VOLUME 23, NUMBER 2

APRIL 2014

23(2) 207–298 (2014)

ISSN 1059-0145

Journal of Science Education and Technology

Special Issue: International Conference on Science Education (Nanjing, China, October 2012)

Guest Editors: Xiufeng Liu · BaoHui Zhang

Springer

Journal of Science Education and Technology

Aims and Scope

Journal of Science Education and Technology is an interdisciplinary forum for the publication of original peer-reviewed, contributed and invited articles to improve and enhance science education at all levels worldwide. Topics covered can be categorized as *disciplinary* (biology, chemistry, physics, mathematics, computer science and engineering and the learning processes related to their acquisition and assessment of results), *technological* (computer, video, audio and print), and *organizational* (legislation, administration, implementation and teacher enhancement). Insofar as technology is playing an increasing role both in the understanding and the development of science disciplines and in the delivery of information, the journal includes it as a component of science education. The journal provides a stimulating and informative variety of papers geared toward theory and practice in the hope that common information shared among a broad coalition of individuals and groups involved in science education will facilitate future efforts. In addition to works in the fields mentioned above and case studies of exemplary implementations, the journal publishes reviews of books, media, software and relevant products to help reach our common goal: excellence in science education.

Journal of Science Education and Technology uses a double blind review process. We reject articles that are specific to one population or location, that are not easily understandable in English, and/or do not address our main focus.

Please read our Review Criteria for further information.

J Sci Educ Technol (2014) 23:207–210
DOI 10.1007/s10956-013-9485-1

Editorial: Special Issue (SI): International Conference on Science Education (ICSE)

Xiufeng Liu · BaoHui Zhang

Published online: 8 January 2014

Abstract In the context of science education globalization, the International Conference on Science Education was held in Nanjing, China, in October 2012. The purpose of this conference was to provide a forum for science education researchers from China and from the rest of the world to exchange research ideas and best practices in science education. A call for papers for a special issue of the Journal of Science Education and Technology was made to all conference participants, and a set of six articles was resulted from a standard peer review process. This set of six articles provides a snapshot of research in China and in some other countries, and represents a dialogue between Chinese science education researchers and science education researchers from other countries. We call for more exchange and collaboration in science education between China and the rest of the world.

Keywords Science education · Globalization · Comparative education · Chinese education · China

According to the just-released PISA 2012 results, students in Shanghai, China once again, after 2009, achieved the highest average score, the highest percentage of level 5 or above, and the lowest percentage of below level 2 among all participating countries/regions in all three subjects, mathematics, science, and reading. Of course, it is naïve to consider this outstanding performance by students in Shanghai to be representative of students in the country, as Shanghai is among the most economically and educationally developed regions in China, and the disparity between east and west and urban centers and rural areas is significant (Wang et al. 2012). How have students in Shanghai been able to perform so well on such international tests as PISA? How do students in other parts of China compare? Those are the questions educators, policy makers, and even the general public around the world are now asking. However, answers to those questions are not easy to find because scholarly publications about Chinese science education in English are few.

The paucity of English literature on Chinese science education is due to many reasons. Language barrier is an obvious one. Also, difference in research traditions between Chinese science education and Western science education is significant. Jenkins (2001, 2004) defines two research traditions in science education: the empirical tradition and the pedagogical tradition. The empirical tradition is characterized by a primary focus on developing and testing general science education theories, while the pedagogical tradition is characterized by a primary focus on improvement of science curriculum and instruction in specific disciplines. Science education research in Western countries follows primarily the empirical tradition, while Chinese science education research follows primarily the pedagogical tradition. The mismatch between the two research traditions poses a great challenge for Chinese science education researchers to publish in English science education research journals, and it is difficult for Chinese science education researchers even to present at major international science education conferences such as the annual meetings of the NARST.

