Computers in Human Behavior 37 (2014) 31–40

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full Length Article

A case study of Augmented Reality simulation system application in a chemistry course

Su Cai [a,b], Xu Wang [a,1], Feng-Kuang Chiang [a,b,*]

[a] School of Educational Technology, Faculty of Education, Beijing Normal University, No. 19, XinJieKouWai St., HaiDian District, Beijing 100875, China
[b] Joint Laboratory for Mobile Learning, Ministry of Education-China Mobile Communications Corporation, Beijing Normal University, No. 19, XinJieKouWai St., HaiDian District, Beijing 100875, China

ARTICLE INFO

Article history:
Available online 10 May 2014

Keywords:
Augmented Reality
Chemistry learning
Inquiry-based learning

ABSTRACT

The comprehension of micro-worlds has always been the focus and the challenge of chemistry learning. Junior high school students' imaginative abilities are not yet mature. As a result, they are not able to visualize microstructures correctly during the beginning stage of chemistry learning. This study targeted "the composition of substances" segment of junior high school chemistry classes and, furthermore, involved the design and development of a set of inquiry-based Augmented Reality learning tools. Students could control, combine and interact with a 3D model of micro-particles using markers and conduct a series of inquiry-based experiments. The AR tool was tested in practice at a junior high school in Shenzhen, China. Through data analysis and discussion, we conclude that (a) the AR tool has a significant supplemental learning effect as a computer-assisted learning tool; (b) the AR tool is more effective for low-achieving students than high-achieving ones; (c) students generally have positive attitudes toward this software; and (d) students' learning attitudes are positively correlated with their evaluation of the software.

1. Introduction

For many learners across the world, chemistry is introduced for the first time in junior high school. Abstract concepts such as molecules, atoms, and amount-of-substance are formidable to junior high school students; these students are often required to envision across micro- and macro-worlds, which can be extremely challenging. "The composition of substances" is a critical concept in chemistry learning, as it is the foundation of further learning about chemicals and organic chemistry. However, young students' imaginative abilities are limited, and it is difficult for them to imagine how particles such as atoms compose substances. This problem necessitates improvement in the learning methods and tools used in chemistry teaching.

Augmented Reality (AR) is an extension of Virtual Reality (VR). By contrast to traditional VR, AR provides a seamless interface for users that combines both the real world and the virtual world. Users can interact with virtual objects that are interposed on real scenes around them and obtain the most natural and genuine human–computer interaction experience. Only a computer and a camera are needed to construct a local AR environment. The camera detects markers within its vision and then presents the scene it captures and the corresponding virtual objects represented by the markers simultaneously on the computer screen. Users can move the markers to interact with the interposed virtual objects. In the three Horizon Reports released during 2010–2012, the New Media Consortium predicted that Augmented Reality will be applied on a large scale in the near future (Johnson, Adams, & Cummins, 2012; Johnson, Levine, Smith, & Stone, 2010; Johnson, Smith, Willis, Levine, & Haywood, 2011).

With the rapid development of Augmented Reality, the integration of AR into disciplinary teaching has emerged to a significant extent. AR is most applicable in the following two cases: (1) When the phenomenon cannot be simulated in reality, such as the solar system in "the book of the futures" (Cai, Wang, Gao, & Yu, 2012). (2) When real experiments have conspicuous shortcomings, such as the convex imaging experiment (Cai, Chiang, & Wang, 2013), as it is dangerous to keep a lighted candle in a classroom. Another example is a serious game for the treatment for a 25-year-old woman with cockroach phobia through a mobile phone (Botella et al., 2011). The use of the game reduced her level of fear and

* Corresponding author at: Joint Laboratory for Mobile Learning, Ministry of Education-China Mobile Communications Corporation, Beijing Normal University, No. 19, XinJieKouWai St., HaiDian District, Beijing 100875, China. Tel.: +86 1 58804632.

E-mail addresses: caisu@bnu.edu.cn (S. Cai), xu_wang@mail.harvard.edu (X. Wang), fkchiang@bnu.edu.cn (F.-K. Chiang).

[1] Harvard Graduate School of Education, Appian Way, Cambridge, MA 02138, United Kingdom.

http://dx.doi.org/10.1016/j.chb.2014.04.018
0747-5632/© 2014 The Authors. Published by Elsevier Ltd.

avoidance before a "one-session" AR exposure treatment was applied.

After a review of the related computer-assisted tools in chemistry education, we consider AR the most suitable and appropriate solution for the present problems we are faced with in instruction on chemistry micro-worlds, as micro-particles cannot be observed in reality.

This research aims to develop an inquiry-based AR learning tool for junior high school chemistry courses, examine its effect on students' cognitive performance, compare its effects on high-achieving and low-achieving students and investigate students' attitudes toward the software.

2. Literature review

Spatial ability plays an important role in chemistry learning, as students are required to visualize specific microstructures, but the visualization of microstructures is a difficult task for students. According to Harle and Towns (2011), research that has focused on visuospatial skills in chemistry has uncovered specific student difficulties in comprehending, interpreting, and translating molecular representations. The study of Tuckey, Selvaratnam, and Bradley (1991) indicated that even at the university level, many students have difficulties with three-dimensional thinking. These difficulties are caused by a misunderstanding of merely a few relatively simple concepts and skills. Sorby (2009) concluded that the implementation of a course aimed toward the development of the 3D spatial skills of first-year engineering students, it appears to have had a positive impact on student success, especially women. This result suggests that spatial skills can be improved through practice and may result in better academic performance. Based on these studies, we aim to eliminate the difficulties faced in chemistry microstructure instruction with regard to spatial skills.

A considerable number of computer-assisted learning tools are used in chemistry teaching, and a great number of researchers have designed specific scenarios using these tools and tested their learning effect on students. In recent years, the most highly praised of these tools for microstructure learning are Virtual Reality- and Augmented Reality-based learning tools.

