, one week after, and eight weeks after learning. In addition, the qualitative data of the experimental group students' on-line argumentation process were collected. Those four classes were taught by the same physical sciences teacher. All 140 students received the introduction of five components of argumentation and engaged in argumentation activities in science classes for three class-periods before they received the two different Web-learning programs and the assessment.

6.2. *Instruments*

6.2.1. *Chemical reaction achievement test (CRAT)*

The CRAT is a multiple choice diagnostic instrument that was developed to measure students' chemical reaction-related conceptions before, directly after, and eight weeks after receiving the chemical reaction topic of the digital learning program. The content validity was established by the same panel of eight evaluators, ensuring that the items were properly constructed and relevant to the Web-learning materials we developed. There are eight items for each topic, for a total of 32 items. Students receive one point for each question they answer correctly, so the highest score is 32. The Cronbach's α of the CRAT was .74 for the pre-test, .79 for the post-test, and .84 for the retention-test.

6.2.2. *Chemical reaction conceptual change test (CRCCT)*

The CRCCT is a two-tier multiple choice diagnostic instrument that was developed to measure the degree of students' conceptual change involving chemical reaction conceptions before, one week after, and eight weeks after receiving the on-line science learning program. The content validity was established by the same panel of six evaluators, ensuring that the items were properly constructed and relevant to the seven topics of physical science Web-learning materials that we developed. There are 32 items and each item has two tiers. There are eight items for each topic, and each item contains two tiers. In the first tier, students choose the correct scientific concepts, and in the second tier they choose the correct reason for the specific chemical reaction concepts. Students need to answer both tiers of each question correctly in order to receive one point, so the highest score can be 32. The Cronbach α of ADRT was .81 for the pre-test, .90 for the post-test, and .92 for the retention-test.

6.2.3. *Chemical reaction dependent argumentation test (CRDAT)*

The CRDAT is a two-tier multiple choice diagnostic instrument that was developed to measure the degree of students' argumentation ability involving chemical reaction-related concepts before, one week after, and eight weeks after receiving the four topics of the digital learning program. Students were not familiar with the terms of data, claim, warrant, backing and rebuttal in their daily life. We introduced students to what we considered to be a complete claim, warrant, backing and rebuttal argument; that is, our level 2 arguments. According to our observations, students learned argumentation within the three class periods; however, it was still very difficult for them to identify these five components from their own statements or from group discourses. Therefore, the test specifically measured students' ability to identify data, claim, warrant, backing, and rebuttal from a set of discourses. We believed that if students are able to determine which statements can be considered to be complete claim arguments, warrant arguments, backing arguments or rebuttal arguments, they are more likely to use these appropriately in their dialogical discourse.

There are five scenarios, covering four topics. There are five questions under each scenario, for a total of 25 questions. Each question contains two tiers. The first tier of each question is for identifying a specific statement from a set of discourses as data, claim, warrant, backing, and rebuttal, respectively. The second tier is for choosing an appropriate explanation for categorizing the specific statement as data, claim, warrant, backing or rebuttal, respectively. The content validity was established by the same panel of six evaluators, ensuring that the items were properly constructed and relevant to the four topics of the chemical reaction digital learning program we developed. Students need to answer both tiers correctly in order to receive one point, so the highest score can be 25. The Cronbach α of CRDAT was .88 for the pre-test, .91 for the post-test, and .93 for the retention-test.

6.2.4. *Qualitative analysis of on-line scientific argumentation*

We modified the analytic frameworks of Osborne et al. (2004) in order to assess the quality and quantity of argumentation and scientific conceptions for each statement produced by each individual student. Therefore, each statement produced by each individual was assessed according to level and type of argumentation. Each statement was classified into two different levels of claim, warrant, backing and rebuttal, respectively. Data is considered to be non-argumentative statements. A level 1 claim is an argument consisting of a claim without any data or fact. A level 2 claim is an argument consisting of a claim with data or fact. A level 1 warrant is an argument consisting of a theory or principle without connection to the claim, or one which does not clearly describe the theory. A level 2 warrant is an argument consisting of a claim with a clearly described theory or principle. A level 1 backing is an argument only consisting of a backing without any connection to claim/warrant, or one which does not clearly describe the connection among them. A level 2 backing is an argument consisting of a claim with backing, and or with data or warrant. A level 1 rebuttal is an argument consisting of a weak rebuttal without clear explanation. A level 2 rebuttal is an argument consisting of a claim with a clearly identifiable rebuttal. The cross-coder reliability is .89.

6.2.5. *Analysis of conceptual change*

The degree and nature of conceptual change was judged and classified into four categories according to the students' answers to the driving questions before and after learning. Both before and after learning, the driving question was classified as either correct or incorrect. The success rate of conceptual change yields is derived from the correct answers to the post-driving question. In addition, students' on-line scientific argumentation discourses were analyzed from a conceptual change perspective. Each statement was determined to be correct, partially correct, or incorrect. The nature and extent of students' conceptual change from the analysis of their discourses also was examined. The cross-coder reliability is .91.

7. Results

7.1. *Multivariate analysis of the chemical reaction achievement test (CRAT)*

One-factor MANCOVA was conducted to examine the effects of instructional approaches using post- and retention-CRAT scores as the dependent measures, and students' pre-CRAT scores as the covariate. Table 2 summarizes the results of the one-factor MANCOVA:

Table 2
Multivariate analysis of covariance (MANCOVA) of post- and retention- of chemical reaction achievement test (CRAT) scores.

Source of variance	Wilk's Λ	Multivariate F	Univariate F		Post-hoc
			Post-test	Retention-test	
Covariates					
Pre-test scores	.53*** (.000)	60.20			
Group memberships					
Instructional approaches	.84*** (.000)	13.22	14.86*** (.000)	26.06*** (.000)	Post: $E^a > C^b$ (.000) Retention: $E^a > C^b$ (.000)

***$p < 0.0001$, **$p < 0.001$, *$p < 0.01$.
[a] Experimental group.
[b] Control group.

specifically, instructional approaches (Wilk's $\Lambda = .84$, $p = 0.000$) reach a statistically significant effect on the performance of post- and retention-CRAT.

