

Facilitating Students' Conceptual Change and Scientific Reasoning Involving the Unit of Combustion

Chin-Quen Lee · Hsiao-Ching She

Published online: 23 June 2009

© Springer Science + Business Media B.V. 2009

Abstract This article reports research from a 3 year digital learning project to unite conceptual change and scientific reasoning in the learning unit of combustion. One group of students had completed the course combining conceptual change and scientific reasoning. The other group of students received conventional instruction. In addition to the quantitative data, six students from each group were interviewed to evaluate their conceptual change, correct concepts and scientific reasoning. Results indicate that the experimental group's students significantly outperformed the conventional group on the Combustion Achievement Test (CAT), Scientific Reasoning Test (SRT) and Combustion Dependent Reasoning Test (CDRT). Moreover, the experimental group's students use higher levels of scientific reasoning more frequently and changed their alternative concepts more successfully than did the conventional group. Furthermore, once the experimental group's students successfully changed their conceptions, their concepts tended to be more stable than the conventional group's students, even after the 6th week of learning. These results demonstrate that combining conceptual change and scientific reasoning indeed improves students' conceptual change and scientific reasoning ability more effectively than conventional instruction.

Keywords Scientific reasoning · Conceptual change · Combustion · Adaptive digital learning

Theoretical Framework

Reasoning is the process of drawing conclusions from principles and from evidence (Wason and Johnson-Laird 1972), moving on from what is already known to infer new conclusions or to evaluate a proposed conclusion. Anderson (1990) refers to reasoning as the mental processes involved in generating and evaluating logical arguments. Ennis (1987) proposed that reasoning skills include clarification, basis, inference and evaluation. Clarification requires identifying and formulating questions, analyzing elements, and defining terms.

C.-Q. Lee · H.-C. She (✉)

Institute of Education, National Chiao Tung University, 1001 Ta-Hsueh Rd.,
Hsin Chu City, Taiwan, Republic of China
e-mail: hcshe@mail.nctu.edu.tw

Basis refers to conclusions about a problem being supported by information from personal observations, statements by others, and previous inferences. Inference skills differ, depending on whether the scientific reasoning involved proceeds inductively or deductively. Inductive reasoning proceeds from specific facts to general conclusions, and it is the main reasoning process used by scientists to arrive at generalizations or scientific laws (Wadsworth 1996, p. 113). Deductive reasoning is the process of reasoning from one or more general statements on what is known to then reach a logically certain conclusion (Johnson-Laird 2000; Rips 1999). Evaluation involves using criteria to judge the adequacy of a problem solution. Individuals use reasoning to analyze complex information efficiently, to make associations with existing cognitive structure, and to form long term memory (LTM) networks.

Many researchers have suggested that conceptual change involves deep restructuring, not only in the concepts but also in the way of reasoning itself (Gil-Perez and Carrascosa-Alis 1994; Furio et al. 2000). In general, these conceptual changes are associated with other changes, such as new ways to view phenomena; or changes in interests, attitudes, or assumed values by the community. These changes correspond with new ways of reasoning and new approaches for solving scientific problems (Duschl and Gitomer 1991; Hashweh 1986). Conceptual change must not be considered merely as a change in content (Hewson and Thorley 1989). Thus, it is necessary to associate conceptual change with scientific reasoning, because science teaching is usually centered on declarative knowledge (knowing “what”) and it pays less attention to the procedural and explicative knowledge (knowing “how” and “why”) (Kuhn 1993).

Lawson (2003) has discussed the relationship existing between students’ misconceptions and their reasoning ability. In order to modify misconceptions, students need to be aware of their own misconceptions and scientific conceptions, together with the evidence and reasoning that questions the validity of the misconceptions. In other words, they must be able logically to see how the evidence supports the scientific conceptions and contradicts the prior misconceptions. Lawson and Weser (1990) measured pre- to post-instruction change among college students and found that less skilled reasoners were initially more likely to hold a variety of nonscientific beliefs about life and were less likely to change some, though not all, of those beliefs. The way they identified students as skilled or less skilled reasoners was based upon Lawson’s Scientific Reasoning Test (SRT), which measures students’ conservation, proportional thinking, identification and control of variables, probabilistic thinking, correlative thinking and hypothetic-deductive thinking ability. Lawson and Worsnop (1992) again found that the more skilled reasoners in a sample of high school students were less likely to hold pre-instruction misconceptions regarding evolution and special creation. Those studies suggested that more skilled reasoners would be less likely to hold alternative conceptions.

Several researchers further proposed possible explanations why reasoning could impact conceptual change. Oliva (2003) examined the relationship between the degree of structural coherence in students’ conceptions (structuralization) and the probability of conceptual change. The results showed that students with higher levels of formal reasoning (logical thinking) change their alternative conceptions more easily when they display a higher level of initial structuralization. It is possible that students with higher levels of formal reasoning may have the necessary cognitive and metacognitive abilities to effect a sudden and global change, in which case structuralization may facilitate this process. Oliva’s levels of formal reasoning basically were measured through the use of Test of Logical Thinking (TOLT) which measures students’ proportional reasoning, probabilistic reasoning, controlling variable, correlational reasoning, and combinatorial reasoning ability, similar to Piaget’s

level of intellectual development clinical interview test. Trumper and Gorsky (1993) also presented similar ideas that those students with higher levels of formal reasoning present preconceptions that are more structured, though they also tend to change their alternative conceptions with greater ease. On the other hand, Park and Han (2002) suggested that deductive reasoning is a potential factor to help students recognize cognitive conflict and resolve it. Reasoning can help students to recognize their own changed ideas and why they would change their alternative conceptions. These studies all support the central important idea that reasoning can stimulate students' ability to recognize their own conceptions and cognitive conflicts and can further restructure their alternative conceptions, regardless of formal reasoning or deductive reasoning aspects.

The studies reported above indicate that scientific reasoning plays an important role in the process of conceptual change. Nevertheless, none of the empirical studies have found evidence to determine whether forcing students' use of scientific reasoning would improve their success in conceptual change. It is also questionable whether a well-developed conceptual change model with a scientific reasoning emphasis would foster students' conceptual change and scientific reasoning ability. It seems there are still unsolved and challenging questions. Will a conceptual change model with stronger emphasis on scientific reasoning foster students' conceptual change as well as their scientific reasoning ability? Can students' scientific reasoning ability improve along with conceptual change over time?

Most previous studies considered only one perspective of conceptual changes in their model or theory: epistemological, ontological, or social/effective. However, several more recent researchers have studied conceptual change from multidimensional perspectives (Mbajjorgu et al. 2007; Savinainen et al. 2005; She 2004b; Venville and Treagust 1998). The Dual Situated Learning Model (DSLML) (She 2002, 2003, 2004a, b; Tang et al. 2005) clearly shares similar ideas of multidimensional perspectives in terms of epistemological, ontological, and motivational perspectives for its theoretical construct.

The DSLML (She 2004a, b) emphasizes the nature of science concepts, together with students' ontological and epistemological beliefs of science concepts as its major theoretical constructs for conceptual change. In the DSLML (She 2004a, b), situated learning indicates that learning is a process of conceptual change; and moreover, that the learning process should be situated in the nature of science concepts and students' beliefs of these science concepts in order to determine what essential mental sets are needed for constructing a more scientific view of these concepts. The term "dual" indicates that the learning events serve three dual functions: the first dual is to consider both the nature of science concepts and students' beliefs of these science concepts; the second dual is creating dissonance with students pre-existing knowledge and providing a new mental set for them to achieve a more scientific view of the concept; and the third dual is arousing students' motivation, while also challenging their ontological and epistemological beliefs of science concepts.

Based on the DSLML, all of the dual situated learning events should be designed according to the nature of the science concepts and students' beliefs of scientific concepts. In other words, the information on how many and which particular mental sets the students lack for restructuring the science concepts would determine how many and what specific dual situated learning events would be designed to foster students' conceptual change. She (2002) proposed that concepts having a higher hierarchical level subsume more essential underlying concepts, thus making it more difficult for higher-level conceptual changes to occur. Thus conceptual change can not happen only due to a discrepant event; however, the changes need a series of events to increase the success of students' conceptual change (She 2002). The number of dual situated learning events required would depend on the number of mental sets that students lack for constructing a more scientific view of the concepts.

More importantly, each dual situated learning event should be connected with the others, and it needs to build upon a prior dual situated learning event.

Moreover, the author claims that conceptual change can not occur simply by creating cognitive conflict, but that conceptual change needs to provide students with new mental sets in addition to creating dissonance. Creating dissonance with students' pre-existing knowledge can arouse students' curiosity and interest, as well as challenge their epistemological and ontological beliefs of the science concepts. Motivation is embodied in the design and the learning process of a dual situated learning event, which requires students to actively engage in the event prediction, visualize what actually happens and explain why it is different from their prediction, thereby stimulating their curiosity and interest. Providing a new mental set should be the platform where knowledge reconstruction can occur. Students must comprehend and believe the new mental set in order for conceptual change to happen. This can be fostered by any type of instructional activity, such as analogy, modeling, discrepant events, and inquiry activities, as long as it provides students with opportunities to visualize what actually happens in order to reconstruct new mental sets. DSLM emphasizes that students should be actively engaged in the conceptual change learning process. All of these criteria above should serve as the basis for developing each dual-situated learning event in order for it to be successful.