Dalgarno, Bishop, and Adlong (2009) used a Virtual Laboratory to prepare new university chemistry students through distance learning. Most students found it to be a valuable preparatory tool and would recommend it for future use. These VR applications have been determined to be effective, whereas interactive VR methods are considered unnatural and limited. Merchant et al. (2012) examined the impact of 3D desktop Virtual Reality environments on learner characteristics using three Second Life simulations. The interactive features of these applications include the ability to interact with an object by zooming in and out, rotating the object and programming the object to behave in a certain manner. They found that the 3D virtual environment would promote student chemistry learning. In the work of Stull, Barrett, and Hegarty (2013), they examined the perceptual differences between using virtual and concrete model to learn organic chemistry. The learning task includes matching and comparing molecule structure and diagrams. It's discovered that there is no difference in the accuracy of task completion using two models, but virtual model provides a higher efficiency.

Compared with VR, AR demonstrates a more natural and innovative interactive concept, which provides students with opportunities to perform. El Sayed, Zayed, and Sharawy (2011) devised an Augmented Reality Student Card (ARSC), which can represent any lesson in a 3D format that aids students in visualizing different learning objects, interact with theories and manage information in a totally new way. The research suggests that ARSCs increase students' visualization abilities using a minimum number of tools. Nunez, Quiros, Nunez, Carda, and Camahort (2008) presented an AR system for teaching spatial relationships and chemical problems with university-level students. In the experiment, students could manipulate crystal structures of certain substances, such as $ZrSiO_4$ with markers. However, in the studies above, only static images or structures are rendered. Some more recent studies indicate more interesting and engaging interactions between students and the computer, taking full advantage of AR technology.

An Augmented Reality Teaching Platform (ARTP) in chemistry was proposed in Iordache, Pribeanu, and Balog (2012). A periodic table is provided where students could place colored balls to complete tasks. The researcher found the activity of placing colored balls onto different chemical elements on the table give the children the feeling of freedom and control, which is beneficial for their mastery. The results show that students understand more comprehensively and easily with this tool. Wojciechowski and Cellary (2013) constructed an AR environment in which students could conduct chemistry experiments, for example, hydrochloric acid (HCl) and sodium hydroxide (NaOH) react producing table salt (NaCl) and water. The results show that "The active participation of learners in hands-on activities has a particularly positive effect on the perceived enjoyment, resulting in their increased motivation for learning", as such seamless AR environments combine learning materials and the real scene around students, providing them with opportunities to manipulate the objects on their own.

In Mayer's multimedia learning theories research, he presented seven principles to involve animation in multimedia learning, the first principle is that students learn deeply from narration and animation than narration alone (Mayer & Moreno, 2002). Guided by this, in the AR application, we seamlessly incorporate interactive animation into the learning scenario.

Besides science disciplines, the AR environment also works well with art disciplines. An AR system for library instruction was developed, which resulted in significant learning performance improvement and was indeed helpful in promoting learner motivation and willingness to learn. "Obviously, learners were very satisfied with the proposed ARLIS for library instruction." (Chen & Tsai, 2012). In a visual art course, Di Serio, Ibáñez, and Kloos (2013) discussed the impact exerted by an AR system on students' motivation, which showed that AR has a positive impact on the motivation of middle school students. Fonseca et al. (2014) offered an opportunity to visualize different stages of a constructive process by AR on mobile devices, in order to improve the understanding of the process and to investigate the relationship among the usability of the tool, students' participation, academic performance after using AR. The results pointed out that the use of mobile devices in the classroom as well as motivation and academic achievement are highly correlated.

Our research targets "The composition of substances" segment of junior high school chemistry syllabus, which requires instruction in microstructures. Considering the difficulty that may exist in the teaching of these abstract materials and the important role that spatial ability plays in molecular geometry learning, we choose to develop an AR-based learning tool. Traditional 2D pictures and textbooks place great cognitive loads on students. Using AR to learn, students can observe a molecule or crystal model from each angle. Furthermore, Piaget (1972) said that "knowledge originates from activities and recognition starts from practice". With prevalent chemistry learning software, students can only observe structures instead of interacting with them. In the proposed AR environment, students can control particles in micro-worlds with markers, construct molecules and substances with these particles and, furthermore, comprehend and conclude the process of substance composition.

3. Research methods

3.1. Research subject

This study involved 29 students in Grade 2 including 16 boys and 13 girls. The experiment of the software's impact was conducted in a junior high school in Shenzhen, China.

3.2. Research rationale and experiment preparation

This study mainly focuses on the supplemental learning effect of AR-based learning tools in a chemistry course. The testing class was taught the content of "The composition of substances". We interviewed the chemistry teacher before the design and development of this AR tool, when she pointed out that her previous experiences showed that students were not very motivated and did not completely comprehend the materials in this chapter, as the content is dull and abstract. Therefore, she expressed a wish to review the content using an AR tool in order to stimulate learning interests and promote learning outcome of "The composition of substances" content. For such reasons, the experiment did not include a control group. Pre-test scores will represent students' learning outcomes when using textbooks, and post-test scores will represent students' learning outcomes after using an AR inquiry-based learning tool. None of the tools used in the activity, including the software, markers and activity form, presented the exact knowledge points included on the test, which means that students' test answers had to be conclusions that they reached by themselves while observing and exploring during the inquiry-based learning process. Additionally, in this case, we believe that the vertical difference between pre-test and post-test scores will represent the AR tool's learning effect. The questionnaire primarily surveys students' learning attitudes toward this AR learning tool.

Before the experiment, researchers installed the AR software on each computer in the classroom. The experiment contains five sections, as shown in Table 1.

3.3. Research question and hypothesis

We proposed three research hypotheses to be further tested and examined by the experiment.

Hypothesis 1. We expect there to be a statistically significant improvement in students' scores on the quiz after using this AR tool.

Table 1
Experiment design.

Experiment content	Operation methods
Pre-test	Each student is required to complete a paper and pencil test independently
Grouping and inquiry-based leaning using activity form	The class is randomly divided into groups of three. Each group is required to use the AR tool to learn as indicated on the activity form and cooperatively complete the form without the teacher's guidance (a set of AR tools contains AR-based software, markers and the activity form)
Post-test	Each student is required to independently complete the same test used for the pre-test
Questionnaire survey	Each student is required to independently complete a paper and pencil questionnaire
Interview	Five students are chosen randomly and interviewed about their feelings during the inquiry-based learning process

Hypothesis 2. As this is a review class of the content, we wish to examine the tool's remedial influence on students. Thus, we want to compare the influence of the AR tool between high-achieving and low-achieving students. And we expect to discover a more significant improvement in low-achieving students.