X. Liu (✉)
University at Buffalo, State University of New York, Buffalo, NY, USA
e-mail: xliu5@buffalo.edu

B. Zhang
Nanjing University, Nanjing, China

The limited interactions between Chinese science education researchers and that in the rest of the world are in contrast to the rapid integration of Chinese economy into the world economy, particularly over the past decade. Since 2010 when China officially joined the World Trade Organization (WTO), Chinese economy has now become the world's second largest economy only after that of the USA. China has now also a large foreign reserve and is the USA's largest foreign creditor. With its rapid economic growth over a long period and increasing financial capacity, China has been making unprecedented investment in scientific research and development as well as in science, technology, engineering, and mathematics (STEM) education. According to the *Science and Engineering Indicators* (National Science Board 2012), a biannual factual and policy neutral compilation of data related to science and engineering in the USA and internationally, in 2008, five million first university degrees (i.e., bachelor's degrees) were awarded in S&E worldwide, 23 % of which were earned by Chinese students, compared to 19 % earned by Europe Union students and 10 % earned by American students. Furthermore, in 2008, while S&E degrees accounted for about one-third of all bachelor's degrees in the USA, in China this number was more than half. In 2007, China overtook the USA as the world leader in the number of doctoral degrees awarded in natural sciences and engineering. At the same time, international student mobility has been increasing. From 1989–2009, the USA had the largest number of foreign students worldwide, with over 60 % of them studying in S&E fields. Among those foreign students studying S&E in the USA, about one in five doctoral degrees was earned by Chinese students. Chinese students now represent a large source of S&E talents in both China and internationally. S&E talents have become global.

Chinese science education can no longer stay in isolation from the rest of the world, and the rest of the world can no longer ignore what is happening in Chinese science education as the world economy is becoming more and more integrated. In an effort to promote dialogue in science education research between China and the rest of the world, in October 2012, the *International Journal of Science Education* published a special issue entitled "Science Education Research in China: Challenges and Promises" (Liu et al. 2012). Consistent with the above effort, the International Conference on Science Education 2012 (ICSE 2012) was held from October 12–15 at Nanjing University, Nanjing, China (http://edu.nju.edu.cn/zbh/icse2012). ICSE 2012 was co-organized by the National Association for Science Education, a branch of the Chinese Society of Education (CNASE) and the Institute of Education of Nanjing University. ICSE 2012 intended to be a forum for science education researchers from around the world to exchange experiences, challenges, and strategies in science education research around a common theme of "Science Education: Policies and Social Responsibilities." ICSE 2012 was the first large international conference organized by the Chinese National Association for Science Education since it was founded in 2009. The conference international organizing committee was composed of noted science education researchers from 22 countries over the five continents. There were 122 representatives from 15 countries attending the conference. They came from China mainland, Taiwan, Macau (China), USA, UK, Australia, Russia, Germany, Japan, Singapore, Malaysia, Korea, Iran, Pakistan, and Nigeria. There were also more than fifty graduate students attending the conference. The conference program included 12 invited plenary presentations—six invited talks from overseas and six invited talks from China—followed by 55 concurrent presentations and 14 posters. Four participated online. There were 45 papers in Chinese and 33 papers in English (Zhang et al. 2013).

This special issue, *International Conference of Science Education 2012*, is one of the products of the above conference. Under an agreement with the Springer and Editor-in-Chief of the *Journal of Science Education and Technology*, Dr. Karen Cohen, we announced a call for papers for this special issue to all participants of the conference. All submitted manuscripts went through the standard peer review process established by the journal. After a little over one year's review, revision and re-review, six articles have been finally accepted for inclusion into the special issue.

"Chemistry Teachers' Knowledge and Application of Models" by Zuhao Wang, Shaohui Chi, Kaiyan Hu, and Wenting Chen reports a study on Chinese chemistry teachers' knowledge and application of models. They found that chemistry teachers' knowledge of some known chemistry models was limited and their modeling process was incomplete. Teachers followed a general pattern when they used models in chemistry teaching. The findings have implications for pre-service and inservice teacher education.