To implement this theory, the DSLM has been transformed into a six stage learning/instructional model. Stage 1: Examining attributes of the science concept, which provides information about the essential mental sets needed to construct a scientific view of the concept. Stage 2: Probing students' alternative scientific conceptions which require probing the students' beliefs concerning the science concept. Stage 3: Determining which mental sets students lack for the construction of a more scientific view of the concepts. Stage 4: Designing dual situated learning events, according to the stage 3 results, which showed which mental sets students lack. Stage 5: Instructing with dual situated learning events, to give students the opportunity to make predictions, provide explanations, confront dissonance, and construct a more scientific view of the concepts. Stage 6: Instruction using a challenging situated learning event, so the students can apply the mental sets they have acquired to a new situation in order to ensure that successful conceptual change has occurred.

The DSLM designs each dual-situated learning event beginning with a driving question, which is targeted to the alternative conceptions commonly found in students. The students are required to provide both an answer and their explanation. Followed by the driving questions, various activities such as graphic and textual illustrations, simulated animations, simulated experiments, and analogies are provided. The design of these activities must create dissonance, challenge students' ontological and epistemological beliefs, and help the students reconstruct the mental sets they lack. Then the same driving question is asked again, and students are required to provide an answer with explanations for what happens in the events.

The Dual Situated Learning Model (DSLM) has evidenced effective conceptual change in middle school students for the topics of air pressure and buoyancy, thermal expansion, heat transfer, dissolution and diffusion, as well as meiosis and mitosis (She 2002, 2003, 2004a, b; Tang et al. 2005). The author believes there is great potential to enhance students' conceptual change as well as scientific reasoning ability through uniting scientific reasoning and conceptual change. Although the conceptual change model is based on the ideas of reasoning and the design of each dual situated learning event has requested students to provide explanations, it has not yet been considered whether the DSLM has any impact on students' scientific reasoning ability. Park and Han (2002) assumed that deductive reasoning plays an important role in conceptual change. But for unknown reasons,

deductive reasoning is not always activated and used spontaneously in students' minds in the process of conceptual change. Thus the DSLM may not be efficient for fostering students' use of scientific reasoning skill. It is clear that the critical point is how to enable the instructional design to activate scientific reasoning.

Therefore, we have modified the dual-situated learning event to engage students in scientific reasoning. The students are asked to provide an answer and explain their reasoning for the driving question before the various activities. The same driving question is asked again after the event to examine the students' conceptual change as well as their reasoning changes. In order to activate students' use of scientific reasoning, we have specifically restructured the process to let students view their answers and reasoning before and after events, as well as the correct answers. We also ask students to reason why they changed or stayed with their original concepts after the events. This is intended to force students to be actively involved with the process of scientific reasoning as well as conceptual change. It is also intended that enhancing scientific reasoning to maximize DSLM can facilitate students' conceptual change and their scientific reasoning ability.

Many students have problems understanding the requirements for combustion and the process of combustion. Many studies have reported that students who describe combustible materials as being reduced to ashes consider that much of the combustible material simply disappears (Andersson 1990; Boujaoude 1991), and ashes get lighter when they are burnt (Driver 1985). Students are not clear that oxygen (gas) is needed for burning and its function for the burning (Andersson 1990; Driver 1985; Watson et al. 1995). They also consider that flames contain only the combustible substance or oxygen and air, or possibly both, but with no interaction. Furthermore, the flame or fire is considered a source of heat to make the modification occur (Watson et al. 1995). Prieto et al. (1992) noticed that students concentrated on the perceptually obvious features of changes such as wax melting and did not see the wax as being used up or interacting with the air. They also reported that students saw the wood changing to ash, and did not mention oxygen or air being involved in the burning process, and ignored any gases produced. Students also think that a change in form of a substance can cause a change in mass, and in particular, that gases have zero or negative weight. Boujaoude (1991) reported that students' understanding about burning are: wax, alcohol, and oxygen are not actively involved in burning; substances undergo no chemical change during burning; terms such as evaporation and burning can be used interchangeably when describing burning alcohol; phrases such as physical change and chemical change can be used interchangeably when describing burning things; burning wax was actually melting and its mass would not change; the loss of weight of the candle would be minimal due to the consumption of the wick or evaporation of wax. Schollum and Happs (1982) reported that students alternative conceptions of combustion and rusting are: air is not involved with burning; burning does not produce new substances; the evaporation of a substance is the process of burning; no molecules exist in a burning flame; rusting is a protective layer so a metal nail will not be destroyed; rustiness exists in the air and it would attach to a metal nail through some process; water and some other materials would produce rusting; rusting occurs because the water destroys a metal nail. These students' alternative conceptions from the previous studies described above serves as one of the major sources for the design of learning events to facilitate students' conceptual change involving combustion-related conceptions.

Many studies clearly point out that science teaching and learning theories are absent in most of the web-learning programs for science (Mioduser et al. 2000; Tuvi and Nachmias 2001). Second, pedagogical considerations are also not considered in the design of most web-learning programs. Third, the capability of the web and multimedia is not fully used to

enhance students' science learning. Therefore, this current study attempts to unite the Dual Situated Learning Model (DSL_M) (She 2002, 2003, 2004a, b; Tang et al. 2005) and scientific reasoning for the development of a SCCR (Scientific Concept Construction and Reconstruction) digital learning program in order to facilitate students' knowledge construction and reconstruction for the concept of combustion. The SCCR project includes the theories of DSL_M, scientific reasoning and adaptive learning into the design of the learning content. DSL_M is a conceptual change theory and also an instructional design guidance that is tailored to foster students' conceptual change in scientific concepts. The DSL_M has been applied in the classroom for several topics and has demonstrated successful fostering of students' conceptual change within the classroom setting. Many studies have proposed that scientific reasoning is pivotal for conceptual change. The limitation of employing DSL_M in a classroom setting is that it has not been able to provide different students with different paths of learning according to their alternative conceptions and reasoning patterns. Adaptive learning is important because it enables learners to select their modular components to customize their learner-centered learning environments. Secondly, it enables the system to offer flexible solutions that dynamically adapt content to fit individual real-time learning needs. Therefore, it is necessary to move beyond the classroom instruction to an adaptive web-learning approach. This can provide different learning events according to the response learners made regarding either accuracy of the concepts or the reasoning behind them, and it can further determine which learning pathway the students need to learn by. It is hoped that by combining DSL_M and scientific reasoning, the efficiency and effectiveness of students' scientific conceptual change and scientific reasoning could be maximized.

Research Questions

This study examines the potential of maximizing students' success in both conceptual change and scientific reasoning ability involving the topic of "combustion" by combining scientific reasoning and conceptual change theories into the design of a SCCR adaptive digital learning program. The following research questions serve as guidelines for this study. First, do students who engage in SCCR adaptive digital learning (experimental group) perform significantly better than the conventional group, regardless of combustion concepts understanding, conceptual change and scientific reasoning ability? Second, does students' level of scientific reasoning have any impact on their understanding of combustion concepts, conceptual change, and scientific reasoning? Third, are there any interactions between instructional approaches and students' levels of scientific reasoning on these two factors? Fourth, what are the relationships between students' combustion concepts understanding, conceptual change, and scientific reasoning? Fifth, what are the nature and extent of students' use of scientific reasoning, correct concepts and conceptual change between the experimental and the conventional group throughout the learning?

Methods

Design of the Combustion Learning Unit

The design of combustion instructional content is based on the six stages of the DSL_M described above. Experienced teachers' suggestions for the nature of the combustion content, together with previous studies of students' alternative conceptions (Andersson

1990; Boujaoude 1991; Prieto et al. 1992; Schollum and Happs 1982; Watson et al. 1995) serve as the major sources to decide what mental sets students lack for constructing a more scientific view of combustion.

The mental sets students lack were categorized into five topics to further develop a series of dual situated learning events: (1) Requirements for combustion: combustible materials, flammable gas, and temperature increase to the flash point. (2) Oxygen: production of oxygen, oxygen is a combustion-supporting gas, oxygen is needed for burning and its function in the burning, and the nature of oxygen such as the density and its solubility. (3) Carbon dioxide: production of carbon dioxide, testing for the existence of carbon dioxide, carbon dioxide is produced during burning, the qualities of carbon dioxide such as density and solubility. (4) Fire extinguishing: the methods and reasons for a fire being extinguished. (5). Rust: nature of rusting substances, reasons for rusting, requirements for rusting, and protection from rusting. It is clear that students' alternative conceptions of burning originate from their personal observations, and since there are many substances they can not visualize, they can not use conservation ideas to explain the products of burning. Therefore, the design of this combustion unit provides students with many demonstrations and multiple representations to visualize the real events in order to help students develop a more scientific view of combustion and corrosion.