Hypothesis 3. Informed by other studies concerning the application of AR in education, we found several studies mentioning the improvement of students' learning motivation instead of significant improvement in grades. We're interested to explore whether this AR tool has similar effect on students' learning motivation towards chemistry, and we expect students to find the experience meaningful and interesting.

3.4. Measurement instruments

3.4.1. Pre–post test

The quiz was devised by a junior high school chemistry teacher and further examined by a group of chemistry education experts, including two junior high school chemistry teachers and three professors specializing in science education. The paper and pencil test includes 32 effective blanks, associated to the learning content "the composition of substances" and has a full mark of 32 points (1 blank 1 point).

Before the experiment, a pre-test was conducted with the whole class—"Quiz of micro particles". Afterwards, the researchers randomly divided the class into 10 groups, 2–3 students in a group. Each group then used the AR tool proposed in this paper to complete an inquiry-based learning task. After learning with AR, a post-test was conducted, using the same quiz—"Quiz of micro particles".

The quiz examines students' understanding and memorizing of several key knowledge points in this chapter. The quiz takes the form of "fill in the blanks". The questions are briefly summarized as below.

General concept questions include, what are the three particles that could form materials? What does a molecule consist of? What does an atom consist of? And determine whether the given materials are elemental or compounded.

Questions regarding the water molecule case include, how many extranuclear electrons does a hydrogen (and oxygen) atom contain, how are they distributed on each electron sheath? What does water (and CO_2) as a material consist of? What is a water (and CO_2) molecule composed of?

Questions regarding the diamond and graphite case include, what does diamond (and graphite) consist of?

Questions regarding the salt case include, what is the chemical representation of salt? What does it consist of?

3.4.2. Instruction and activity form

In this study, we designed an inquiry-based group learning scenario in which students were required to conduct explorations in groups of three without teacher instruction. They must use the AR tool and conclude the principles by themselves. The activity form is designed to assist them in the learning process, providing operation steps showing them what to do and asking introductory questions encouraging them to think and draw conclusions. The activity form was devised by us according to our application design. We want to clarify that the design of this activity form is not going to limit students' operations with the markers. As it is the first time students get knowledge of AR technology, and we put them in a self-exploration scenario without teachers' guidance, the activity form aims to show them how to use markers to interact with the computer and observe properly. They are also encouraged to operate markers freely to see what they could discover.

3.4.3. Post-questionnaire

The questionnaire adopts Likert scale with six options: a scale of 1, or "Strongly Disagree", to 6, or "Strongly Agree". The questionnaire consists of four constructs: learning attitude, satisfaction with the software, cognitive validity and cognitive accessibility. The "learning attitude" construct includes seven items, which are revisions of items taken from Hwang and Chang (2011). The "satisfaction with the software" construct is covered by 14 items, which were revised and appended from Chu, Hwang, & Tsai, 2010. The "cognitive validity" construct is covered by five items, and the "cognitive accessibility" construct is covered by four items, all of which are revisions of items taken from Chu, Hwang, Tsai, and Tseng (2010).

29 copies of the questionnaire survey were distributed and 29 were received, all of which were considered valid. We conducted a reliability analysis of each construct and the entire questionnaire; the results are shown in Table 2. The Cronbach's Alpha coefficient for the questionnaire is 0.974, suggesting that the questionnaire is reliable. As the Cronbach's Alpha coefficient for each construct is above 0.70, we consider each construct to have high inner consistency and reliability.

3.4.4. Interview protocol

By doing the interview, we aim to further explore students' learning experiences through AR tools. During the interview, we asked the students the following questions.

Do you think this AR tool facilitates your chemistry learning?
Why do you think this tool is helpful? In what areas does the tool help you?
Do you wish to use AR software to learn chemistry in the future? Why?
How does this AR software compare with other learning tools you have used, such as flash applications and 3D learning software, for which you use a mouse to interact with the computer? In what areas do you think AR is better than those tools?
Do you think that AR software has any disadvantages? What are they?
Can you offer some advice for improving this AR learning tool?

3.5. AR tool introduction

The set of AR Tools we developed during this study contains AR software, six markers and an activity form. The software contains four specific substance composition applications: (1) hydrogen atoms and oxygen atoms compose water molecules, and water molecules compose water; (2) carbon atoms compose diamond crystals; (3) carbon atoms compose graphite crystals; and (4) chloridion and sodion compose NaCl.

The interaction tool used with this software is the marker. A set contains six markers printed with the numbers 1–6, which are selectively applicable to different applications. After the software is installed, students can use different markers to control microparticles and conduct inquiry-based learning as instructed on the activity form and further generalize concepts and conclusions.

Table 2
Cronbach's alpha for each construct.

Variable	Number of items	Cronbach's alpha
Learning attitude	7	0.822
Satisfaction	14	0.963
Cognitive validity	5	0.965
Cognitive accessibility	4	0.911
Overall	30	0.974

3.5.1. Instructional design of the experiment

By contrast to previous studies of AR, we adopted a student-centered scenario, in which students are divided into groups and learn in an AR environment on their own. They use markers to represent particles and construct molecules and substances as if conducting real experiments in a laboratory. We expect students to explore and reach conclusions through group effort without benefiting from the teacher's direct instruction. In this case, the activity form shows operational steps, and present questions corresponding to the required knowledge points, which encourage them to think and investigate.

We present a detailed description of the activity form content of activity 1, the case "hydrogen atoms and oxygen atoms compose water molecules, and water molecules compose water". In this case markers "2" "3" "4" represent three different particles.

Activity 1 takes 8–10 min to complete. The activity form instructs operation procedure of this activity and presents questions on what students observed in each step.

With the 3D models presented from different markers, we expect students to master the structure of the Hydrogen atom, the Oxygen atom and the water molecule, which they could hardly visualize with traditional 2D materials. With this hands-on activity, we expect students to understand the procedural knowledge that 2 Hydrogen atoms and 1 Oxygen atom compose a water molecule, and water molecules compose water. We also positively predict that the process can be better memorized with students' visualized operation.