Therefore, the main effect was performed to independently examine the effect of the instructional approaches on post- and retention-CRAT. This indicated that the effects for instructional approaches on both post-CRAT ($F = 14.86$, $p = 0.000$) and retention-CRAT ($F = 26.06$, $p = 0.000$) were significant. Thus, the students' post- and retention-CRAT were significantly affected by the instructional approach. The post-hoc analysis for the main effect suggests that the on-line argumentation learning program group performed significantly better than the on-line learning program group ($p_{(\text{post})} = .000$, $p_{(\text{retention})} = .000$) on post- and retention-CRAT. In summary, the on-line argumentation learning group outperformed the on-line learning group on both post- and retention-performance of chemical reaction achievement.

7.2. Multivariate analysis of the Chemical Reaction Conceptual Change Test (CRCCT)

One-factor MANCOVA was conducted to examine the effects of instructional approaches using post- and retention-CRCCT scores as the dependent measures and students' pre-CRCCT scores as the covariate. Table 3 summarizes the results of the one-factor MANCOVA: specifically, instructional approaches (Wilk's $\Lambda = .78$, $p = 0.000$) reach a statistically significant effect on the performance of post- and retention-CRCCT.

Therefore, the main effect was performed to independently examine the effect of the instructional approaches on post- and retention-CRCCT. This indicated that the effects for instructional approaches on both post-CRCCT ($F = 32.91$, $p = 0.000$) and retention-CRCCT ($F = 28.51$, $p = 0.000$) were significant. Thus, the students' post- and retention-CRCCT were significantly affected by the instructional approach. The post-hoc analysis for the main effect suggests that the on-line argumentation learning program group performed significantly better than the on-line learning program group ($p_{(\text{post})} = .000$, $p_{(\text{retention})} = .000$) on post- and retention-CRCCT. In summary, the on-line argumentation learning group outperformed the on-line individual learning group on both post- and retention-performance of chemical reaction conceptual change.

7.3. Multivariate analysis of the Chemical Reaction Dependent Argumentation Test (CRDAT)

One-factor MANCOVA was conducted to examine the effects of instructional approaches using post- and retention-CRDAT scores as the dependent measures and students' pre-CRDAT scores as the covariate. Table 4 summarizes the results of the one-factor MANCOVA: specifically, instructional approaches (Wilk's $\Lambda = .90$, $p = 0.001$) reach a statistically significant effect on the performance of post- and retention-CRDAT.

Then main effect was performed to independently examine the effect of the instructional approaches on post- and retention-CRDAT. This indicated that the effects for instructional approaches on both post-CRDAT ($F = 4.47$, $p = 0.036$) and retention-CRDAT ($F = 15.34$, $p = 0.000$) were significant. Thus, the students' post- and retention-CRCCT were significantly affected by the instructional approach. The post-hoc analysis for the main effect suggests that the on-line argumentation learning program group performed significantly better than the on-line learning program group ($p_{(\text{post})} = .036$, $p_{(\text{retention})} = .000$) on post- and retention-CRDAT. In summary, the on-line argumentation learning group outperformed the on-line learning group on both post- and retention-performance of chemical reaction dependent argumentation.

7.4. Multiple regression analysis

This section examines whether or not students' degree of conceptual change would increase as their achievement and argumentation ability increased. Therefore, the stepwise regression method was used to explore whether the post-CRAT test or post-CRDAT would be most important for predicting the post-CRCCT scores, and whether retention-CRAT or retention-CRDAT would be the most important for predicting the retention-CRCCT scores. Results indicated that the best single predictor for post-CRCCT sores was the post-CRAT, followed by post-CRDAT scores. The standardized regression coefficient for post-CRAT and post-CRDAT were .66 and .23. Together post-CRAT and post-CRDAT accounted for 65% of the variance in post-ADRT scores (Table 5). Results indicated that the best single predictor for retention-CRCCT sores was the retention-CRAT, followed by retention-CRDAT. The standardized regression coefficient for retention-CRAT and retention-CRDAT were .64 and .23. Together retention-CRAT and retention-CRDAT accounted for 55% of the variance in retention-CRDRT scores.

7.5. The quantity and quality of on-line scientific argumentation analyses

The experimental group's students' on-line scientific argumentation process was analyzed in two aspects: nature and extent of argumentation and of conceptual change. The quality and quantity of students' argumentation and conceptual change were presented in the following in order to manifest the nature and extent of experimental group's on-line scientific argumentation process.

Table 3
Multivariate analysis of covariance (MANCOVA) of post- and retention- of chemical reaction conceptual change test (CRCCT) scores.

Source of variance	Wilk's Λ	Multivariate *F*	Univariate *F*		Post-hoc
			Post-test	Retention-test	
Covariates					
Pre-test scores	.28*** (.000)	175.13			
Group memberships					
Instructional approaches	.78*** (.000)	19.66	32.91*** (.000)	28.51*** (.000)	Post : E[a] $>$ C[b] (.000) Retention: E[a] $>$ C[b] (.000)

***$p < 0.0001$, **$p < 0.001$, *$p < 0.01$.
[a] Experimental group.
[b] Control group.

7.5.1. Quantity of argumentation

Repeated measures of ANOVA were used to examine increases in mean frequency of arguments from topic 1 to topic 4. Table 6 summarizes the quantity of argumentation generated by each group at each question across the four topics. All argumentation questions were designed to require 10–15 min for students to argue. The mean frequency of arguments generated by each group in each question increased progressively from 3.42 to 9.04 from topic 1 to 4 (Fig. 5). This clearly indicated that the amount of arguments generated by each group increased progressively throughout the program. The mean frequency of arguments generated by each student in each question increased from .67 to 1.78 from topic 1 to topic 4. Closely examined, this reveals a similar pattern to the group argumentation pattern described above (Fig. 6). This clearly demonstrates that students' ability to generate arguments indeed increased from topic 1 to topic 3 and decreased a little at topic 4. Post hoc shows that the number of arguments each group generated at topic 2, 3 and 4 are significantly greater than for topic 1. This indicates that students' ability to generate arguments increased as the program went on.