"Students' Perceptions of Their Science Teachers' Pedagogical Content Knowledge" by Lilia Halim reports a study on science teachers' pedagogical content knowledge from students' perspective. A questionnaire was used to collect data from 316 Form 4 (16 years old) students. One-way ANOVA analysis revealed that the differences in science teachers' PCK identified by students of different achieving abilities were statistically significant. Overall, students of various academic achieving abilities considered all the components of PCK as important. The low-achieving students viewed all the components of PCK as being less important compared to the high and moderate achievers. In particular, low-achieving students did not view "Knowledge of Concept Representation" as important for effective teaching. They valued the fact that teachers should be alert to their needs, such as being sensitive to students' reactions and preparing additional learning materials. This study has revealed that PCK of science teachers should be different for high- and low-achieving students and knowledge of students' understanding played a critical role in shaping teachers' PCK.

"Secondary Students' Stable and Unstable Optics Conceptions Using Contextualised Questions" by Hye-Eun Chu and David F. Treagust focuses on the stability of and interrelationships between students' conceptions about Light Propagation and Visibility of Objects. Using the Light Propagation Diagnostic Instrument, they surveyed 1,233 Korean and 1,149 Singapore students across 3 years of secondary schooling from years 7 to 9. They found that only about 10–45 % of students could apply their conceptions of basic optics in contextualized problem situations giving rise to both stable and unstable alternative conceptions. Students' understanding of Light Propagation concepts compared to Visibility of Objects concepts was more stable in different problem situations. The concepts of Light Propagation and Visibility of Objects were only moderately correlated. School grade was not a strong predictive variable, but students' school achievement correlated strongly with their conceptual understanding in optics. Possible influence of the teaching and learning approach and education systems in the two countries was discussed.

"On the Evolution of a Lesson: Group Preparation for Teaching Contest as Teacher Professional Development Activity for Chinese Elementary Science Teachers" by Xiaowei Tang and Faxian Shao reports a study on an inservice science teacher professional development through group lesson preparation for teaching contest. Through participant observation and discourse analysis, they examined how a science lesson evolved through lesson-polishing process and how such process influenced individual learning and the development of local teaching community. Although lesson-polishing activity opened up space for critical yet cooperative professional interactions and tryouts of different designs and teaching strategies, thus opportunities for individual learning and development of practical rationalities within local community, such activities were greatly limited by the tendency of refining every detail in lesson design, the existence of overriding dispositions and authorities with overriding power, as well as the focus on practical suggestions that could be directly implemented.

"Development of an Instrument for Assessing the Effectiveness of Chemistry Classroom Teaching" by Changlong Zheng and Peng He reports a study to measure the effectiveness of chemistry lessons. Using focus group interviews, the study investigated the variables on the effectiveness of lessons. They found a total of 21 such variables that were related to five main factors: rational use of time (RUT), quality of teaching behavior chain (QTBC), match degree (MD), quality of using resource & technology (QUR&T), and rationality of primitive content (RPrC). Based on these findings, they constructed a scale for measuring the effectiveness of chemistry lessons.

"Enactment of Scientific Inquiry: A Case Study in China Mainland" by Lei Wang, Ronghui Zhang, and David Clarke reports a collective case study on how two Chinese science teachers implemented inquiry science teaching in their classrooms. Based on analyses of pre-instructional and post-instructional interviews, classroom observations, as well as student and teacher lesson artifacts, they found that both teachers implemented a range of inquiry activities. Differences in the implemented inquiry activities between the two teachers were also noticeable. Factors influencing teachers' implementation of inquiry science teaching were discussed.

Four of the above articles were written by researchers from China, one from Malaysia, and one from Singapore with data from Korea and Singapore. We do not claim that this set of articles represents the science education research status in those countries, nor do they represent all the papers presented at the ICSE 2012. Instead, they provide a snapshot of research in those countries and represent a dialogue between Chinese science education researchers and science education researchers from other countries. Through this set of articles, we show our commitment to promoting interactions in science education research between China and the rest of the world. We call for more exchange and collaboration in science education between China and the rest of the world.

In concluding this editorial, we would like to thank Springer for sharing our vision for science education globalization. We particularly thank Dr. Karen Cohen, Editor-in-Chief of the *Journal of Science Education and Technology*, for her support to publish this special issue. We also thank the reviewers from around the world for their time and expertise in reviewing the submissions and revisions. Enjoy reading the articles in this special issue!