3.5.2. Application design and development

The software is programmed in Java, and the extra packages used include NyARToolkit, Java3D and JMF (Java Media Framework). In addition to accurate modeling, the essence of human–computer interaction with this software is to detect and record the position of each marker in the camera's view, as the application will trigger different animations when the marker is at different positions. That is, the interaction between users and computer is position-based. In other words, we use position of markers to present different phase of a structure and various combinations of atoms. The markers' behavior can be consistent with real particle behaviors in some cases, while inconsistent in other cases. For example, when two markers get closer, a new molecule can be formulated, which is what really happens in micro-world. In another example, when lifting a marker, the molecule changes from molecular structure into substantial form, or specifically from H_2O molecular structure into a water drop. The behavior "lifting a/an molecule/atom" does not really happen in micro-world, whereas with these special behaviors and operations, we expect students to acknowledge the transformation between atoms, molecules and substances through tangible movement. The following figure shows operation screens from two applications, the water and the diamond cases.

When students move marker "2" within the camera's view, they will observe the model of the hydrogen atom, as shown in Fig. 1, and when lifting the marker, they will observe that the electron is revolving irregularly around the nucleus, as shown in Fig. 2.

As shown in Fig. 3, three atoms, including 2 hydrogen atoms and 1 oxygen atom are interposed in the scene. When we move the two hydrogen atoms close to the oxygen atom, a water molecule is formed, as shown in Fig. 4. Users are allowed to lift the water molecule closer to the camera to view its structures, and if we keep lifting, it turns into a water drop, as shown in Figs. 5 and 6.

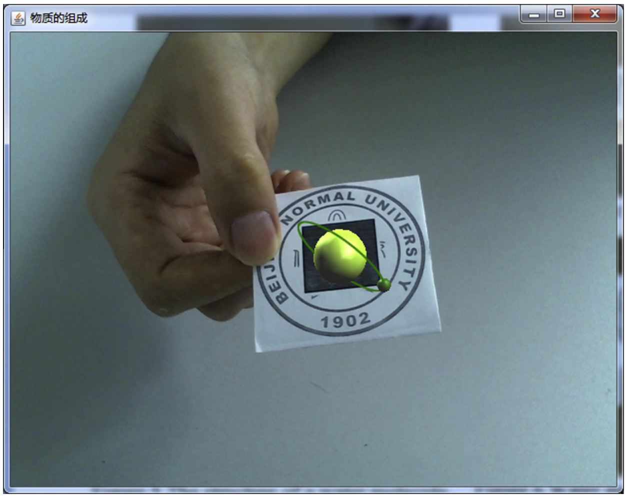

Fig. 1. Hydrogen atom model.

Fig. 2. The electron revolves around the nucleus.

Fig. 3. Models of three atoms.

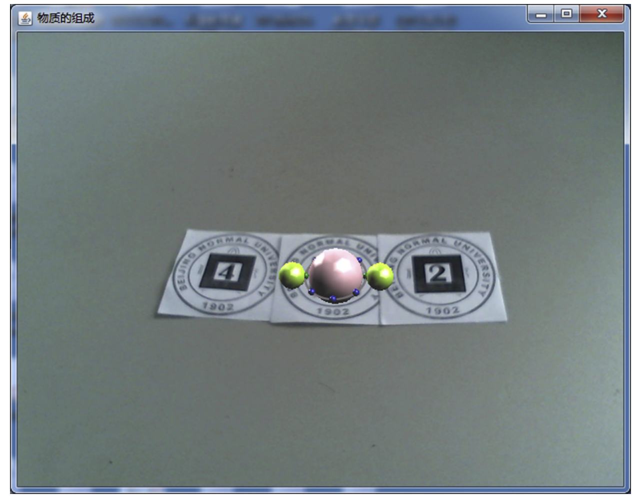

Fig. 4. Three atoms form a water molecule.

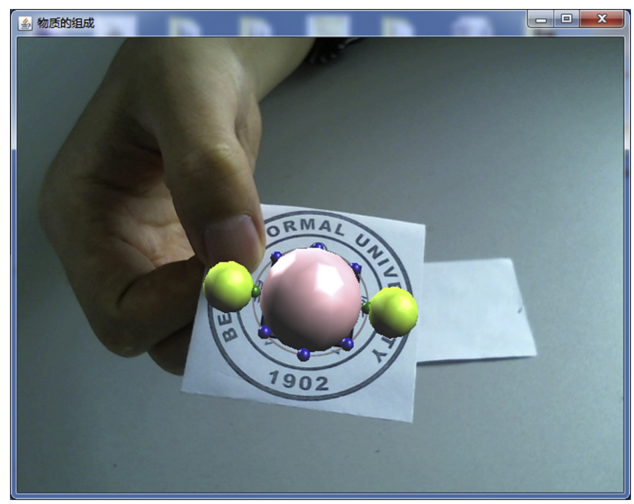

Fig. 5. The structure of a water molecule.

Fig. 6. Water molecules form a real water drop.

In the case of the diamond, users are allowed to construct a diamond crystal element using carbon atoms and bonds. A basic tetrahedron unit is constructed, which is elemental to the diamond crystal. Further, we use this tetrahedron unit to extend and construct a bigger diamond crystal element. When we put another marker within the scene, a piece of diamond is interposed, allowing students to visualize the material formed by this chemical structure.

After students finish the inquiry-based activity, researchers expect them to (1) know that there are three particles that can compose substances, explain the formulation of water, graphite, diamond and NaCl, know the structure of atoms of different elements, and connect the features of substances with microstructures; (from the cognitive perspective) (2) be able to generalize abstract concepts and master basic chemistry research methods; (from the behavioral perspective) (3) form the habit of respecting objective facts and having a serious attitude toward science, inspiring interest in learning chemistry (from the emotional perspective).

3.6. Data analysis methods

In this experiment, we adopted quantitative methods to explore the change in students' scores, whereas we adopted qualitative methods at the same time to delve deeper into their feelings and experiences throughout the process.

3.6.1. Quantitative research methods

We used quantitative methods to analyze the data obtained from the quiz test and the questionnaire. We conducted a paired t-test on the pre-test and post-test scores and an independent t-test on the scores of high-achieving and low-achieving students to determine the differences between them. We calculated descriptive statistics for each item on the questionnaire and each construct as a whole, including the average score, standard deviation, and maximum and minimum values. Furthermore, we calculated a Pearson correlation coefficient between learning attitude and the other three constructs.