Each statement generated by students was further categorized into two levels of claim, warrant, backing and rebuttal arguments. With an average of five students in a group, the mean frequency of two levels of claim, warrant, backing and rebuttal arguments generated by each group in each question from topic 1 to 4 ranged from 2.14 to 6.18, .64–2.39, .18–1.11, and .25–.82 (Fig. 7). Repeated measures of ANOVA were used to examine increases in mean frequency of the two levels of claim, warrant, backing and rebuttal arguments from topic 1 to topic 4 (Table 7). There was a trend of increase from topic 1 to topic 4 in all aspects, including two levels of claim, warrant, backing and rebuttal, respectively. The increase also reached a statistically significant difference level when comparing later topics with earlier topics in general. In detail, it shows that the mean frequency of the two levels of claim generated by each group at topic 4 is greater than topic 1 and 2, but only topic 1 reached a statistically significant level. The mean frequency of the two levels of warrant generated by each group at topic 4 is greater than topic 1, 2, and 3; but only topic 1 reached a statistically significant level. The mean frequency of the two levels of backing generated by each group at topic 2 is, statistically, significantly greater than at topics 1, 3, and 4. For rebuttal, the mean frequency of the two levels of rebuttal generated by each student's questions at topic 3 is higher than at topic 1, 2, and 4; but only topic 1 and 4 reached a statistically significant difference level.

7.5.2. Quality of argumentation

Table 7 summarizes the quality of argumentation for the two different levels of claim, warrant, backing and rebuttal arguments generated by each group for each question (10–15 min) across four topics. All of the argumentation questions were designed to require approximately 10–15 min for students to argue. The mean frequency of arguments generated by each group in each question shows a growing pattern overall, regardless of levels of claim, warrant, backing and rebuttal argument. It shows that students produce more level 2 warrant, backing and rebuttal arguments than level 1 warrant, backing and rebuttal arguments as they moved toward topic 4, except level 2 rebuttal arguments went up from topic 3. Students produced more level 1 claim than level 2 claims from beginning to the end.

Repeated measures of ANOVA were used to examine any increases in mean frequency of level 1 claim argument from topic 1 to topic 4. It shows there was an increase of producing level 1 warrant, backing or rebuttal arguments from topic 1 to topic 2 or 3. The increase also reaches a statistically significant difference level when comparing later topics with earlier topics. In detail, it shows that the mean frequency of level 1 claims generated by the questions of each group at topic 3 is greater than at topics 1, 2 and 4; and only topic 1 and 2 reached statistical significance. The mean frequency of level 1 warrants generated by the questions of each group at topic 2 is greater than at topics 1, 3 and 4; and only topic 1 and 4 reached statistical significance. The mean frequency of level 1 backings generated by the questions of each group at topic 2 is, statistically, significant greater than at topics 1, 3, and 4. For rebuttals, the mean frequency of level 1 rebuttals generated by the questions of each group at topic 3 is higher than at topic 1, 2 and 4; and only topic 4 reached a statistically significance level.

Table 4
Multivariate analysis of covariance (MANCOVA) of post- and retention- of chemical reaction dependent argumentation test (CRDAT) scores.

Source of variance	Wilk's Λ	Multivariate *F*	Univariate *F*		Post-hoc
			Post-test	Retention-test	
Covariates					
Pre-test scores	.32***(.000)	148.01			
Group memberships					
Instructional approaches	.90***(.001)	7.70	4.47*(.036)	15.34***(.000)	Post : E[a] $>$ C[b] (.000) Retention: E[a] $>$ C[b] (.000)

***$p < 0.0001$, **$p < 0.001$, *$p < 0.01$.
[a] Experimental group.
[b] Control group.

Table 5
Stepwise Multiple Regression Summary for Relationships among Chemical Reaction Achievement Test (CRAT), Chemical Reaction Conceptual Change Test (CRCCT), and Physical Science Dependent Argumentation Test (PSDAT).

Significant predictor variable	Standardized regression coefficients	SE	*t*
Dependent variable: post-CRCCT			
Post-test of chemical reaction achievement test (CRAT)	.66	.11	7.97***
Post-test chemical reaction dependent argumentation test (CRDAT)	.23	.09	2.78
R	.80		
R Square	.65		
Dependent variable: retention-CRCCT			
Post-test of chemical reaction achievement test (CRAT)	.64	.11	7.38***
Post-test chemical reaction dependent argumentation test (CRDAT)	.23	.09	2.64**
R	.74		
R square	.55		

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Repeated measures of ANOVA were used to examine increases in mean frequency of level 2 claim, warrant, backing, and rebuttal arguments from topic 1 to topic 4. The increase of level 2 claim and warrant arguments reached a statistically significant difference level when comparing later topics with earlier topics. In detail, it shows that the mean frequency of level 2 claims generated by each student's questions at topic 2 is, statistically, significantly greater than at topics 1 and 4. The mean frequency of level 2 warrants generated by each student's questions at topic 4 is, statistically, significantly greater than topics 1, 2 and 3. The mean frequency of level 2 backings generated by each student's questions at topic 2 is greater than at topics 1, 3 and 4. For rebuttals, the mean frequency of level 2 rebuttals generated by each student's questions at topic 3 is, statistically, significantly higher than at topics 1 and 2.

7.5.3. Conceptual change

Table 8 summarizes the mean frequency of students' correct and incorrect answers to the same driving questions before and after argumentation. Adding both correct–correct and incorrect-correct categories together yields the success rate of conceptual change. It indicates that students improved their correct conceptions across four topics in seven questions, and six of those questions reached a statistically significant difference level. In addition, students decreased their incorrect conceptions across four topics in seven questions, and six of them reached a statistically significant difference level.