Each dual situated learning event requires students to provide an answer and their reasons to the driving question before proceeding with various activities. The same driving question is asked again after the event to examine the students' conceptual change as well as their reasoning changes. The SCCR adaptive digital learning system will provide students with a different learning pathway according to the accuracy of each individual's answers and reasons. In order to activate students' use of scientific reasoning, therefore, we specifically restructure the process to let students view their answers and their written reasons both before and after events, as well as the correct answers. The system then requests students to justify why they changed or stayed with their original thoughts after the learning events. This draws students into the process of scientific reasoning as well as conceptual change (Appendix 1).

Instruments

Combustion Achievement Test (CAT)

The CAT is a multiple-choice diagnostic instrument that was developed to evaluate students' concepts of combustion related concepts. The content validity was established by a panel composed of two science teachers and one science educator, ensuring that the items were properly constructed and relevant to the combustion unit we developed. The questions require students to use deeper information processing ability and are mainly for analysis and synthesis. There are 96 items (Appendix 2). Students receive one point if they answer one question right, so the highest possible score is 96. The Cronbach's α for the CAT was 0.90 for the pretest, 0.93 for the posttest, and 0.94 for the retention-test.

Scientific Reasoning Test (SRT)

The SRT is a two-tier multiple choice diagnostic instrument developed to measure students' scientific reasoning, and it was modified from Lawson's scientific reasoning test. It measures students' conservation, proportional thinking, identification and control of variables, probabilistic thinking, correlative thinking and hypothetic-deductive ability.

There are 12 items, each of which contains two tiers; the first tier requires students to choose the answer, and the second tier requires students to use the thinking mentioned above (Appendix 3). Students need to answer both tiers correctly in order to receive one point, so the highest score can be 12. Students SRT scores of 0–2 were classified at the lower level corresponding generally to Piaget's pre-concrete operational level. Scores of 3–4 were classified at the lower level corresponding generally to Piaget's post-concrete operational level. Scores of 5–8 were classified as transitional level reasoners and those who scored from 9–12 were classified as the higher level that corresponds generally to Piaget's formal operational stage (Lawson 1978). The Cronbach's α of the Scientific Reasoning Test (SRT) was 0.71 for the pretest, 0.61 for the posttest, and 0.76 for the retention-test, which is close to Lawson's result for the Cronbach's α .

Combustion Dependent Reasoning Test (CDRT)

The CDRT is a two-tier multiple choice diagnostic instrument that was developed to measure the degree of students' conceptual change involving combustion. The content validity was established by the same panel of three evaluators, ensuring that the items were properly constructed and relevant to the combustion unit. There are 41 items, and each item contains two tiers, the first tier to diagnose students' concepts, and the second tier to diagnose students' reasons for their concepts (Appendix 4). Thus, students need to answer two tiers correctly in order to receive one point. The highest score can be 41. The Cronbach's α of the Combustion Dependent Reasoning Test (CDRT) was 0.89 for the pretest, 0.92 for the posttest, and 0.94 for the retention-test.

Interview Results Analyses

Students' interview results were transcribed and then analyzed by a flow map method (Anderson and Demetrius 1993). There were 31 questions used for interviewing students to determine their concepts related to the processes of combustion, oxygen, carbon dioxide, extinguishing fire, and rustiness (Appendix 5). A flow map is constructed by diagramming the interviewee's verbalization of thought as it unfolds, and this map also can display the sequence and complex thought patterns expressed by the interviewee. Flow maps have been employed in several studies and it has been suggested that flow map analyses are a useful and valid method of representing students' conceptual frameworks for science (Anderson and Demetrius 1993; Demetrius 1998).

Students' flow maps were further analyzed into two major components: complexity of scientific reasoning and quantity of conceptual change.

To evaluate the complexity of scientific reasoning, this study uses four dimensions of scientific reasoning modified from Hogans' study (2000): Generativity (G), Elaboration (EL), Justification (J), and Explanation (Ex).

1. Generativity (G): the number of students' own observations, ideas, or conjectures. For instance, "I do not know" was coded as G0 because it did not answer anything relevant; while "water can be used to extinguish the fire" was coded as G1.
2. Elaboration (EL): the number of scientific concepts used to describe and explain the concepts. For instance, "oxygen can help with burning." was coded as EL1; "air is composed of many different gases, oxygen is one of them, and carbon dioxide is one of them." was coded as EL2.
3. Justification (J): the number of evidence-based or inference-based or experiments that were used to support the ideas or assertions. For instance, "the solubility of carbon

dioxide is higher than oxygen because we did the experiment and it shows that the water level in the carbon dioxide test tube is higher than in the tube with oxygen.” was coded as J1; “A balloon would float in the air if we fill it with oxygen, because the density of oxygen is less than air. Air is composed of many different gases and their density is different, some of them are denser than oxygen, and some of them are lighter than oxygen; when mixed together, the density of air is greater than oxygen.” was coded as J2.

4. Explanation (EX): the number of proposed rules or mechanisms to explain the concepts or assertions. For instance, “The reason for putting the cap on the alcohol burner is to extinguish the fire; it cuts off one of the requirements for burning, which is the combustion-supporting gas of oxygen.” was coded as EX 1; “To prove that the mechanism of a metal nail rusting is related to oxygen, the experimental design uses a paraffin seal around a metal nail with oxygen inside and another paraffin seal with air inside. It is hoped to determine whether oxygen or air is the major component of causing serious rusting. If the nail with oxygen has more serious rusting than the other one, then it proves that oxygen would foster corrosion more than the air.” was coded as EX2.

To quantify (measure) conceptual change, the following dimensions were considered:

1. Correct concept: whether students’ alternative conceptions decrease or increase throughout the learning.
2. Nature and quantity of conceptual change: the number of linkages between pre- and post-flow maps, or between post- and retention-flow maps shows the nature and quantity of conceptual change. (e.g., Fig. 1)
 - a. Progress (PG): the number of linkages shows how much the student’s concept progressed, this indicates how much that they changed their alternative concept to the correct concept.
 - b. Maintain Correctness (MCT): the number of linkages shows that the student’s correct concept was maintained correct.
 - c. Maintain incorrect/partial correct (MI/PC): the number of linkage shows the student’s concepts were maintained as either partially correct or incorrect, which means their alternative concept did not change.
 - d. Retrogression (RTG): the number of linkage shows that the student’s concept retrogressed, which means that their original concept was correct, but, alternative concepts emerged after learning.

Samples, Design, and Data Analysis Procedures

A total of 61 sixth graders recruited from two average-achievement classes in a middle school participated in this study. One class of students (experimental group) received the SCCR digital learning program and the other class of students received conventional instruction. Both groups of students also did the same experiments in a laboratory. These two classes were taught by two different teachers. Both programs covered the same content and sequence but with different ways of presenting the content and different ways of instruction. The major topics were: requirements for combustion, ways of producing oxygen and its characteristics, ways of producing carbon dioxide and its characteristics, the methods and reasons for extinguishing fire, and requirements for rusting. For both classes, the combustion section lasted eight class periods (each class period is about 40 min) over

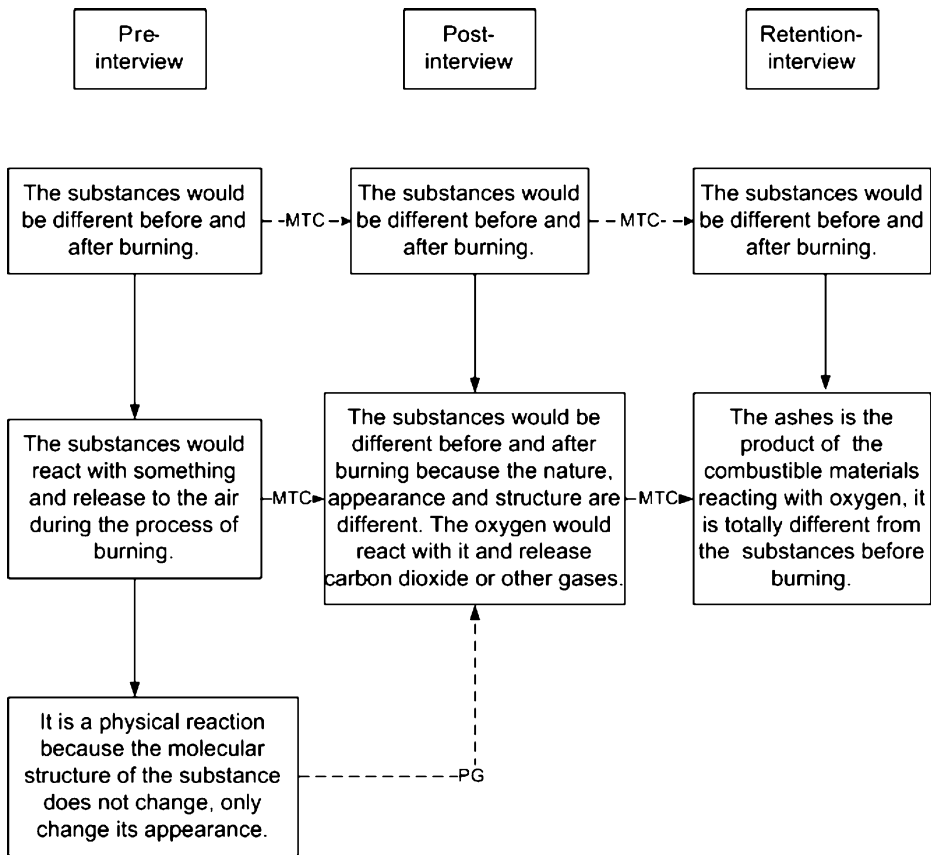


Fig. 1 Flow Map of Combustion. Note 1. Scientific Reasoning: Generativity (G), Elaboration (EL), Justification (J), Explanation (EX). Note 2. Conceptual Change: Progress (PG), Maintain-Correct (MTC), Maintain incorrect /Partial Correct (MI/PC), Retrogress(RTG)