3.6.2. Qualitative research methods

We video-recorded the entire 40-min class and took pictures during the process. In addition, the researchers kept literal records of exceptional student performance during the class. After the class, we interviewed five students and recorded the interviews. We analyzed the video, pictures and notes acquired from the case study. Several experiment scenes are displayed in Fig. 7.

4. Data analysis and findings

4.1. Overall cognitive performance

The experiment produced 29 * 2 test copies (29 for the pre-test and 29 for the post-test), all of which are considered effective. The full mark of the test is 32 points. We conducted a paired t-test for the pre-test and post-test score variables. The tested variable is post-test score minus pre-test score, which stands for the difference yielded after using the AR tool for each student. The results are shown in Table 3.

Table 3 shows that the p-value (two-tailed) of the mean is close to zero ($t = 4.332$, p-value = 0.000). When the significance level is 0.05, we should reject the null hypothesis, which suggests that students' scores after using the AR inquiry-based learning tool are significantly higher than those attained before the learning activity. As a result, we conclude that with other unobserved variables controlled, the AR inquiry-based learning tool has a statistically significant improvement on the score of the adopted cognitive quiz test, and students' averages scores increased by 3.310 points.

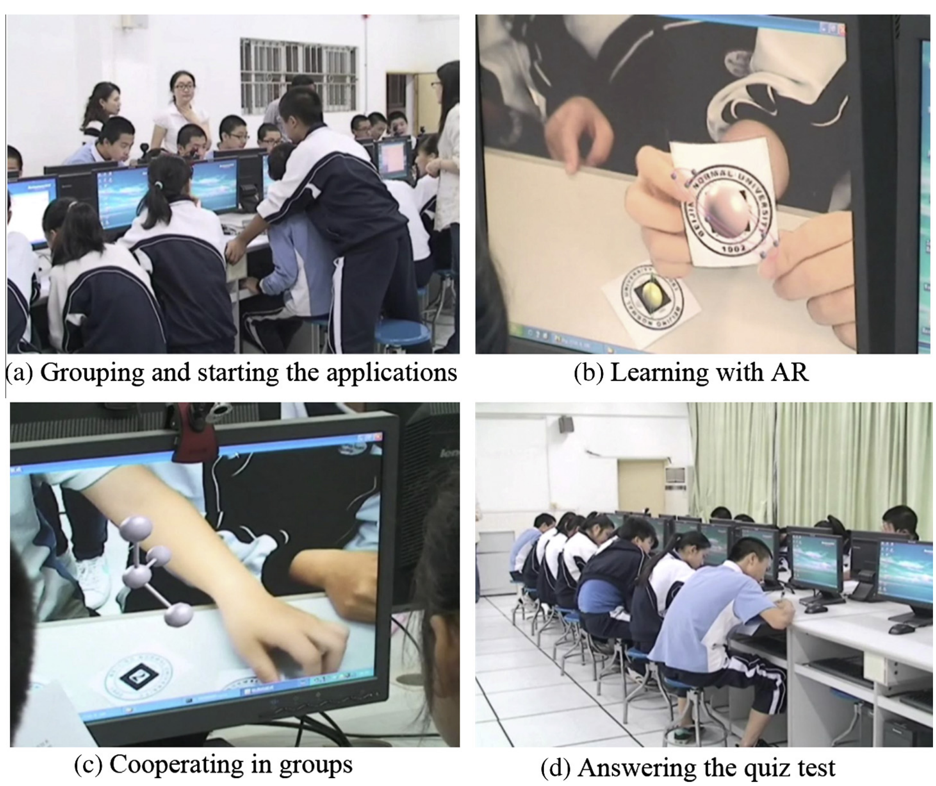

Fig. 7. Experiment scenes.

Table 3
Paired t-test for pre-test and post-test score variables.

	Paired differences					t	df	Sig (2-tailed)
	Mean	Std. deviation	Std. error mean	95% Confidence interval of the difference				
				Lower	Upper			
Post-test score –pre-test score	3.310	4.115	.764	1.745	4.876	4.332	28	.000

Table 4
Average scores for low-achieving and high-achieving groups.

Group	Pre-test average	Post-test average	Gain average
High-achieving	27	28.6	1.6
Low-achieving	12.7	18.7	6

4.2. Comparison of the learning gains of high-achieving and low-achieving students

Placing the students' pre-test scores in a high-to-low order, we categorized the first 33% as high-achieving students and the last 33% as low-achieving students. We calculated the average learning gains of both groups, as shown in Table 4. Then, we conducted an independent t-test with the difference between high-achieving students' learning gains and low-achieving students' learning gains, as shown in Table 5.

For the Levene's test for the equality of variances shown above, $F = 2.652$, $p = 0.120 > 0.05$, which suggests that we cannot reject the null hypothesis and should accept that the variance difference is not significant at the 0.05 significance level; we should consult the "Equal variances assumed" row. In the upper row, $t = -2.302$ and $p = 0.033 < 0.05$, suggesting that the difference of the mean learning gains between the low-achieving group and high-achieving group is significant at the 0.05 level. In the two Tables above, we can observe that the average learning gains of low-achieving students is 4.218 points higher than that of high-achieving students.

4.3. Attitudes toward the AR tool through questionnaire analysis

In the questionnaire analysis, we calculated the score of each construct by averaging all of the corresponding items within each construct. The descriptive statistics obtained are shown in Table 6.

In the table above, we can observe that the "Learning Attitude" construct has the highest mean value, which suggests that the students tested generally have positive learning attitudes. By contrast, the "Cognitive Accessibility" construct has the lowest mean value, which suggests that the usability of this AR-based learning tool is not as satisfactory as that of the other three constructs and must be further improved. Furthermore, we displayed the descriptive statistics of each item in the four constructs, respectively in the following tables. Descriptive statistics for items relating to the "Learning Attitude" construct are shown in Table 7.

Table 6
Descriptive statistics for the four questionnaire constructs.

Variable	Sample size	Max	Min	Mean	Std. deviation
Learning attitude	29	3	6	4.9655	0.90565
Satisfaction	29	1	6	4.4138	1.45202
Cognitive validity	29	1	6	4.4138	1.57020
Cognitive accessibility	29	1	6	4.2759	1.53289

Table 7
Descriptive statistics for "Learning Attitude" construct.