References

Jenkins WE (2001) Research in science education in Europe: Retrospect and prospect. In: Behrend H et al (eds) Research in science education: Past, present, and future. Kluwer Academic Publishers, Hingham, pp 17–25

Jenkins WE (2004) Science education: Research, practice and policy. In: Scanlon E et al (eds) Reconsidering science learning. Routledge/Falmer, London

Liu X, Liang LL, Liu E (2012) (eds.). Science education research in China: Challenges and promises. A special issue of the Int J Sci Edu 34(13), 1961–2088

National Science Board (2012) Science and Engineering Indicators 2012. National Science Foundation, Arlington VA

Wang S, Liu X, Zhao Y (2012) Opportunities to learn in school and at home: how can they predict students' understanding of basic science concepts and principles. Int J Sci Edu 34(13):2061–2088

Zhang BH, Zhang H, Peng S (2013). The development and changes of science education in a global age: an overview of the International Conference on Science Education 2012. Glob Edu (4): 120–128. [In Chinese]

J Sci Educ Technol (2014) 23:211–226
DOI 10.1007/s10956-013-9455-7

Chemistry Teachers' Knowledge and Application of Models

Zuhao Wang · Shaohui Chi · Kaiyan Hu · Wenting Chen

Published online: 12 June 2013

Abstract Teachers' knowledge and application of model play an important role in students' development of modeling ability and scientific literacy. In this study, we investigated Chinese chemistry teachers' knowledge and application of models. Data were collected through test questionnaire and analyzed quantitatively and qualitatively. The result indicated as follows: (1) Chemistry teachers' knowledge of some known chemistry models was limited; (2) Chemistry teachers preferred those models that were vivid when they chose models; (3) Teachers' modeling process was incomplete; (4) Teachers adopted a general pattern when applying models in chemistry teaching. The findings have implications for teacher education.

Keywords Chemistry Teachers · Model · Knowledge · Application

Introduction

Model is a simplified representation of a system or phenomenon that focuses attention on specific aspects or components of a system (e.g., prototype), such as ideas, objects, events or processes (Gilbert et al. 1998; Ingham and Gilbert 1991). Because models play an important role in the formation and the justification of scientific knowledge, which are the core components of scientific theories, science can be viewed as a complex and dynamic network of models (Koponen 2007; Pluta et al. 2011; Schwarz et al. 2009; Windschitl et al. 2008).

Models take a central role in science education, fulfilling a series of purposes, such as making abstract entities visible (Francoeur 1997), and deriving hypotheses from the model (Van Driel and Verloop 1999b). In some countries, for example, USA [American Association for the Advancement of Science (AAAS) 1989 and 1993; National Research Council (NRC), 1996] required students to be knowledgeable in varied aspects of scientific inquiry and the nature of science, including the role of models and modeling (Crawford and Cullin 2005).

Explaining and modeling are just two of the many facets of chemistry (Talanquer 2011). The process of developing models modeling is central to scientists' daily practice; thinking and reasoning with models enables scientists to visualize the abstract processes and entities they are investigating, to provide explanations for them and to make predictions about them (Gilbert et al. 2000), which also allow scientists to represent their current understanding of a system under study (Jungck and Calley 1985), and to communicate ideas to others (NRC 2012).

Modeling also provides opportunities for students to learn about science inquiry (Crawford and Cullin 2005; Schwarz and White 1998; Wisnudel-Spitulnik et al. 1999). Specifically, teaching models are created to support the learning of some abstract topics, especially concepts related to bonding and structure (Kozma and Russell 2005). What's more, engaging students in modeling leads to more sophisticated understanding of key models in science, as well as helps them understand the nature of disciplinary knowledge (Schwarz et al. 2009).

Over the past decades, many scholars have paid attention to model and modeling in science education, which

Z. Wang (✉)
Institute of Curriculum and Instruction, East China Normal University, Shanghai, China
e-mail: zhwang@chem.ecnu.edu.cn

S. Chi · K. Hu · W. Chen
Department of Chemistry, East China Normal University, Shanghai, China

has led to deep thinking and various relevant researches (e.g., Erduran and Duschl 2004; Klein 2003; Suckling et al. 1978). They have found that students face a number of difficulties with models for science learning including immature understanding of models and limited experience of applying and creating models (Halloum 1996; Schwarz et al. 2009; Treagust 2002; Wei 2011).