4 weeks. The conventional group's teacher mainly used lectures and discussions for teaching. The experimental group's students received different learning pathways according to the accuracy of their individual responses on concepts and reasoning while learning with the SCCR digital learning program. Each student in the experimental group must be actively involved in the learning process, and the system provides feedback to the students. In particular, they need to justify their concepts and reasons if they change after learning the events. The teacher serves as a facilitator to solve their problems while they are learning. The teacher only provides them with SCCR and the lab experiment, without extra lectures or discussions in order to simplify the experimental design. One limitation of this study is that two different teachers were responsible for the two groups.

Two-factor MANCOVA design was used for this study. In addition to the first factor of instructional approach, the students' level of scientific reasoning was included as the second factor. The pre-SRT scores were used to differentiate students' levels of scientific reasoning. Combustion-Dependent Reasoning Test (CDRT), Scientific Reasoning Test (SRT), and Combustion Achievement Test (CAT) were administered to the students in both groups before, 1 week after and sixth weeks after the instruction.

In addition, students' interview data were used to support the quantitative results. A total of 12 students (six students from each group) were interviewed with full audio tape recordings before, directly after and 6 weeks after instruction. Students were selected for interviews as follows: two high achievers (one male and one female), two middle achievers (one male and one female), and two low achievers (one male and one female). Students' interview results were transcribed and then analyzed using a flow map. Students' flow maps were further analyzed into two major parts: complexity of scientific reasoning and quantity of conceptual change. MANCOVA was used to compare the two groups of students' complexity of scientific reasoning and correct concepts. A T test was used to compare both experimental and conventional groups' quantity of conceptual change, using the mean difference scores of the pre-post flow map and post-retention flow map for each category of conceptual change (PG, MCT, MI/PC, RTG).

Due to the sample size limitation of interviewing, there were only six students from each group, so the effect size (Cohen's d for T test and $\text{partial}\eta^2$ for MANCOVA) was used to measure its practical significance. Cohen (1988), and Vacha-Haase and Thompson (2004) defined effect size as small ($0.06 > \text{Partial}\eta^2 \geq 0.01$), medium ($0.14 > \text{Partial}\eta^2 \geq 0.06$), and Large ($\text{Partial}\eta^2 \geq 0.14$), in light of the nature and characteristics of behavioral or social science studies.

Results

Combustion Achievement Test (CAT)

Two-factor MANCOVA was conducted to examine the effects of instructional approaches, scientific reasoning levels, and using post- and retention- of CAT scores as the dependent measures, and students' pre-CAT scores as the covariate. Table 1 summarizes the results of the two-factor MANCOVA, indicating that only the instructional approach (Wilk's $\Lambda=0.55$, $p=0.000$) has a statistically significant effect on the performance of post- and retention-CAT. There is no interaction between instructional approaches and scientific reasoning levels. Therefore, the following main effect for the instructional approaches was evaluated.

First, one-factor MANCOVA was performed to examine the effect of the instructional approach factor on both post-test and retention-test, as shown in Table 2. Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the instructional approach on post- and retention-CAT. It can be seen that there is a significant effect for instructional approaches on both post-CAT ($F=43.98$, $p=0.000$) and retention-test ($F=10.59$, $p=0.002$). Thus, the students' post- and retention-CAT were significantly affected by the instructional approaches. The post-hoc analysis for main effect suggests that the experimental group performed significantly better than the conventional group on post- and retention-CAT (experimental group > conventional group, $p_{(\text{post})}=0.000$, $p_{(\text{retention})}=0.002$) (Table 2). This clearly demonstrates that the experimental group students outperformed the conventional group on both post- and retention-CAT.

Scientific Reasoning Test (SRT)

Two-factor MANCOVA was conducted to examine the effects of instructional approaches and scientific reasoning levels using post- and retention-SRT scores as the dependent measures, and students' pre-SRT scores as the covariate. Table 1 summarizes the results of the two-factor MANCOVA, indicating that only the instructional approach has a statistically significant

Table 1 Multivariate Analysis of Covariate (MANCOVA) of post- and retention- of Combustion Achievement Test (CAT), Scientific Reasoning Test (SRT), and Combustion Dependent Reasoning Test (CDRT)

Source of variance	Wilk's Λ	Partial η^2	Effect size	Hypothesis df	Error df	Multivariate F
Combustion Achievement Test (CAT)						
Covariates						
Pre-test scores	0.78	0.218	L	2	53	7.38**
Group memberships						
Instructional approaches	0.55	0.449	L	2	53	21.70***
Levels of Scientific Reasoning	0.90	0.050	S	4	106	1.38
Scientific Reasoning Test (SRT)						
Covariates						
Pre-test scores	0.75	0.254	L	2	53	9.01***
Group memberships						
Instructional approaches	0.81	0.192	L	2	53	6.29*
Levels of Scientific Reasoning	0.89	0.058	S	4	106	1.63
Combustion Dependent Reasoning Test (CDRT)						
Covariates						
Pre-test scores	0.67	0.328	L	2	53	12.91***
Group memberships						
Instructional approaches	0.72	0.284	L	2	53	10.54***
Levels of Scientific Reasoning	0.132	0.132	M	4	106	4.04*

*** $p < 0.0001$, ** $p < 0.001$, * $p < 0.01$

* $0.0099 < \text{partial } \eta^2 < 0.0588$, effect size = small; ** $0.0588 < \text{partial } \eta^2 < 0.1379$, effect size = medium; *** $\text{partial } \eta^2 > 0.1379$, effect size = large

effect on the performance of post- and retention-SRT (Wilk's $\Lambda = 0.81$, $p = 0.004$). There is no interaction between instructional approaches and scientific reasoning levels. Therefore, the following main effect for the instructional approaches was examined.

One-factor MANCOVA was performed to examine the effect of instructional approach on both post-test and retention-test, as shown in Table 2. Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the instructional approaches on post- and retention-SRT. It can be seen that there was a significant effect for instructional approaches on post-SRT ($F = 12.21$, $p = 0.000$). Thus, the students' post-SRT was significantly affected by the instructional approach. The post-hoc analysis for main effect suggests that the experimental group performed significantly better than conventional group on the post-test (experimental group > conventional group, $p_{(\text{post})} = 0.001$) (Table 2). This clearly demonstrates that experimental group students outperformed the conventional group on the post-SRT.

Combustion Dependent Reasoning Test (CDRT)

Two-factor MANCOVA was conducted to examine the effects of instructional approaches and scientific reasoning levels using post- and retention- CDRT scores as the dependent measures, and students' pre-CDRT scores as the covariate. Table 1 summarizes the results of the two-factor MANCOVA indicating that both instructional approaches (Wilk's $\Lambda = 0.72$, $p = 0.000$) and students' level of scientific reasoning (Wilk's $\Lambda = 0.75$, $p = 0.004$)

Table 2 MANCOVA and Post-hoc of Instructional Approaches and 3-Levels of Scientific Reasoning of Post- and Retention- of Combustion Achievement Test (CAT), Scientific Reasoning Test (SRT), and Combustion Dependent Reasoning Test (CDRT)

	Wilk's Λ	Partial η^2	Effect size	Univariate F		
				Post-test	Retention-test	Post-hoc
Combustion Achievement Test (CAT)						
Instructional Approaches	0.55*** (0.000)	0.449	L	43.98**	10.59*	Post: E>C (0.000) Retention: E>C (0.002)
Scientific Reasoning Test (SRT)						
Instructional Approaches	0.81* (0.004)	0.192	L	12.21**	0.01	Post: E>C (0.001)
Combustion Dependent Reasoning Test (CDRT)						
Instructional Approaches	0.72*** (0.000)	0.284	L	18.27***	11.22**	Post: E>C (0.000) Retention: E>C (0.001)
Levels of Scientific Reasoning	0.75* (0.004)	0.132	M	1.21	6.50*	Retention: Post-concrete operational > pre-concrete operational (0.001); transitional > pre-concrete operational(0.018)

E experimental group, C conventional group

*** $p < 0.0001$, ** $p < 0.001$, * $p < 0.01$

reached a statistically significant effect on the performance of post- and retention- of CDRT. There was no interaction between instructional approaches and scientific reasoning levels. Therefore, the following main effects were performed for the instructional approaches and level of scientific reasoning separately.