Item	Mean	Std. deviation
I view learning about "the composition of substances" as rewarding	5.24	1.023
I think that learning chemistry is rewarding	5.55	0.736
I think that learning things related to chemistry is meaningful	5.34	0.814
I think that learning and observing chemistry-related content in addition to that in textbooks is meaningful	5.34	0.974
I will actively search for information related to chemistry in books or on the internet	3.28	1.907
When I come across problems in learning chemistry, I will actively reach out to teachers, classmates, books or the internet for solutions	4.07	1.689
I think that learning chemistry is important for everyone	5.03	1.149

The statement "I think that learning chemistry is rewarding" has the highest value (Mean = 5.55), which is very close to the full mark of 6, suggesting that most students consider it important to learn chemistry. Although the item "I will actively search for information related to chemistry in books or on the internet" has the lowest score (Mean = 3.28), which means that although they find learning chemistry important, most students only learn chemistry in class and do not actively search for information.

Descriptive statistics for items related to the "Satisfaction" construct are shown in Table 8.

In this table, the item "this game-like learning tool can aid me in discovering new questions" has the highest score, 4.79, which suggests that students enjoy the "game-like" learning tool and that this AR-tool introduces new questions for them to solve. Although the statement "I will recommend the AR learning tool to other classmates" has the lowest value, 3.90, suggesting that although the students may find the AR tool interesting, there still exist problems discouraging them from recommending the software to other classmates.

Table 5
Independent t-test.

	Levene's test for equality of variances		t-test for equality of means						
	F	Sig	t	df	Sig (2-tailed)	Mean difference	Std. error difference	95% Confidence interval of the difference	
								Lower	Upper
Equal variances assumed	2.652	.120	−2.302	19	.033	−4.218	1.832	−8.053	−.384
Equal variances not assumed			−2.383	13.908	.032	−4.218	1.770	−8.018	−.419

Table 8
Descriptive statistics for the "Satisfaction" construct.

Item	Mean	Std. deviation
AR-based learning software is more interesting than previously used learning methods	4.59	1.701
This game-like learning tool can aid me in discovering new questions	4.79	1.634
Using AR-based software enables me to view chemistry concepts such as the micro-world, atoms, and molecules in a different way	4.59	1.701
I like learning chemistry using AR	4.07	1.811
I like game-like learning methods	4.55	1.804
I hope that other disciplines such as physics and biology will apply AR tools to learning as well	4.55	1.594
I hope to use similar AR tools to learn chemistry in the future if possible	4.17	1.983
I will recommend the AR learning tool to other classmates	3.90	1.877
I'm interested in using AR-based learning tools	4.59	1.524
The content of this software is closely related to the "the composition of substances", which is a very interesting topic to me	4.21	1.780
The AR-based learning tool enables me to learn not only on own but also with my friends and classmates	4.72	1.386
The design of this software is pleasing and genuine	4.31	1.713
The color of this software is appropriate, as it is attractive and does not distract me	4.34	1.610
I think that learning about the composition of substances and micro-worlds using an AR-based learning tool is necessary	4.66	1.471

Table 9
Descriptive statistics for the "Cognitive Validity" construct.

Item	Mean	Std. deviation
I believe that AR demonstration renders learning materials more detailed and understandable	4.38	1.699
I think that using this game-like AR learning tool is very helpful for learning chemistry	4.41	1.722
This AR learning tool is more effective than any other software I have ever used	4.34	1.696
Using this AR software enables me to master important knowledge points in an in-depth manner and comprehend the principles I did not understand in the past	4.28	1.688
The AR learning tool provides abundant space for me to think and try, which aids me in solving problems	4.38	1.720

Table 10
Descriptive statistics for the "Cognitive Accessibility" construct.

Item	Mean	Std. deviation
Operating AR software is not difficult	4.10	1.760
Learning to use this AR tool does not cost me a great deal of time and energy	3.76	1.845
The content of and procedures for this learning activity are clear and understandable to me	4.34	1.738
I can grasp how to operate AR software within a very short timeframe	4.69	1.561

Descriptive statistics for items related to the "Cognitive Validity" construct are shown in Table 9.

In this table, the mean scores for the five items are very close, suggesting that most students judge the AR tool to be helpful to their learning.

Table 11
Pearson Correlations for "Learning Attitude", "Satisfaction", "Cognitive Validity", and "Cognitive Accessibility".

	Learning attitude	Satisfaction	Cognitive validity	Cognitive accessibility
Learning attitude		0.554	0.513	0.573
		0.002	0.004	0.001
Satisfaction	0.554			
	0.002			
Cognitive validity	0.513			
	0.004			
Cognitive accessibility	0.573			
	0.001			

Descriptive statistics for the items related to the "Cognitive Accessibility" construct are shown in Table 10.

In this table, the item "I can grasp how to operate AR software within a very short timeframe" has the highest score, 4.69, whereas the item "Learning to use this AR tool does not cost me a great deal of time and energy" has the lowest score, 3.76. These results suggest that although students think that they can master how to operate the software very quickly, it still requires energy and time. Thus, to improve the efficiency of the AR tool, we need to lessen the cognitive load that the software itself places on students.

The above tables show that students generally have a positive learning attitude and provide positive evaluations of the software.

Next, the correlation between students' learning attitudes and their evaluation of the AR tool was analyzed. Table 11 shows that the Pearson correlation coefficient between "Learning Attitude" and "Satisfaction" is 0.554, $p = 0.002 < 0.05$. The Pearson Correlation coefficient between "Learning Attitude" and "Cognitive Validity" is 0.513, $p = 0.004 < 0.05$. The Pearson Correlation coefficient between "Learning Attitude" and "Cognitive Accessibility" is 0.573, $p = 0.001 < 0.05$. These coefficients suggest that learning attitude has a significant positive correlation with students' satisfaction with the AR tool and students' evaluation of the AR tool's cognitive validity and accessibility. In other words, the more important a student thinks learning chemistry is, the more useful and satisfactory he will believe the AR tool is, which is consistent with our expectations.

4.4. Students' experience exploration through observations and interviews

Throughout the entire experiment, researchers made careful observations and kept records of students' performance. Most of the students appeared excited, curious and motivated during the inquiry-based learning activity. The first two groups to finish the activity were comprised of all boys. At first, two girls did not participate in the learning activity, instead completing homework on the other side of the classroom; after the teacher's encouragement, they joined the experiment. Based on this observation, we also did t-test with difference in average learning gains between boys and girls and expect boys to yield more average leaning gains than girls. But the result shows that although boys' average learning gains is higher than that of girls, the difference is not statistically significant.