In between answering the driving questions, students were involved in the argumentation process. The nature of each argument generated in argumentation discourses was judged and classified into three categories of correct, partially correct, and incorrect. Table 9 summarizes the correctness of conceptions of each argument generated by each group from topic 1 to topic 4. It shows that the mean score of correct conceptions gradually increased from topic 1 to topic 4 and reached a statistically significant difference level. The partially correct conceptions gradually increased from topic 1 to topic 3 and significantly decreased at topic 4. Incorrect conceptions increased slightly from topic 1 to topic 3 and decreased significantly at topic 4. This indicated that students gradually changed their conceptions as time went on.

Fig. 8 presents the success rate of conceptual change across four topics in seven questions, showing that the success rate of conceptual change ranges from 70% to 83% for most of the questions after went through argumentation process, although three questions ranged from 41% to 60%.

7.5.4. Argumentation discourses

In order to support the nature and extent of students' conceptual change during the on-line synchronous argumentation, the argumentation discourse of a group of students is provided below.

Example 1: **Topic 3–2** At the condition of 100 °C, if hydrogen and oxygen are mixed and heated with fire, the product is steam (water molecule). The result indicated that the ratio of H_2: O_2:H_2O is 2:1:2. The chemical equation is $2H_2 + 1O_2 \rightarrow 2H_2O$. Please discuss with your group members regarding the relationships of this chemical reaction from the aspects of molecular levels."

Mary: I saw the ratio of H_2: O_2:H_2O are 2:1:2. (Data)
John: I saw 2 hydrogen atoms react with one oxygen atom and produce 2 water molecules. (Data)
Steven: I think the number of atoms would be the same before and after the reaction; only the arrangement of atoms changed. (Data)
Joyce: I saw the kinds of atoms would be the same before and afterthe chemical reaction. (Data)
Mary: I saw that at the condition of 100 °C, mix hydrogen with oxygen and heat with fire, and then produce steam (water molecule) (Data).

Table 6
Repeated measures of ANOVA of arguments generated by each group across four topics.

	M[a]	SD	*F* value of repeated measures ANOVA	Post hoc comparisons
Topic 1	3.42	1.25	40.06*** ($p = 0.000$)	
Topic 2	8.21	3.07		2 > 1***
Topic 3	9.04	1.51		3 > 1***
Topic 4	7.89	3.03		4 > 1***

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $N = 14$ groups.
[a] (mean frequency): the mean frequency of arguments generated by each groups' students at each argumentation question.

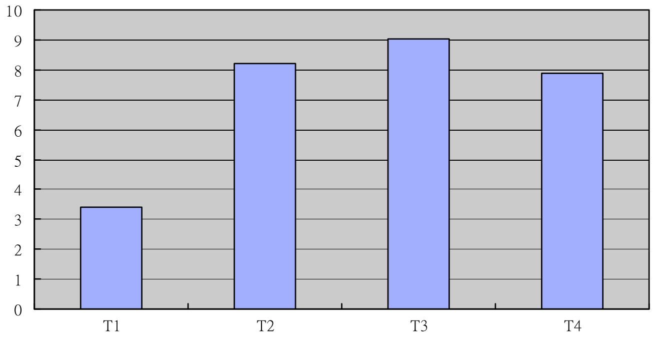


Fig. 5. Distribution of mean frequency of arguments generated by each group of students across four topics.

Steven: I did not see any changes on the numbers and kinds of atoms after the reaction; therefore, I think it would not change or produce new atoms during the chemical reaction,. Only the arrangement of atoms changed. (Claim)
Joyce: I support my ideas, the kinds and numbers of atoms stay the same after the chemical reaction according to the law of conservation of mass. (Warrant)
John: I support Steven's ideas; the atom would not disappear according to the law of conservation of mass. (Warrant)
Mary: I think the arrangement of atoms would be changed after the chemical reaction according to the atomic theory of Dalton. (Warrant)
Mary: The number of atoms would be the same after the chemical reaction according to the atomic theory of Dalton, but the arrangement of atoms would be different. (Warrant)
John: I support Mary's ideas; the atoms would not disappear, but the arrangement of atoms would be changed according to the law of conservation of mass. (Warrant)
Joyce: I found that the number of molecules changed after the chemical reaction, thus I think the number of molecules would change after the reaction. (Claim)
Mary: I disagree with Joyce, because only the arrangement of atoms would be changed according to the principle of chemical reaction. (Rebuttal)
Mary: Chemical reaction only involves the rearrangement of atoms, but the number of atoms would be the same. (Claim)
Joyce: I saw the number of oxygen molecules has changed. (Data)
John: I disagree with Joyce's opinion, because it does not fit the law of conservation of mass. (Rebuttal)
Steven: Chemical reaction follows the law of conservation of mass, therefore, the kinds and numbers of atoms would not change and only the atoms rearrange. (Warrant)

There are 11 arguments and 6 non-argumentative statements initiated by students. Students began with a series of non-argumentative statements to describe what they have observed and known. Gradually, various types of arguments were provided which clearly disclosed how students' argumentation facilitated their conceptual development and evolvement. Most of the arguments were classified as level 2 arguments regardless of claim and warrant. These arguments either consisted of a claim with data or fact (level 2 claim) or consisting of a claim with a clear description of the law of conservation of mass or atomic theory of Dalton (level 2 warant). During the conceptual development and evolvement process, many supporting arguments were displayed to recognize others' opinions. This is fairly positive for learning through argumentation. Near the end of the argumentation, Joyce raised a claim argument which consisted of an alternative conception stating that the number of molecules changed after the chemical reaction. Mary and John disagreed with Joyce's conceptions and provided rebuttal arguments trying to persuade her. In order to make it clearly to Joyce, Steven further generated a warrant argument specifically emphasizing that the kinds and number of atoms would not change, but only be rearranged, with the support of law of conservation. This demonstrated that students are able to accurately determine what alternative conceptions their peer had and further provide rebuttal or warrant to employ, through the use of either evidence or theory, in order to persuade their group member. More

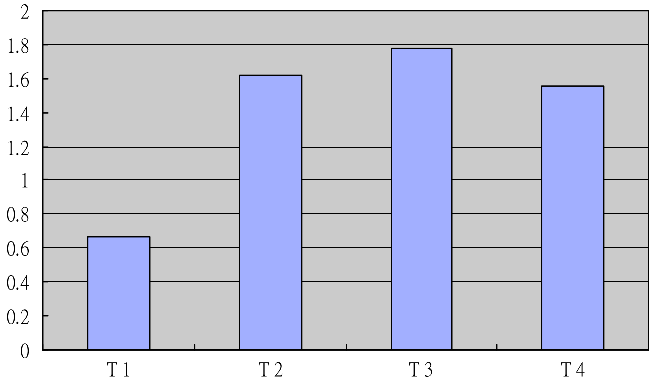


Fig. 6. Distribution of mean frequency of arguments generated by each student across four topics.