First, one-factor MANCOVA was performed to examine the effect of instructional approach factor on both post- and retention-CDRT, as shown in Table 2. Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the instructional approaches on post- and retention-CDRT. The results indicated a significant effect for instructional approaches on both post-CDRT scores ($F = 18.27, p = 0.000$) and retention-CDRT ($F = 11.22, p = 0.001$). Thus, the students' post- and retention-ADRT were significantly affected by the instructional approach. The post-hoc analysis for main effect suggests that the web-based instructional group performed significantly better than the conventional group on post- and retention-CDRT (experimental group > conventional group, $p_{\text{post}} = 0.000, p_{\text{retention}} = 0.001$) (Table 2). This clearly demonstrates that experimental group students outperformed the conventional group on both post- and retention-CDRT.

Second, the same procedures were performed to examine the effect of the level of scientific reasoning, as shown in Table 2. Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the level of scientific reasoning on post- and retention-CDRT. This indicated a significant effect for the level of scientific reasoning on retention-test ($F = 6.50, p = 0.003$). Thus, the students' retention-CDRT was significantly affected by the level of the scientific reasoning. The post-hoc analysis for main effect suggests that the transitional level students performed significantly better than pre-concrete operational level students, and post-concrete operational level students performed better than pre-concrete operational students on retention-CDRT (transitional > pre-concrete

operational, $p_{(\text{retention})}=0.018$; post-concrete operational > pre-concrete operational, $p_{(\text{retention})}=0.001$) (Table 2). This clearly demonstrates that transitional level students outperformed those in the pre-concrete operational level, and post-concrete operational level students outperformed the pre-concrete operational level on retention-CDRT.

Correlation Results

Table 3 summarizes the correlations among students' conceptual change, scientific reasoning, and combustion concepts. The results show that CAT, SRT, and CDRT are highly correlated with each other and reached a 0.0001 statistically significant difference level for pre- or post- and retention-tests. The only two exceptions are that pre-CAT correlated with retention-SRT and retention-SRT correlated with retention-CDRT, and both reached 0.005 statistically significant levels (Table 3). This indicates that students' conceptual change is highly correlated with their level of scientific reasoning and combustion concepts.

Interview Results

Students were interviewed before, soon after and 6 weeks following the instruction. Students' interview results were transcribed and then analyzed by a flow map method. The flow map data were further analyzed in two parts: complexity of scientific reasoning and degree of conceptual change.

Complexity of Scientific Reasoning

The MANCOVA and post-hoc test were used to compare both experimental and conventional students' complexity of scientific reasoning across the 31 interview questions, using the mean scores of post-flow map and retention-flow map for each category of scientific reasoning (G, EL, J, EX) as the dependent measures, and students' mean scores of pre-flow map for each category of scientific reasoning serve as covariate (Table 4). The better results for the experimental group over the conventional group on the post-flow map dimensions of Generativity (G), Elaboration (EL), Justification (J) and Explanation (EX) for nine, 15, seven, and one out of 31 interview questions were practically significant. In contrast, the higher mean scores for the conventional group on the post-flow map dimensions of Generativity (G), Elaboration (EL), and Justification (J) were practically significant for 14, eight, and two questions. This indicates that the students in the experimental group used a higher level of scientific reasoning (Elaboration and Justification) more frequently than the conventional group. In contrast, the conventional group used a lower level of scientific reasoning (Generativity) more frequently than the experimental group immediately after learning.

The better results for the experimental group over the conventional group on the retention-flow map dimensions of Generativity (G), Elaboration (EL), Justification (J) and Explanation (EX) for 11, 14, seven and one questions were practically significant. In contrast, the higher mean scores for the control group on the retention-flow map dimensions of Generativity (G), Elaboration (EL), and Justification (J) were practically significant for 16, 6 and one questions. After the 6th week of learning, the students in the experimental group used higher levels of scientific reasoning (Elaboration and Justification) more frequently than did the conventional group. In contrast, the conventional group used a lower level of scientific reasoning (Generativity) more frequently than in the experimental group.

Table 3 Correlation results of pre-, post-, and retention- of CAT, SRT, and CDRT

	Pre-CAT	Post-CAT	Retention-CAT	Pre-SRT	Post-SRT	Retention-SRT	Pre-CDRT	Post-CDRT	Retention-CDRT
Pre-CAT	1***	0.76***	0.82***	0.66***	0.59***	0.51**	0.70***	0.74***	0.71***
Post-CAT		1***	0.92***	0.74***	0.67***	0.68***	0.83***	0.86***	0.86***
Retention-CAT			1***	0.73***	0.63***	0.66***	0.81***	0.85***	0.89***
Pre-SRT				1***	0.70***	0.62***	0.73***	0.68***	0.73***
Post-SRT					1***	0.71***	0.71***	0.70***	0.62***
Retention-SRT						1***	0.67***	0.62***	0.57***
Pre-CDRT							1***	0.77***	0.78***
Post-CDRT								1***	0.81***
Retention-CDRT									1***

*** $p < 0.0001$, ** $p < 0.005$, * $p < 0.01$

Table 4 Summary of post hoc results for students' complexity of scientific reasoning

Complexity of scientific reasoning Experimental vs. Conventional	Generativity	Elaboration	Justification	Explanation
	No. of questions	No. of questions	No. of questions	No. of questions
Post-flow map				
Experimental > Conventional	9	15	7	1
Conventional > Experimental	14	8	2	0
Retention-flow map				
Experimental > Conventional	11	14	7	1
Conventional > Experimental	16	6	1	0

Experimental: experimental group, Conventional: conventional group

No. of questions: number of interviewing questions at specific scientific reasoning category reached practically significant difference level between experimental group and conventional group

Measure of Conceptual Change

Quantity of Correct Concept

MANCOVA was used to compare both experimental and conventional students' quantity of correct concepts, using the mean scores of post-flow map and retention-flow map for correct concepts while the dependent measures and students' mean scores of pre-flow map for correct concepts served as covariate (Table 5). The better result for the experimental group over the conventional group on the post-flow map and retention-flow map dimension of correct concepts for 20 and 19 out of the 31 interview questions were practically significant. On the contrary, the higher mean scores for the conventional group on the post-flow map and retention-flow map dimension of correct concepts were practically significant for seven and six questions. In summary, the students in the experimental group performed better than the conventional group on the quantity of correct concepts across all 31 interview questions, either 1 week or 6 weeks after learning.

Quantity of Conceptual Change

T test was used to compare both experimental and conventional groups' quantity of conceptual change, using the mean difference scores of the pre-post flow map and post-retention flow map for each category of conceptual change (PG, MCT, MI/PC, RTG). The better results for the experimental group over the conventional group on the pre-post flow map dimensions of Progress (PG), Maintain Correct (MCT), and Maintain Incorrect/Partially Correct (MI/PC) for 16, eight, and six questions out of the 31 interview questions were practically significant. In contrast, the higher mean scores for the conventional group

Table 5 Analysis of quantity of correct concepts

	Wilk's Λ	Partial η^2	Effect size	Univariate F		
				Post-test	Retention-test	Post-hoc
Combustion 1	0.656	0.344***	L	0.897**	3.065***	Post: E > C; Retention: E > C
Combustion 2	0.743	0.257***	L	1.849***	0.021	Post: C > E
Combustion 3	0.955	0.045*	S	0.004	0.297*	Retention: C > E
Combustion 4	0.851	0.149***	L	0.341*	0.341*	Post: E > C; Retention: E > C
Combustion 5	0.754	0.246***	L	0.769**	2.185***	Post: E > C; Retention: E > C
Combustion 6	0.077	0.923***	L	11.607***	0.613**	Post: E > C; Retention: E > C
Oxygen1	0.763	0.237***	L	1.544***	0.121*	Post: E > C; Retention: E > C
Oxygen2	0.727	0.273***	L	1.219***	0.061**	Post: C > E; Retention: E > C
Oxygen3	0.758	0.242***	L	1.227***	0.923***	Post: E > C; Retention: E > C
Oxygen4	0.073	0.927***	L	8.939***	60.680***	Post: E > C; Retention: E > C
Oxygen5	0.841	0.159***	L	0.799**	0.107*	Post: C > E; Retention: E > C
Oxygen6	0.998	0.002	M	0.005	0.000	–
Oxygen7	0.911	0.089**	M	0.193*	0.118*	Post: C > E; Retention: C > E
Carbon dioxide 1	0.179	0.821***	L	0.675***	3.179***	Post: E > C; Retention: C > E
Carbon dioxide 2	0.826	0.174**	M	1.461***	0.143*	Post: E > C; Retention: E > C
Carbon dioxide 3	0.704	0.296***	L	1.618***	1.397***	Post: E > C; Retention: E > C
Carbon dioxide 4	0.824	0.176***	L	2.558***	0.310*	Post: E > C; Retention: E > C
Carbon dioxide 5	0.889	0.111**	M	.875**	–	Post: E > C
Carbon dioxide 6	0.795	0.205***	L	1.487***	0.002	Post: C > E
Extinguish fire 1	0.710	0.290***	L	0.350*	3.645***	Post: E > C; Retention: E > C
Extinguish fire 2	0.736	0.264***	L	0.068*	1.629***	Post: E > C; Retention: E > C
Extinguish fire 3	0.820	0.180***	L	0.937***	0.610**	Post: E > C; Retention: E > C
Extinguish fire 4	0.900	0.100**	M	–	0.444**	Retention: E > C
Extinguish fire 5	–	–	–	–	–	–
Extinguish fire 6	0.871	0.129**	M	0.533**	0.221*	Post: E > C; Retention: C > E
Extinguish fire 7	0.492	0.508***	L	0.733***	0.043*	Post: E > C; Retention: E > C