We found that most students do not like to consult the paper activity form; on the contrary, they like to interact with the AR software on their own. The responses to the activity form denoted conspicuous mistakes, but these can be avoided with careful observation and proper teacher guidance. According to this observation, we've further reflected on the usability of this software and decided to abandon long and detailed papery instruction in the future; instead, try to incorporate video or audio instructions and improve the learnability of this tool.

After the experiment, we randomly picked five students, including four boys (S1–S4) and one girl (S5), to be tested using interviews and communication. During the interview, we asked them five questions according to the interview protocol and encouraged them to discuss their feelings regarding the learning tool. We presented the questions and some of their responses as follows,

> *Questions 1, 2 and 3: Do you think this AR tool facilitates your chemistry learning? Why do you think this tool is helpful? In what areas does the tool help you?*
> *The AR tool can help me remember the structure of atoms. Chemistry is relatively difficult, and sometimes, we are not able to imagine the structures correctly in class with merely the teacher's simple instruction. The software is more attractive, which leaves a deeper impression in our mind. The more interesting the material is, the longer it will be remembered. (S1)*
> *I used to think that learning is always planar, but after using AR software, I found that learning in real space can be really exciting. (S2)*
> *AR software makes learning materials more clear and understandable, which helps me remember knowledge points more directly. (S5)*

When the 5 students were asked whether they would like to use the AR tool for future study, they all said "yes". S3 said that the AR tool is much more interesting than traditional learning materials and that he would like to use it in the future.

> *Question 4: How does this AR software compare with other learning tools you have used, such as flash applications and 3D learning software, for which you use a mouse to interact with the computer? In what areas do you think AR is better than those tools?*
> *Compared with previous flash courseware and other 3D modeling software, the AR tool could help us develop operation capabilities. The natural and direct interaction is better than keyboard and mouse interaction for remembering procedural knowledge especially. (S1)*
> *Question 5 and 6: Do you think that the AR software has any disadvantages? What are they? Can you offer some advice for improving this AR learning tool?*
> *The model can be instable, and the display can flicker at times. (S4)*
> *I hope the simulation of substances can be more real. (S1)*
> *I suggest that the software add some cartoon or animation elements to be more fascinating. (S2)*

According to this short interview, interviewed students all presented a positive attitude towards the AR tool and expressed their willingness to use the software again in future learning process. They mentioned several advantages of the AR tool as we expected, including the material is interesting, clear, understandable, the experience is spatial instead of planar, it helps to visualize, etc. While when asked to compare the AR tool with other learning tools, one student proposed that it's more helpful to remember procedural knowledge, this is also consistent with our expectation that AR can be used to enable students to accomplish procedural experiments which they may not be able to conduct otherwise. In the end, we asked students to talk about the disadvantages of the AR tool, most of them mentioned that the model can be instable at times. And one student expected the simulations to be more real, another suggested adding some cartoon or animation elements. These suggestions provided by students are precious and invaluable data for us to further revise and complete our product.

5. Discussions and conclusions

5.1. Preliminary findings and discussions

Based on the data analysis of learning effect, students' attitudes, and the observations and interviews, we drew the following preliminary conclusions.

First, the AR inquiry-based learning tool has significant supplemental learning effects. The empirical study tested the software's supplemental effect. The pre-test scores represent students' learning outcomes when using textbooks. Although students learned the topic in class within a week, memory decay is inevitable. As a result, the pre-test score should be lower than students' learning outcomes when using textbooks. We did not arrange control group experiments, and we must admit that review of the same content using other tools or materials may also result in score increases.

Second, the AR tool results in more significant learning gains for low-achieving students than for high-achieving students. The AR tool is more effective for low-achieving students, and we analyzed the possible reasons for this disparity. (1) The original scores of high-achieving students are very high, some even close to full marks. The space for improvement is quite limited. (2) The test was relatively basic and was already mastered by high-achieving students at the start point. (3) The AR tool aims to aid students in exploring and generalizing concepts; its effect may not be entirely demonstrated on a paper and pencil test. In addition, representing learning gains with the differences in paper and pencil test scores can be biased, especially when the main goal of this AR tool is to help students develop problem-solving skills, inquiry-based exploration skills instead of merely cognitive memorization. In future experiments, open-ended questions targeted at students' inquiry process with AR might better represent the outcome of the experiment. Through further analysis of the test, we found that students' attitudes toward the test are not consistent. Some students provided correct answers on the pre-test but made careless mistakes on the post-test. Additionally, the AR tool provides a new cognitive method and is expected to have a long-term memory effect on students through their inquiry-based observation and operation. However, the post-test was conducted immediately after the learning activity, so the time effect of the AR tool cannot be measured.

Several groups made conspicuous mistakes in following the activity form because of carelessness. On the one hand, students are required to explore by themselves without the teacher's guidance, thus incorrect procedures and lack of feedback might account for these mistakes. On the other hand, we noticed during the experiment that most students did not like to consult the operation procedures provided on the paper activity form, perhaps because young children prefer images over texts. When letters and pictures are presented at the same time, students may focus on the computer screen and ignore the activity form. From the perspective of usability of the AR tool, providing long and detailed instructions for middle school students may not be appropriate. In further design, we aim to increase the learnability and flexibility of this AR tool, enabling students to master it without much learning and in a shortest time.

Third, in general, students possess a positive learning attitude and provide positive evaluations of the AR tool, as is suggested by questionnaire statistics. Among the four constructs, "Cognitive Accessibility" received a relatively lower score, while still within the positive category, which echoes the usability issue we discussed above. The results found in student attitudes toward the AR tool is consistent with the results of Nunez et al. (2008), in which AR is considered as a powerful tool that helped the students surveyed comprehend 3D structures. In Iordache et al. (2012), it is also indicated that ARTP is a useful teaching aid able to increase the students' motivation to learn.