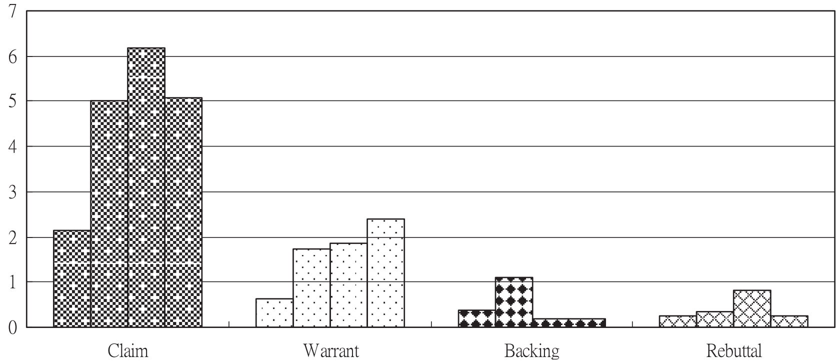


Fig. 7. Distribution of mean frequency of claim, warrant, backing and rebuttal arguments generated by each group across four topics.

importantly, it was found that all five students correctly answered that the number of hydrogen and oxygen atoms before the reaction is equal to the number of hydrogen and oxygen atoms after the reaction according their post-driving question responses.

8. Discussions

The results of this study are quite encouraging as they demonstrate that the on-line argumentation learning group is far more effective than the on-line learning group for promoting students' conceptions of chemical reaction, conceptual change related to chemical reaction, and chemical reaction dependent argumentation. The Author argues that most computer-assisted learning studies can not effectively change students' alternative conceptions or promote students' science learning outcome because their instructional materials are not developed based on solid theories or models of conceptual change or science learning (She & Lee, 2008; Liao & She, 2009). It supports the results which found that the on-line argumentation learning group performed far better than the on-line learning group because the on-line learning group did not include argumentation theory and conceptual change theories into the design.

Our results further demonstrate that students' argumentation in a science context can be nurtured through the use of the on-line synchronous scientific argumentation program. It supports some of the previous studies that engaging in argumentation improves conceptual understanding in science (Duschl et al., 1999; Zohar & Nemet, 2002), and promotes conceptual change in university students (Nussbaum & Sinatra, 2003) and secondary school students (Ravenscroft, 2000, 2007; Ravenscroft & Pilkington, 2000). Our study adds new evidence that providing argumentation templates enhances students' argumentation skills in science, in addition to the previous studies involving history or social studies (Li & Lim, 2008; Saye & Brush, 2002).

Our regression results support the idea that argumentation not only allows students to consolidate existing scientific knowledge, but also to construct new knowledge for themselves based on the ideas of others (Brown & Campione, 1998; Driver et al., 2000; Nussbaum & Sinatra, 2003). Some studies suggest that conceptual change involves deep restructuring, not only in the concepts, but also in the way of reasoning (Furió, Calatayud, Bárcena, & Padilla, 2000; Gil-Pérez & Carrascosa, 1994). Other studies have reported that conceptual change would be

Table 7
Repeated measures ANOVA of two levels of claim, warrant, backing, and rebuttal arguments generated by each group across four topics.

Topics	Level 1		Level 2		Total		F^{b}	Post hoc comparisons
	M^{a}	SD	M	SD	M	SD		
Claim								
Topic 1	1.91	.85	.22	.19	2.14	.85	16.12^{***}	
Topic 2	3.21	1.70	1.71	1.52	5.00	2.08		$2 > 1^{**}$
Topic 3	5.07	2.49	1.07	.81	6.18	1.72		$3 > 1^{***}$
Topic 4	4.18	2.52	.86	.77	5.07	2.20		$4 > 1^{**}$
Warrant								
Topic 1	.52	.29	.11	.15	.64	.40	6.61^{**}	
Topic 2	1.21	.85	.54	.57	1.75	1.09		$2 > 1^{**}$
Topic 3	1.14	.84	.68	.91	1.86	1.31		$3 > 1^{**}$
Topic 4	.50	.59	1.75	1.74	2.39	1.86		$4 > 1^{**}$
Backing								
Topic 1	.34	.25	.03	.05	.39	.27	4.48^{*}	
Topic 2	.93	.65	.18	.37	1.11	.88		$2 > 1^{**}$, $2 > 3^{**}$,$2 > 4^{**}$
Topic 3	.11	.21	.07	.18	.18	.25		
Topic 4	.07	.18	.11	.40	.18	.54		
Rebuttal								
Topic 1	.19	.21	.06	.07	.26	.30	2.27	
Topic 2	.36	.46	.04	.13	.36	.46		
Topic 3	.39	.59	.43	.58	.82	1.08		$3 > 4^{*}$, $3 > 1^{*}$
Topic 4	.11	.21	.18	.42	.25	.51		

$^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.000$, $N = 14$ groups.

[a] Mean frequency: the mean frequency of five arguments generated by each individual students at each argumentation question.

[b] F value of repeated measures ANOVA.

Table 8
The mean frequency of correctness of pre- and post-driving question across four topics with seven questions.