Table 5 (continued)

	Wilk's Λ	Partial η^2	Effect size	Univariate F		
				Post-test	Retention-test	Post-hoc
Rustiness1	0.713	0.287***	L	2.210***	0.111*	Post: C > E; Retention: E > C
Rustiness2	0.765	0.235***	L	0.172*	1.005***	Post: C > E; Retention: C > E
Rustiness3	0.873	0.127***	L	0.563**	0.176*	Post: E > C; Retention: C > E
Rustiness4	0.922	0.078**	M	0.508**	0.145*	Post: E > C; Retention: E > C
Rustiness5	0.508	0.492*	S	1.005***	0.005	Post: E > C

E Experimental group, C conventional group

*0.0099 < partial η^2 < 0.0588, effect size = small; **0.0588 < partial η^2 < 0.1379, effect size = medium; ***partial η^2 > 0.1379, effect size = large

on the pre-post flow map dimensions of Progress (PG), Maintain Correct (MCT), Maintain incorrect/partial correct (MI/PC), and Retrogress (RTG) were practically significant for four, 15, ten, and two questions, respectively. In summary, the students in the experimental group change their alternative concepts more successfully than the conventional group, whereas the students in the conventional group mainly stay with their original correct concepts right after learning (Figs. 2, 3). This indicates that the students in the experimental group made more progress than conventional group in changing their alternative concepts after learning.

The better results for the experimental group over the conventional group on the post-retention flow map dimensions of Progress (PG), Maintain Correct (MCT), Maintain Incorrect/Partially Correct (MI/PC), and Retrogression (RTG) were practically significant for four, 16, three, and four questions. In contrast, the higher mean scores for the conventional group on the post-retention flow map dimensions of Progress (PG), Maintain

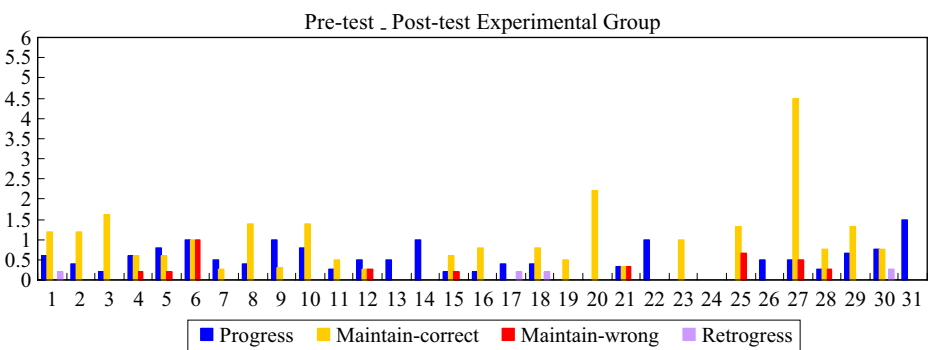


Fig. 2 Distribution of means scores of students' conceptual change across 31 interview results from pre- to post-test of experimental group

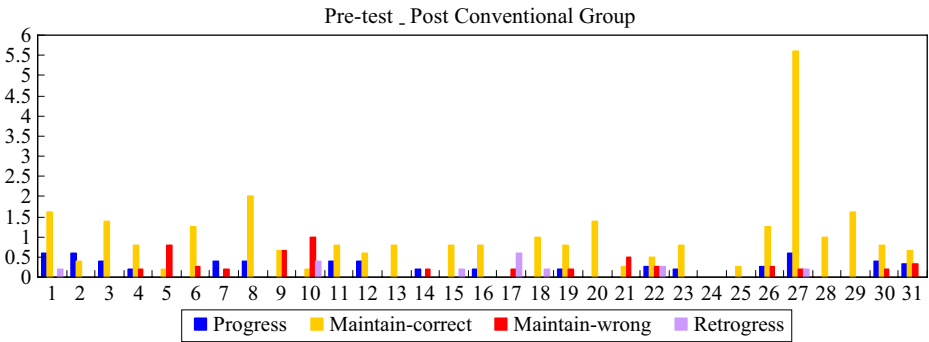


Fig. 3 Distribution of means scores of students’ conceptual change across 31 interview results from pre- to post-test of conventional group

Correct (MCT), Maintain incorrect/partial correct (MI/PC), and Retrogress (RTG) were practically significant for six, seven, 14, and seven questions, respectively. Six weeks after learning, it shows that the students in the experimental group change their alternative concepts more successfully than the conventional group because once the students in the experimental group acquire the scientific concepts, then these are maintained beyond the 6th week after testing (Figs. 4, 5). On the contrary, the conventional group’s alternative concepts increased greatly 6 weeks after learning which indicated that conventional students’ conceptual change is very unstable after a period of time. In short, once the students in the experimental group successfully change their alternative concepts, it is more likely to be stable even after a period of time.

Discussion and Conclusions

The two-factor MANCOVA also indicates that the instructional approach has a significant effect on students’ performance on post- and retention-tests of CAT, SRT and CDRT; and post hoc analysis further suggests that the experimental group performed significantly better

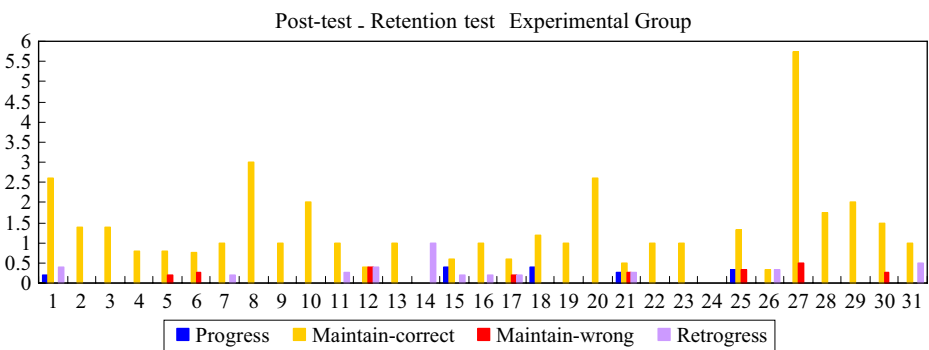


Fig. 4 Distribution of means scores of students’ of conceptual change across 31 interview results from post- to retention-test of experimental group

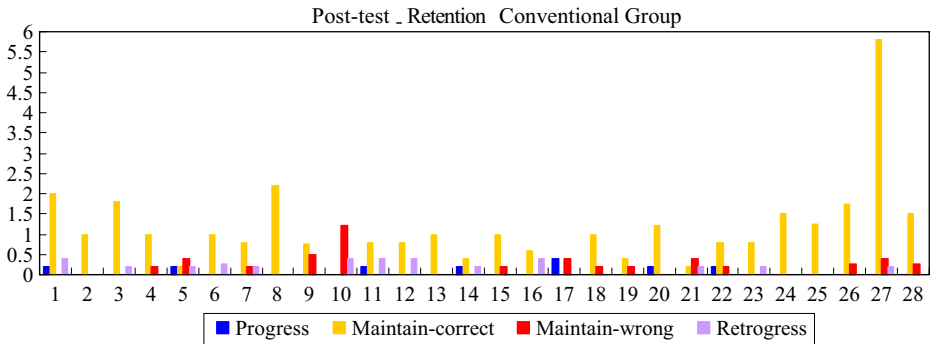


Fig. 5 Distribution of means scores of students' conceptual change across 31 interview results from post- to retention-test of conventional group

than the conventional group on the post- CAT, SRT and CDRT and retention-CAT and CDRT. These quantitative results clearly reveal that the experimental group outperformed the conventional group, regardless of students' concepts of combustion, scientific reasoning ability and conceptual change. These results are quite significant for the idea that experimental students' conceptual change and their scientific reasoning ability can be promoted through learning from a well-designed and theory-based SCCR adaptive learning program.