Consequently, educators have made efforts to improve students' understanding of models and modeling practice (Schwarz et al. 2009; Treagust 2002; Yang 2012). Recently, there has been a call for model-based teaching and learning (MBTL) (Buckley et al. 2004; NRC 2012; Wei 2011). Model-based teaching is any implementation that brings together information resource, learning activities, and instructional strategies intended to change student conceptions and improve student scientific understanding (Zhang et al. 2006). Therefore, questions arise: How good is teachers' knowledge of model? How do teachers apply model in classroom?

Everyone would expect that teachers' understanding and application of models has a close relationship with students' understanding of model and mastery of modeling skill. On the other hand, teachers' knowledge and ideas determine their teaching methods. In recent years, an extensive body of research concerning teacher's knowledge has accumulated in the field of teacher education (Fang 2003); however, little has been known about teachers' knowledge and application of model in science education (Van Driel and Verloop 1999a). Therefore, it is necessary for us to explore teachers' knowledge and application of models, to be more specific, their understanding of known models, their teaching practice of models in classroom including model selection, construction and application. This research may provide feasible suggestions for improvement of MBTL and thus contribute to students' improvement of scientific knowledge and modeling ability. In this research, we investigated chemistry teachers' knowledge and application of models. We intended to answer the following questions:

(1) What is chemistry teachers' knowledge of known model?
(2) How do chemistry teachers select model?
(3) How do chemistry teachers construct model?
(4) How do chemistry teachers use model in teaching?

Theoretical Framework

Models and Modeling in Science Education

In science education, models and modeling can help learners build subject matter expertise, epistemological understanding, and practices and skills such as systems thinking (Lehrer and Schauble 2000; Lesh and Doerr 2003; Schwarz and White 2005), which also can help them understand some abstract scientific conceptions, theories, and phenomena. Modeling not only can help students learn to demonstrate important thinking strategies (Stratford 1996), but also make them to learn science subject matter (Harrison and Treagust 1996).

Modeling-based learning is the approach of using modeling during learning in science, which can provide the context in which the construction and refinement of models can achieve better conceptual and operational understanding of the nature of science (Bell 1995; Grosslight et al. 1991; Harrison and Treagust 1998; Louca et al. 2011; Schwarz 2009; Sins et al. 2009; Windschitl et al. 2008). In recent years, models and the process of modeling have been indicated as core components of scientific endeavors (Gilbert 1991; Linn 2003).

Teachers' understanding and application of model has a close relationship with students' learning achievement (Duit and Glynn 1996). Modeling activities can also provide especially valuable opportunities for teachers to monitor students' progress from their initial mental models to an understanding of established scientific or historical models (Justi and van Driel 2005). Dori and Barak (2001) investigated the effect that teaching organic chemistry using virtual and physical models had on students' understanding of both new concepts and the spatial structure of new molecules. They found that experimental students who worked with two kinds of models gained better understanding of the model concept and they were more capable of defining and implementing new concepts and were able to transfer between the chemistry understanding levels: symbolic, macroscopic, and submicroscopic.

Therefore, teachers really need an adequate grasp of subject matter and the purpose and nature of scientific models in order to teach their own students (Justi and Gilbert 2001). So far, there have been a lot of researches concerning students' learning of models; however, the number of investigations into teachers' knowledge of models and how they apply models in classroom is limited (Van Driel and Verloop 1999b). Nonetheless, as recognition of the role of models and modeling in science education is rather recent, major research studies on this theme have only been published in the last two decades (Justi and van Driel 2005).

In order to answer our research questions and consider the characteristics of chemistry education in China, we referenced the few relevant research studies and build the theoretical framework.