	Pre-driving question		Post-driving question		Mean difference	*T*
	M	SD	*M*	SD		
Topic 1						
Question 1						
Correct	.23	.42	.59	.50	.36	5.22***
Incorrect	.77	.42	.41	.50	−.36	−5.22***
Question 2						
Correct	.28	.45	.72	.45	.44	7.37***
Incorrect	.72	.45	.28	.45	−.44	−7.37***
Topic 2						
Question 1						
Correct	.80	.40	.83	.38	.03	.58
Incorrect	.20	.40	.17	.38	−.03	−.58
Question 2						
Correct	.65	.48	.80	.40	.15	2.49*
Incorrect	.35	.48	.20	.40	−.15	−2.49*
Topic 3						
Question 1						
Correct	.27	.45	.41	.50	.14	2.19*
Incorrect	.73	.45	.59	.50	−.14	−2.19*
Topic 4						
Question 1						
Correct	.27	.45	.46	.50	.19	3.02**
Incorrect	.73	.45	.54	.50	−.19	−3.02**
Question 2						
Correct	.73	.45	.86	.35	.13	2.12*
Incorrect	.27	.45	.14	.35	−.13	−2.12*

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.000$; $N = 74$.

greatly enhanced by students' grasp of correct conception and increased scientific reasoning ability (She & Liao, 2010). Krummheuer (1995) provides a helpful definition of argument as “the intentional explication of the reasoning of a solution during its development or after it”. In other words, argumentation can be interpreted as being a number of people who are involved in the reasoning process together. Our research adds a new interpretation which is that conceptual change involves deep restructuring, not only in the concepts but also in the way of argumentation.

Osborne et al. (2004) found that developing argumentation in a scientific context is significantly more difficult than engaging argumentation in a socio-scientific context. Our study was set up to investigate students' argumentation in a scientific context with supporting on-line argumentation templates. Our results indicate that the quantity and quality of arguments generated by students indeed improved gradually from topic 1 to topic 4, and reached a statistically significant difference level. Our study partially supported Osborne's studies that developing students' argumentation ability in a scientific context is rather difficult and indeed takes time in the beginning of their learning; however, with supporting argumentation templates, the on-line learning environment indeed fosters their ability to generate more and higher level arguments from topic 2. In short, our finding provides evidence that the obstacle to developing argumentation ability in a scientific context can be overcome through providing students with a framework of templates for argumentation.

Table 9
Analysis of the correctness of conceptions for each argument generated at each argumentation question by each group across four topics.

Topics	*M*	SD	*F* value of repeated measures ANOVA	Post hoc comparisons
Correct conceptions				
Topic 1	1.10	.58	29.79*** ($p = 0.000$)	$2 > 1^{***}$
Topic 2	4.82	2.26		$3 > 1^{***}$
Topic 3	3.86	1.49		$4 > 1^{***}$, $4 > 3^{**}$
Topic 4	6.14	2.84		
Partial correct conceptions				
Topic 1	2.17	.89	17.26*** ($p = 0.000$)	
Topic 2	4.07	2.12		$2 > 1^{**}$, $2 > 4^{**}$
Topic 3	5.25	1.64		$3 > 1^{***}$, $3 > 2^{*}$, $3 > 4^{***}$
Topic 4	2.25	1.70		
Incorrect conceptions				
Topic 1	1.05	.51	14.52*** ($p = 0.000$)	
Topic 2	1.96	1.46		$2 > 1^{*}$
Topic 3	2.96	1.49		$3 > 1^{**}$, $3 > 4^{***}$
Topic 4	1.25	1.46		

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.000$, $N = 14$ groups.

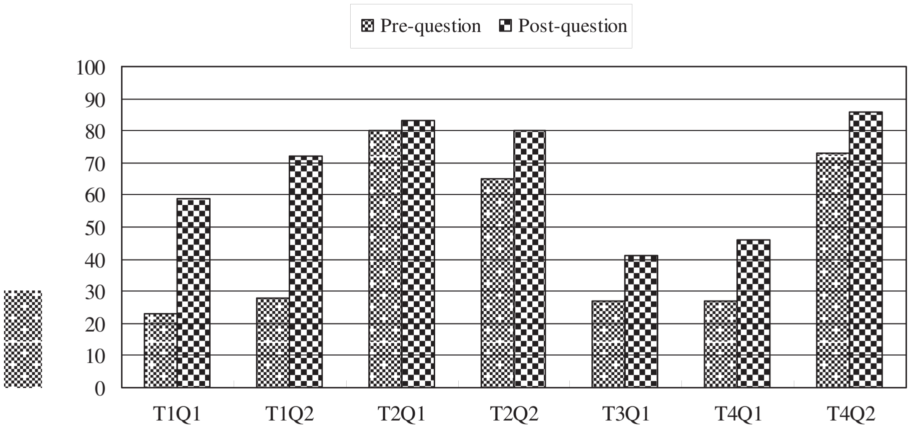


Fig. 8. The successful rate of conceptual change across four topics with seven questions.

It also indicates that the difficulties of learning to generate claim, warrant, backing, and rebuttal arguments are varied, which means that it might take longer to nurture students' ability to generate rebuttal than claim, warrant and backing arguments. A possible explanation is that the ability to generate level 2 rebuttal needs a longer time to cultivate than level 2 claim, warrant, and backing. The findings also demonstrate a fairly positive and significant improvement in students' argumentation as the program went on. One of the interesting findings is that the ability to generate claim, backing, and rebuttal arguments drops at some point before reaching a plateau. Further research might be useful to explore how long it would take for students to reach their optimum plateau in generating claim, warrant, backing, and rebuttal arguments.

Our findings also indicated that students successfully changed their alternative conceptions throughout this program. One possible explanation for such success in conceptual change is that the design of the on-line synchronous scientific argumentation learning program is based on the integration of argumentation theory, which considers different conceptual change theories as the basis for building the social constructivism aspects of conceptual change (She, 2002, 2004b; Clement, 1993; Driver, 1995; Driver & Oldham, 1986; Hewson & Hewson, 1988; Savinainen et al., 2005; Strike & Posner, 1985; Venville & Treagust, 1998; Vosniadou & Brewer, 1994). The program itself was purposely designed with a series of activities consisting of a pre-driving question, group argumentation, a post-driving question, and finally a series of well-designed multimedia learning materials. The intention was to bring students into the process of actively generating arguments according to what they know, see and believe, and then to lead them to further face dissonance, interchange opinions and change their alternative conceptions. The results confirm our initial belief that it is possible to promote students' conceptual change through strongly theory-based argumentation.