The two-factor MANCOVA also indicates that the level of scientific reasoning has a significant effect on students' performance on post- and retention-test of CDRT, and the post hoc test shows that the level of scientific reasoning has effect only on retention-CDRT. Students with higher levels of scientific reasoning ability would change their alternative conceptions related to combustion more efficiently than students with lower levels of scientific reasoning ability. This may be because students with higher levels of scientific reasoning ability would be more able to spontaneously use their higher level of scientific reasoning to explain logically both how and why the evidence supports their scientific conceptions, and why it contradicts their alternative conceptions. Several previous studies have proposed similar ideas (Lawson and Weser 1990; Oliva 2003).

The correlation results shows that CAT, SRT and CDRT are highly correlated with each other, for both pre- and post- or retention-tests. This explains why students with a high score on SRT would also have a high score on their CDRT, and vice versa. In short, it provides empirical evidence to support the previous researchers' suggestions that conceptual change involves deep restructuring in both the concepts and also in the way of reasoning (Gil-Perez and Carrascosa-Alis 1994; Furio et al. 2000). Moreover, it also implies that students' scientific reasoning ability increases along with conceptual change.

The interview data were transformed into flow maps and further analyzed according to three major dimensions: complexity of scientific reasoning, quantity of correct concepts and quantity of conceptual change. The dimension of complexity of scientific reasoning indicated that the experimental group used a higher level of scientific reasoning (Elaboration and Justification) more frequently than the conventional group, both immediately after learning and on the 6th week after learning. However, the conventional group used a lower level of scientific reasoning (Generativity) more frequently than experimental group both immediately after learning and on the 6th week after learning. This

again supports the quantitative data of SRT results indicating that experimental students outperform conventional students on the post-SRT. It implies that students' scientific reasoning ability improved within a very short period by combining conceptual change and scientific reasoning in the design of a digital learning program. Lawson (2003) has reported that students' scientific reasoning ability can be improved within a semester, which is longer than our study, where we found students' scientific reasoning could be enhanced within a number of weeks.

Our study further shows that the experimental group performed better than the conventional group on the quantity of correct concepts both immediately after learning and on the 6th week after learning. Again, experimental students acquired correct concepts more efficiently than conventional students. This supports the observation of Cepni and Keles (2006) that computer-assisted instructional materials are effective for reaching comprehension and application levels of cognitive domain. Our measurement of correct concepts covers comprehension, analyses, application, and synthesis levels of cognitive domain. In addition, it also supports the results of the quantitative results for CAT that the experimental group outperforms the conventional group on both post- and retention-test.

The students in the experimental group changed their alternative concepts more successfully than the conventional group, although conventional students mainly stayed with their original correct concepts right after learning. By the 6th week after learning, it can be seen that once the experimental students acquired the scientific concepts, this understanding lasted longer than 6 weeks after learning. In contrast, conventional students' alternative concepts increased after 6 weeks of learning. This again supports the quantitative data of CDRT results that students in the experimental group changed their alternative concepts more effectively than the conventional group either right after learning or 6th weeks after learning. Both our qualitative and quantitative results are different from the study by Cepni and Keles (2006), which found that computer assisted materials (CAIM) did not change students' alternative conceptions related to photosynthesis, although they changed students' comprehension and application levels of cognitive domain and their attitudes. Our results indicate that SCCR digital learning is very efficient for improving students' conceptual change as well as scientific reasoning from both quantitative and qualitative data within a very short period. This demonstrates that combining conceptual change and scientific reasoning into the design of adaptive digital learning indeed powerfully facilitates students' conceptual change and their scientific reasoning ability.

Notes This research is based on work supported and funded by the National Science Council (NSC 94, 95-2511-S-009-002), Taiwan, R.O.C.

Appendix 1

Students' answer and reasoning before and after the adaptive digital dual situated learning event

MyOpenText.ca Inc. 中國教育出版集團有限公司 中國教育出版集團有限公司 Microsoft Internet Explorer

Question before instruction 1

What are the requirements for burning? (Multiple Choice)

1. The combustible materials is required burning , such as candles, alcohol, gas.
Your reason is because candles are combustible materials.
2. The air is required to help burning.
Your reason is I have learned that air can support burning.
3. The temperature being high enough is a prerequisite.
Your reason is _____

Submit

MyOpenText.ca Inc. 中國教育出版集團有限公司 中國教育出版集團有限公司 Microsoft Internet Explorer

Question after instruction 1

What are the requirements for burning? (Multiple Choice)

1. The combustible materials is required burning , such as candles, alcohol, gas.
Your reason is because candles and alcohol are combustible materials and they are required for burning.
2. The air is required to help burning.
Your reason is Air contains oxygen is the combustion-supporting gas.
3. The temperature being high enough is a prerequisite.
Your reason is: if the temperature is not high enough then the burning would not happen.

Submit

MyOpenText.ca Inc. 中國教育出版集團有限公司 中國教育出版集團有限公司 Microsoft Internet Explorer

Multiple Choice Question

What are the requirements for burning? (Multiple Choice)

1. The combustible materials is required burning , such as candles, alcohol, gas.
Your reason is _____
2. The air is required to help burning.
Your reason is _____
3. The temperature being high enough is a prerequisite.
Your reason is _____

	Before	After	Correct Answer
[Answer]	1,2	1,2,3	1,2,3
[Reason]	1. Because candles are combustible material. 2. I have learned that air can support burning.	1. Because candles and alcohol are combustible materials and they are required for burning. 2. Air contains oxygen is the combustion-supporting gas. 3. If the temperature is not high enough then the burning would not happen.	-

Your later answer is **not the same** as the original.
Your explanation is: I only know there are three essential factors for burning. After wait:
Send

MyOpenText.ca Inc. 中國教育出版集團有限公司 中國教育出版集團有限公司 Microsoft Internet Explorer

Multiple Choice Question

What are the requirements for burning? (Multiple Choice)

1. The combustible materials is required burning , such as candles, alcohol, gas.
Your reason is _____
2. The air is required to help burning.
Your reason is _____
3. The temperature being high enough is a prerequisite.
Your reason is _____

	Before	After	Correct Answer
[Answer]	1,2	1,2,3	1,2,3
[Reason]	1. Because candles are combustible material. 2. I have learned that air can support burning.	1. Because candles and alcohol are combustible materials and they are required for burning. 2. Air contains oxygen is the combustion-supporting gas. 3. If the temperature is not high enough then the burning would not happen.	-

Your later answer is **not the same** as the original.
Your explanation is: I only know there are three essential factors for burning. After watching the instructional video, I've become more aware of the prerule of burning.
Send

Appendix 2

Combustion Achievement Test (CAT) example question

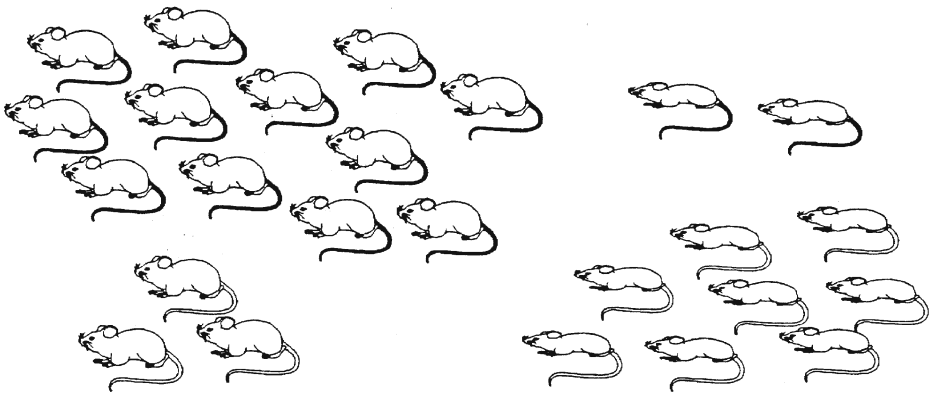
Which gas can help burning?

- vaporized water (steam)
- oxygen
- carbon dioxide
- all of them

Appendix 3

Scientific Reasoning Test (SRT) example question (Lawson 1978)

- Farmer Brown was observing the mice that live in his field. He discovered that all of them were either fat or thin. Also, all of them had either black tails or white tails. This made him wonder if there might be a link between the size of the mice and the color of their tails. So he captured all of the mice in one part of his field and observed them. Below are the mice that he captured.



Do you think there is a link between the size of mice and the color of their tails?

- appears to be a link
- appears not to be a link
- cannot make a reasonable guess

1.1. because

- there are some of each kind of mouse.
- There may be a genetic link between mouse size and tail color.
- There were not enough mice captured.
- Most of the fat mice have black tails while most of the thin mice have white tails.
- As the mice grew fatter, their tails became darker.