Further, there exists a significant positive correlation between students' learning attitudes and their evaluation of the AR tool. This result indicates that the AR tool is like any other learning tool in that the learning gains it may produce are based on students' belief that learning the discipline is important. If researchers want to promote this AR tool, the primary task is to promote students' learning initiative and allow them to genuinely believe that learning chemistry is important and rewarding. With both the interview and the questionnaire, we found that students view the stability of the AR tool as unsatisfactory. The main reason is that the AR software must detect the real scene with the camera, which can be constricted by lighting conditions. There is thus more space for improvement to the software.

5.2. Conclusions

The experiment result shows that the AR tool is beneficial in improving middle school students' cognitive test performance on corresponding content, and has relatively larger influence on low-achieving students. Additionally, students generally hold a positive attitude toward the AR tool and enjoyed the exploration experience.

Based on the findings, we wish to further employ this AR tool as a remedial learning tool and extend the method to other chapters and contents in middle school chemistry course that require students to memorize abstract chemical structures and concepts. With the application and instruction form, teachers can apply this AR tool in inquiry-based learning in their own classes.

5.3. Possible improvements and future work

Through this empirical study, we have witnessed the great potential and acceptance of the inquiry-based AR tool. The research shows that in an inquiry-based chemistry microstructure learning scenario, students' academic performance will be enhanced by the AR learning tool. We also proposed several aspects where the research can be further improved and continued. We wish to observe how this AR tool compares with other learning software and whether the AR learning tool will result in longer-lasting memory. The following are our plans for our future work.

In terms of software design,

- Promote the cognitive accessibility of the software by making it more stable and easier to use.
- Present vocal operation procedures as part of the software rather than text on an activity form.
- Survey target students to determine their preferences. Add elements that might attract them to the software, such as cartoons and animation.
- In terms of experiment design.
- Add a control group, which will learn the same content using textbooks, and analyze the difference between this group and the experiment group using the AR tool.
- During the inquiry-based learning activity, incorporate teacher guidance and enable students to explore and learn using the teacher's instruction and feedback.
- Increase the difficulty level of the pre-test and post-test. Change the structure of the test, avoid having all blanks assess cognitive knowledge points, and add problem-solving topic questions.
- Conduct two post-tests. Launch the first post-test immediately after use of the AR tool and the other a week after use. Attempt to measure the long-time learning effect of the AR tool.

References

Botella, C., Breton-López, J., Quero, S., Baños, R. M., García-Palacios, A., Zaragoza, I., et al. (2011). Treating cockroach phobia using a serious game on a mobile phone and augmented reality exposure: A single case study. *Computers in Human Behavior, 27*(1), 217–227.

Cai, S., Chiang, F. K., & Wang, X. (2013). Using the augmented reality 3D technique for a convex imaging experiment in a physics course. *International Journal of Engineering Education, 29*(4), 856–865.

Cai, S., Wang, X., Gao, M., & Yu, S. (2012). *Simulation teaching in 3D augmented reality environment. In: 1st IIAI international conference on advanced applied informatics*. Fukuoka, Japan: IEEE Computer, Society, pp. 83–88.

Chen, C., & Tsai, Y. (2012). Interactive augmented reality system for enhancing library instruction in elementary schools. *Computers & Education, 59*, 638–652.

Chu, H. C., Hwang, G. J., & Tsai, C. C. (2010a). A knowledge engineering approach to developing mindtools for context-aware ubiquitous learning. *Computers & Education, 54*(1), 289–297.

Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, Judy C. R. (2010b). A two tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers & Education, 55*(4), 1618–1627.

Dalgarno, B., Bishop, A. G., & Adlong, W. (2009). Effectiveness of a virtual laboratory as a preparatory resource for distance education chemistry students. *Computers & Education, 53*, 853–865.

Di Serio, Á., Ibáñez, M. B., & Kloos, C. D. (2013). Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education, 68*, 586–596.

El Sayed, N. A. M., Zayed, H. H., & Sharawy, M. I. (2011). ARSC: Augmented reality student card an augmented reality solution for the education field. *Computers & Education, 56*, 1045–1061.

Fonseca, D., Martí, N., Redondo, E., Navarro, I., & Sánchez, A. (2014). Relationship between student profile, tool use, participation, and academic performance with the use of augmented reality technology for visualized architecture models. *Computers in Human Behavior, 31*, 434–445.

Harle, M., & Towns, M. (2011). A review of spatial ability literature, its connection to chemistry, and implications for instruction. *Journal of Chemical Education, 88*(3), 351–360.

Hwang, G. J., & Chang, H. F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*(1), 1023–1031.

Iordache, D. D., Pribeanu, C., & Balog, A. (2012). Influence of specific AR capabilities on the learning effectiveness and efficiency. *Studies in Informatics and Control, 21*(3), 233–240.

Johnson, L., Adams, S., & Cummins, M. (2012). *Augmented reality time-to-adoption horizon: Four to five years. NMC Horizon Report: 2012 K-12 Edition*. Austin, Texas: The New Media Consortium, pp. 28–31.

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). *Simple augmented reality time-to-adoption horizon: Two to three years. The 2010 Horizon Report*. Austin, Texas: The New Media Consortium, pp. 21–24.

Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). *Augmented reality time-to-adoption horizon: Two to three years. The 2011 Horizon Report*. Austin, Texas: The New Media Consortium, pp. 16–19.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*(1), 87–99.

Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O., Cifuentes, L., & Davis, T. J. (2012). The learner characteristics, features of desktop 3D virtual reality environments, and college chemistry instruction: A structural equation modeling. *Computers & Education, 59*, 551–568.

Nunez, M., Quiros, R., Nunez, I., Carda, J. B., & Camahort, E. (2008). Collaborative augmented reality for inorganic chemistry education. *New Aspects of Engineering Education*, 271–277.

Piaget, J. (1972). *To understand is to invent the future of education*. Grossman Publishers.

Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. *International Journal of Science Education, 31*(3), 459–480.

Stull, A. T., Barrett, T., & Hegarty, M. (2013). Usability of concrete and virtual models in chemistry instruction. *Computers in Human Behavior, 29*(6), 2546–2556.

Tuckey, H., Selvaratnam, M., & Bradley, J. (1991). Identification and rectification of student difficulties concerning three-dimensional structures, rotation, and reflection. *Journal of Chemical Education, 68*(6), 460–464.

Wojciechowski, R., & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Computers & Education, 68*, 570–585.