9. Conclusions

This study reports the development of an on-line synchronous scientific argumentation program based on the argumentation and conceptual change theories in order to promote students' conceptual change and argumentation involving chemical reaction. Our study makes a major step from previous on-line argumentation learning by incorporating argumentation and conceptual change theories and pedagogy into the on-line argumentation learning program. In addition, this program also includes the ideas of synchronous communication which provides students with the opportunity to argue with their group in real time and deliver a higher degree of elaboration and construction of arguments. Moreover, the argumentation template in our interface was developed in order to provide students with the support of scaffolding and building effective and good arguments. Results are very positive as described in the following.

The one-factor MANCOVA indicated that instructional approach has significant effect on students' performance in post- and retention-test of CRAT, CRCCT and CRDAT. Post hoc analysis further suggests that the on-line argumentation learning group significantly outperformed the on-line learning group on post-test and retention-test of CRAT, CRCCT, and CRDAT. This demonstrates that the on-line argumentation learning group is far more efficient than the on-line learning group for promoting students' conceptions of chemical reaction, conceptual change related to chemical reaction, and chemical reaction dependent argumentation.

Regression indicated that the best predictor for post-CRCCT was post-CRAT, followed by post-CRDAT scores. Similarly, the best predictor for retention-CRCCT was retention-CRAT, followed by retention-CRDAT scores. This implies that student' conceptual change increases as their chemical reaction conception and chemical reaction dependent argumentation ability increases.

Results indicate that the amount of arguments categorized as rebuttal increased progressively from topic 1 to topic 3, and reached a statistically significant difference level when comparing later topics with earlier topics. It also indicated that claim arguments generated by students all increased after topic 2, reached a plateau, and then dropped a little. Warrant arguments gradually increased from topic 1 to topic 4. Backing arguments were generated by students increased after topic 2, and then dropped significantly after topic 2. Students start with limited ability to generate higher level arguments, such as Level 2 warrant, backing and rebuttal, at the beginning of receiving the on-line argumentation learning program; however, their ability to generate higher level arguments increased substantially toward topic 4.

Finally, our findings provide evidence that students' conceptual change can be fostered through a series of carefully designed on-line argumentation learning programs. Results show that mean frequency of the correct conception of each argument generated by students gradually increased as they progressed from topic 1 to topic 4 and reached a statistically significant difference level. Their partially correct conceptions gradually increased from topic 1 to topic 3 and significantly decreased at topic 3. Incorrect conceptions increased slightly from topic 1 to topic 3 and decreased significantly at topic 4. This indicated that students gradually changed their conceptions as time went on.

The degree of conceptual change shows that the success rate for conceptual change ranged from 70% to 83% for most of the post-driving questions, although three questions ranged from 41% to 60%. This supports the result that the experimental group students' scores on CRAT and CRCCT made significant progress from pre- to post- and then from post- to retention-test. It also demonstrates that well-designed argumentation and conceptual change program indeed promotes students' conceptual change.

Acknowledgements

This research is based in work supported and funded by the National Science Council (NSC 95-2522-S-009-001-MY3), Taiwan, R.O.C.

References

Andriessen, J., Erkens, G., Van de Laak, C., Peters, N., & Coirier, P. (2003). Argumentation as negotiation in electronic collaborative writing. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 79–115). The Netherlands: Kluwer Academic Publishers.

Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching, 45*(1), 101–131.

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). Netherlands: Kluwer Academic Publishers.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: designing for learning from the web with KIE. *International Journal of Science Education, 22*(8), 797–817.

Brown, A. L., & Campione, J. C. (1998). Designing a community of young learners: theoretical and practical lessons. In N. M. Lambert, & B. L. McCombs (Eds.), *How students learn: Reforming schools through learner-centered education* (pp. 153–186). Washington, DC: American Psychological Association.

Chinn, C. A., & Anderson, R. C. (1998). The structure of discussions intended to promote reasoning. *Teachers College Record, 100*(2), 315–368.

Cho, K. L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology: Research & Development, 50*(3), 5–22.

Clark, D. B., & Sampson, V. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching, 45*(3), 293–321.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30*(1), 1241–1257.

Cobern, W. W. (1993). Contextual constructivism: the impact of culture on the learning and teaching of science. In K. Tobin (Ed.), *The practice of constructivism in science education* (pp. 51–69). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cross, D., Taasoobshirazi, G., Hendricks, S., & Hickey, D. T. (2008). Argumentation: a strategy for improving achievement and revealing scientific identities. *International Journal of Science Education, 30*, 837–861.

De Vries, E., Lund, K., & Baker, M. (2002). Computer-mediated epistemic dialogue: explanation and argumentation as vehicles for understanding scientific notions. *The Journal of the Learning Sciences, 11*(1), 63–103.

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in cognitive construction of knowledge. *Educational Psychologist, 33*(2/3), 109–128.

Driver, R. (1995). Constructivist approaches to science learning. In L. P. Steffe, & J. Gale (Eds.), *Constructivism in education* (pp. 385–400). Hillsdale, NJ: Erlbaum.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Philadelphia, PA: Open University Press.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classroom. *Science Education, 84*(3), 287–312.

Driver, R., & Oldham, V. (1986). A constructivist approach to curriculum development in science. *Studies in Science Education, 13*(1), 105–122.

Duschl, R., Ellenbogen, K., & Erduran, S. (1999, April) Understanding dialogic argumentation among middle school science students. Paper presented at American Educational Research Association Annual Conference, Montreal, Canada

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38*(1), 39–72.

van Eemeren, F. H., & Grootendorst, R. (2004). *A systematic theory of argumentation: the pragma-dialectical approach*. New York: Cambridge University Press.

Furió, C., Calatayud, M. L., Bárcena, S., & Padilla, D. H. (2000). Functional fixedness and functional reduction as common sense reasoning in chemical equilibrium and in geometry and polarity of molecules. *Science Education, 84*(5), 545–565.