Appendix 4

Combustion Dependent Reasoning test (CDRT) example question

How does a cigarette-lighter light a candle?

- (1) Flames move from a cigarette-lighter to the candle, therefore, lighting the candle.
- (2) The temperature of a cigarette-lighter flame is high and reaches the ignition point of a candle, thus facilitating the burning.
- (3) The oil of a cigarette-lighter is heated and evaporated to gas, thus making the candle burn.
- (4) Lighting cigarette-lighter would induce the candle to burn.

Reasons

- (1) It provides combustible materials
- (2) It provides materials to help burning
- (3) It increases the temperature high enough to reach the ignition point.
- (4) all of them

Appendix 5

Interview Protocol

1. How can we make combustion happen? Explain.
2. Which materials can be used as fuels for burning in your daily life? Explain.
3. Can combustion happen without air? Explain.
4. What else do you need to light a candle besides a candle and air? Explain.
5. How does a cigarette-lighter light a candle? Explain.
6. Are all substances the same before and after burning? Explain.
7. Do you know how to produce oxygen? Tell me ways of producing oxygen that you know of. Explain.
8. What elements are present in the air? Which elements found in air can produce combustion? Explain.
9. What would be different if pure oxygen was used to help burning? Explain.
10. What would happen to a balloon if we fill it with pure oxygen? Would it fly up or go down? Explain why.
11. Is there oxygen in the water? Can you prove it?
12. Would pH value change if we add oxygen into water? Explain why.
13. What are the uses for oxygen? Explain.
14. Do you know how to produce carbon dioxide? Tell me ways of producing carbon dioxide that you know of. Explain.
15. Can carbon dioxide help combustion? Explain why.
16. Is there carbon dioxide in water? Can you prove it?
17. Which one would dissolve into water easier, oxygen or carbon dioxide? Explain why.
18. What would happen to a balloon if we fill it with carbon dioxide? Would it fly up or go down? Explain Why.
19. When will we use carbon dioxide in our daily life? Explain.
20. What are the ways to extinguish fire? Explain.

21. What are the theories on extinguishing fires while using gas stove? Explain.
22. What are the principles behind fire extinguishing by water? Explain.
23. What are the principles behind fire extinguishing by putting a cap on the alcohol burner? Explain.
24. How can we extinguish fire if alcohol spilled from the alcohol burner? Explain
25. How can we extinguish oil fire? Explain.
26. How can we extinguish fire in electrical equipment? Explain.
27. What substances will rust? Explain.
28. What phenomena occur while substances rust? Explain.
29. What are the conditions would make substance rusted? Explain Why
30. Can you design an experiment to prove your ideas regarding the conditions that cause rusting of substances?
31. What can be done to protect a substance from rusting? Explain.

References

- Anderson, J. R. (1990). *Cognitive psychology and its implications* (3rd ed.). New York: Freeman.
- Andersson, B. R. (1990). Pupils' conceptions of matter and its transformations (age 12–16). *Studies in Science Education*, 18(1), 53–85. doi:10.1080/03057269008559981.
- Anderson, O. R., & Demetrius, O. J. (1993). A flow-map method of representing cognitive structure based on respondents' narrative using science content. *Journal of Research in Science Teaching*, 30(8), 953–969. doi:10.1002/tea.3660300811.
- Boujaoude, S. B. (1991). A study of the nature of students' understandings about the concept of burning. *Journal of Research in Science Teaching*, 28(8), 689–704. doi:10.1002/tea.3660280806.
- Cepni, S., & Keles, E. (2006). Turkish students' conceptions about the simple electric circuits. *International Journal of Science and Mathematics Education*, 4(2), 269–291. doi:10.1007/s10763-005-9001-z.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Earlbaum Associates.
- Demetrius, O. J. (1998). *An investigation of the relationships between junior high school students' cognitive background variables and structure of recalled biological information. Paper presented at the annual meeting of National Association for Research in Science Teaching*. San Diego, CA.
- Driver, R. (1985). *Children's ideas in science*. Philadelphia, PA: Open University Press.
- Duschl, R., & Gitomer, D. (1991). Epistemological Perspectives on conceptual change: implications for educational practice. *Journal of Research in Science Teaching*, 28(9), 839–858. doi:10.1002/tea.3660280909.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). New York: Freeman.
- Furio, C., Calatayud, M. L., Barcenas, S. L., & Padilla, O. M. (2000). Functional fitness and functional reduction as common sense reasonings in chemical equilibrium and in geometry and polarity of molecules. *Science Education*, 84(5), 545–565. doi:10.1002/1098-237X(200009)84:5<545::AID-SCE1>3.0.CO;2-1.
- Gil-Perez, D., & Carrascosa-Alis, 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. doi:10.1002/scs.3730780310.
- Hashweh, M. Z. (1986). Towards an explanation of conceptual change. *International Journal of Science Education*, 8(3), 229–249.
- Hewson, P. W., & Thorley, N. R. (1989). The conditions of conceptual change in the classroom. *International Journal of Science Education*, 11(5), 541–553. doi:10.1080/0950069890110506.
- Hogan, K. (2000). Exploring a process view of students' knowledge about the nature of science. *Science Education*, 84(1), 51–70. doi:10.1002/(SICI)1098-237X(200001)84:1<51::AID-SCE5>3.0.CO;2-H.
- Johnson-Laird, P. N. (2000). Thinking: Reasoning. In A. Kazdin (Ed.), *Encyclopedia of psychology*, vol. 8 (pp. 75–79). Washington, DC: American Psychological Association.

- Kuhn, D. (1993). Science as Argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3), 319–337. doi:10.1002/sce.3730770306.
- Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15(1), 11–24.
- Lawson, A. E. (2003). *The neurological basis of learning, development and discovery*. London: Kluwer Academic.
- Lawson, A. E., & Weser, J. (1990). The rejection of nonscientific beliefs about life: Effects of instruction and reasoning skills. *Journal of Research in Science Teaching*, 27, 589–606. doi:10.1002/tea.3660270608.
- Lawson, A. E., & Worsnop, W. A. (1992). Learning about evolution and rejecting a belief in special creation: Effects of reflective reasoning skill, prior knowledge, prior belief and religious commitment. *Journal of Research in Science Teaching*, 29(2), 143–166. doi:10.1002/tea.3660290205.
- Mbajjorgu, N. M., Ezechi, N. G., & Idoko, E. C. (2007). Addressing nonscientific presuppositions in genetics using a conceptual change strategy. *Science Education*, 91(3), 419–438.
- Mioduser, D., Nachmias, R., Oren, A., & Lahav, O. (2000). Web-Based learning Environments (WBLE)—Current technological and pedagogical state. *Journal of Research on Computing in Education*, 33(1), 55–76.
- Oliva, J. M. (2003). The structural coherence of students' conceptions in mechanics and conceptual change. *International Journal of Science Education*, 25(5), 539–561.
- Park, J., & Han, S. (2002). Using deductive reasoning to promote the change of students' conceptions about force and motion. *International Journal of Science Education*, 24, 593–609.
- Prieto, T., Watson, R., & Dillon, J. S. (1992). Pupils' understanding of combustion. *Research in Science Education*, 22, 331–340.
- Rips, L. J. (1999). Human modes of quantificational reasoning. In S. B. Cooper & J. K. Truss (Eds.), *Models and computability* (pp. 353–365). Cambridge, England: Cambridge University Press.
- 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, 175–195.
- Schollum, B., & Happs, J. C. (1982). Learners' views about burning. *The Australian Science Teachers' Journal*, 28(3), 84–88.
- 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. (2003). DSLM instructional approach to conceptual change involving thermal expansion. *Research in Science and Technological Education*, 21(1), 43–54.
- She, H. C. (2004a). 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. (2004b). Fostering "Radical" conceptual change through Dual Situated Learning Model. *Journal of Research in Science Teaching*, 41(2).
- Tang, H. Y., She, H. C., & Lee, Y. M. (2005). The impact of DSLM instruction on middle school students' conceptual change involving mitosis and Meiosis. Paper Presented at the National Association for Research in Science Teaching 2005 World Conference, Dallas, Texas. April.
- Trumper, R., & Gorsky, P. (1993). Learning about energy: The influence of alternative frameworks, cognitive levels, and closed-mindedness. *Journal of Research in Science Teaching*, 30(7), 637–648.
- Tuvi, I., & Nachmias, R. (2001). Current state of web sites in science education—Focus on atomic structure. *Journal of Science Education and Technology*, 10, 293–303.
- Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, 51(4), 473–481.
- Venville, G. J., & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Teaching*, 35, 1031–1055.
- Wadsworth, B. J. (1996). *Piaget's theory of cognitive and affective development: White plains*. NY: Longman.
- Wason, P., & Johnson-Laird, P. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Watson, J. R., Prieto, T., & Dillon, J. (1995). The effect of practical work on students' understanding of combustion. *Journal of Research in Science Teaching*, 32(5), 487–502.