Teachers' Understanding of Models

Since model is descriptive, explanatory and predictive, i.e., making the abstract concrete, it is very important for teachers to have a good understanding of model and modeling (Hodson 1993). It is found that, based on the literature regarding teachers' understanding of models, scholars generally consider two aspects. One is teachers' view of models in terms of its nature and function. How teachers view models in terms of its nature and function in science education has a great influence on their model selection, instructional strategy and the effect of model-based instruction (De Jong and Van Driel 2001). Only when becoming aware of the function of models, can teachers take advantage of models and improve students' understanding.

The other one is teachers' knowledge of known model. As scholars, such as Harrison (2000a, b) and Smith and Finegold (1995) agreed, teachers' knowledge of known model, to a certain degree, could determine how well they apply models to improve students' learning achievement. Therefore, it is necessary to refer to research regarding the two aspects to guide us explore teachers' understanding and then their application of models. (Crawford and Cullin 2005; Van Driel and Verloop 1999a).

A large number of researches have explored how teacher view model's nature and function. Van Driel and Verloop (1999b) investigated experienced science teachers' ideas of models. Their findings indicated that teachers could give a general elaboration, that is, model was simplified representation of entity; however, their ideas were diverse and limited. According to Harrison (2000), only 2 of 25 in-service teachers he interviewed expressed that models could be used as thinking tools. Harrison (2001) also reported that some teachers never thought about the nature of models and paid little attention to models in the classroom. On the other hand, teachers tended to directly provide models instead of encouraging students to construct models by themselves, which suggested they had relatively narrow opinion about the nature and function of model.

Other scholars also attained similar findings, De Jong and Van Driel (2001)found that preservice science teachers' idea of models needed to be improved. Justi and Gilbert (2002a) interviewed some teachers and found they lacked awareness of the scope and limitations of models in the presentation of models to students; teachers might acknowledge the usefulness of models as pedagogical tools for teaching information about scientific content rather than see models as tools within a scientific process that could help learners understand the nature of science (Crawford and Cullin 2004; Henze et al. 2007; Justi and Gilbert 2002b) or as thinking tools that could advance students' model-based reasoning (Harrison and Treagust 2000; Henze et al. 2007). Most teachers were found to fail to realize the importance of model in teaching and learning, and they have limited experience and knowledge about the epistemological richness of the pedagogy, such as scientific modeling (Van Driel and Verloop 1999b, 2002) or modeling-centered inquiry (Windschitl and Thompson 2006).

Justi and Gilbert (2002a) investigated 39 in-service teachers and preservice teachers and found the majority viewed models as a kind of instructional strategies or tools. Furthermore, teachers' idea of the function of models in teaching was categorized as follows: (1) Models make science more interesting; (2) models make explanation clearer; (3) models make the abstract concrete, help students reach a better understanding of complex phenomena at the molecular level; and (4) models can improve students' conceptual change, especially improve their self-constructing models.

As for teachers' knowledge of models, the research is quite few. Smit and Finegold (1995) found teachers failed to understand the nature of models fully and their knowledge of models needed to be improved. Van Driel and Verloop (2002) explored how good 74 science teachers' knowledge of familiar models and ability of modeling. They found that some teachers failed to fully understand models and had difficulties integrating their own knowledge into instruction.

To sum up, teachers' understanding of models, including their view of model's nature and function, and knowledge of known models, is found to be limited and should be improved, which may have an influence on teachers' model selection, application, and then the outcome of model-based instruction. For one thing, teachers tend to hold superficial conception of model's nature and function, resulting from their poor knowledge of models, lack of experience of modeling, or disinclination to reflect upon models in science education, and consequently have difficulty selecting the appropriate model, constructing scientific model, let alone create opportunity for students to benefit from model-centered inquire, and fail to make full use of model and modeling in classroom. For another thing, teachers' own knowledge is far from perfect; as a result, their efforts to guide students understand the abstract better and involve students in modeling will be poor. Therefore, teachers' understanding of models should be investigated and efforts and changes needed to be made to improve this situation.

Unfortunately, teachers may acknowledge the usefulness of models as pedagogical tools for teaching information about scientific content rather than see models as tools within a scientific process that can help learners understand the nature of science (Justi and Gilbert 2002a; Crawford and Cullin 2004; Henze et al. 2007) or as thinking tools that can advance students' model-based reasoning (Harrison and Treagust 2000; Henze et al. 2007).