Gil-Pérez, D., & Carrascosa, J. (1994). Bringing pupils learning closer to a scientific construction of knowledge: a permanent feature in innovations in science teaching. *Science Education, 78*(3), 301–315.

Hewson, P. W., & Hewson, M. G. (1988). An appropriate conception of teaching science: a view from studies of science learning. *Science Education, 72*(5), 597–614.

Janssen, J., Erkens, G., & Kanselaar, G. (2006). Visualization participation to facilitate argumentation. Paper presented at the 7th International Conference of the Learning Sciences, Bloomington, IN.

Jimenez-Aleisandre, M. P., & Erduran, S. (2008). Argumentation in Science Education: an Overview. In M. P. Jimenez-Aleisandre, & S. Erduran (Eds.), *Argumentation in Science Education* (pp. 3–25). Springer publisher.

Kirschner, P. A., Buckingham Shum, S. J., & Carr, C. S. (2003). *Visualizing argumentation: Software tools for collaborative and educational sense-making*. London: Springer Verlag.

Kolsto, S. D., & Ratcliffe, M. (2008). Social aspects of argumentation. In S. Erduran, & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 114–133). Berlin, Germany: Springer.

Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb, & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229–269). Hillsdale, NJ: Earlbaum.

Kuhn, D. (1993). Science as argument: implications for teaching and learning scientific thinking. *Science Education, 77*(3), 319–337.

Kuhn, D. (2005). *Education for thinking*. Cambridge, Massachusetts: Harvard University Press.

Leitao, S. (2000). The potential of argument in knowledge building. *Human Development, 43*, 332–360.

Li, D. D., & Lim, C. P. (2008). Scaffolding online historical inquiry tasks: a case study of two secondary school classrooms. *Computers & Education, 50*(4), 1394–1410.

Liao, Y. W., & She, H. C. (2009). Enhancing Eight Grade Students' Scientific Conceptual Change and Scientific Reasoning through a Web-based Learning Program. *Educational Technology & Society, 12*(4), 228–240.

Mercer, N., Littleton, K., & Wegerif, R. (2004). Methods for studying the processes of interaction and collaborative activity in computer-based educational activities. *Technology, Pedagogy and Education, 13*(2), 193–209.

Millar, R., & Osborne, J. F. (1998). *Beyond 2000: Science education for the future.* London: King's College London.

National Research Council (1996). *The National Science Education Standards*. Washington, D.C.: National Academy Press.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation in pedagogy of school science. *International Journal of Science Education, 21*(5), 553–576.

Nussbaum, M. E., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology, 28*(3), 384–395.

Organisation for Economic Cooperation and Development (2007). *PISA 2006: Science Competencies for Tomorrow's World. Vol I: Analysis (pp. 101).* Paris: OECD Publishing.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994–1020.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: the role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167–199.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Science Education, 66*(2), 211–227.

Ravenscroft, A. (2000). Designing argumentation for conceptual development. *Computers & Education, 34*(3–4), 241–255.

Ravenscroft, A. (2007). Promoting thinking and conceptual change with digital dialogue games. *Journal of Computer Assisted Learning, 23*(6), 453–465.

Ravenscroft, A., McAlister, S., & Baur, E. (2006). Development, piloting and evaluation of interloc: An open source tool supporting dialogue games in education. Final Project Report to UK JISC (Joint Information Systems Committee), Bristol, UK.

Ravenscroft, A., & Pilkington, R. M. (2000). Investigation by design: developing dialogue models to support reasoning and conceptual change. *International Journal of Artificial Intelligence in Education, Special Issue on Analyzing Educational Dialogue Interaction: From Analysis to Models that Support Learning, 11*(1), 273–298.

Ravenscroft, A., Wegerif, R. B., & Hartley, J. R. (2007). Reclaiming thinking: dialectic, dialogic and learning in the digital age. *Special Issue of British Journal of Educational Psychology (BJEP): Psychological Insights into the Use of New Technologies in Education, 11*(5), 39–57.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation driven inquiry: integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*(3), 345–372.

Savinainen, A., Scott, P., & Viiri, J. (2005). Using a bridging representation and social interactions to foster conceptual change: designing and evaluating an instructional sequence for Newton's third law. *Science Education, 89*(2), 175–195.

Saye, J. W., & Brush, T. A. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research & Development, 50*(3), 77–96.

Schellens, T., & Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers & Education, 46*(4), 349–370.

She, H. C. (2002). Concepts of higher hierarchical level required more dual situational learning events for conceptual change: a study of students' conceptual changes on air pressure and buoyancy. *International Journal of Science Education, 24*(9), 981–996.

She, H. C. (2004a). Fostering "Radical" conceptual change through Dual Situated Learning Model. *Journal of Research in Science Teaching, 41*(2).

She, H. C. (2004b). Facilitating changes in ninth grade students' understanding of dissolution and diffusion through DSLM instruction. *Research in Science Education, 34*(4), 503–526.

She, H. C., & Lee, C. Q. (2008). SCCR Digital Learning System for Scientific Conceptual change and Scientific Reasoning. *Computers & Education, 51*(2), 724–742.

She, H. C., & Liao, Y. W. (2010). Bridging Scientific Reasoning and Conceptual Change through Adaptive Web-based Learning. *Journal of Research in Science Teaching, 47*(1), 91–119.

Strike, K. A., & Posner, G. J. (1985). A conceptual change view of learning and understanding. In L. H. T. West, & A. L. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 211–231). Orlando, FL: Academic Press.

Toulmin, S. (1958). *The uses of argument*. Cambridge, England: Cambridge University Press.

Venville, G., & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Teaching, 35*(9), 1031–1055.

Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science, 18*(1), 123–183.

Wegerif, R., & Mercer, N. (1996). Computer and reasoning through talk in the classroom. *Language and Education, 10*(1), 47–65.

West, L. H. T., & Pines, A. L. (1983). How "Rational" is rationality? *Science Education, 67*(1), 37–39.

Wray, D., & Lewis, M. (1997). *Extending literacy: Children reading and writing non-fiction*. London: Routledge.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching, 39*(1), 35–62.