However, most teachers were found to fail to realize the importance of model in teaching and learning, and they have limited experience and knowledge about the epistemological richness of the pedagogy (Justi and Gilbert 2002b) such as scientific modeling (van Driel and Verloop 1999a, 2002) or modeling-centered inquiry(Windschitl and Thompson 2006).

Smit and Finegold (1995) found teachers failed to understand the nature of model fully and their knowledge of model needed to be improved. Van Driel and Verloop (1999b) investigated how experienced science teachers view models, and their findings indicated that teachers could give a general elaboration, that is, model was simplified representation of entity. However, their ideas were diverse and limited. According to Harrison (2000), only 2 of 25 in-service teachers he interviewed expressed that models could be used as thinking tools. Harrison (2001) also reported that some teachers never thought about the nature of model and paid little attention to models in classroom. On the other hand, teachers tended to directly provide models instead of encouraging students to construct models by themselves, which suggested they had poor knowledge of the nature and function of model.

Other scholars also attained similar findings, De Jong and Van Driel (2001) found that preservice science teachers' idea of model was narrow. Justi and Gilbert (2002b) interviewed some teachers and found they were lack of the conscious of considering the scope and limitations of models in the presentation of models to students.

Justi and Gilbert (2002a) investigated 39 in-service teachers and preservice teachers and found the majority viewed model as a kind of instructional strategy or tool. Furthermore, teachers' idea of the function of model in teaching was categorized: (1) models make science more interesting; (2) models make explanation clearer; (3) models make the abstract concrete and help students reach a better understanding of complex phenomena at the molecular level; (4) models can improve students' conceptual change, especially improve their self-constructing models.

Van Driel and Verloop (2002) explored how good 74 science teachers' knowledge of familiar model and ability of modeling. They found that some teachers failed to fully understand models and had difficulties integrating their own knowledge into instruction.

Teachers' Application of Models

Model is a useful tool for teachers to improve students' understanding, and model-based instruction has been proved to be effective. An extensive body of research with respect to model-instruction has been accumulated. The majority of researches usually focused on the effect of a certain kind of model-based instruction. In research, the steps of model-based instruction are clearly explained as a teaching routine, and teachers teach according to the routine, and the effect generally referring to students' academic achievement was tested, that is, teacher's application of models in teaching is set beforehand. The role of teachers is weakened in those researches, and we can hardly see teachers' individuality in terms of how they apply models for teaching.

A kind of well-developed model-based instruction is truly important, and the role of teacher is not less important that it directly determines its ultimate success. Application of models in classroom involves a series of activities and practices, which have to be integrated and designed well. For example, using models has to be suitable, which requires teachers to have a clear understanding of the teaching purpose, the nature and function of models, and so on. Moreover, the instructional strategy is very important. How teacher select models, apply models and other teaching-related problems should be solved to improve the model-based instruction. This research considered how teachers select models and apply models in teaching and aimed at exploring intrinsic factors of model-based instruction.

However, quite a few researches focus on teachers' application of models in classroom. Among these few researches, most of them are general. For example, recent studies have revealed that teachers, both experienced and beginning, fail to have a good understanding of model and they meet various difficulties applying models in teaching (Van Driel and Verloop 1999a; Harrison 2001; Crawford and Cullin 2002; Justi and Gilbert 2002a, b, 2003). The detailed information, such as, what difficulties, is unknown; the deep analysis, such as, the characteristics and category of teachers' difficulties, is also unknown.

Other researches usually aim at improvement of model-based instruction from the perspective of teachers. Glynn (1991) put forward the Teaching-with-Analogies Model, and it includes 6 steps: (1) introduce the target concept; (2) remind students of what they know of the analogy concept; (3) identify relevant features of the concept and analogy; (4) connect the relevant features; (5) indicate where the analogy breaks down; (6) draw a conclusion about the concept. Research indicated that the teaching model was effective, for example, it improved students' conceptual change and helped them reach a good understanding of the advantages and disadvantages of models. Harrison and Treagust (2000b, 2001) recommended that teachers should teach modeling skill, encourage students to use multiple analogical models rather than isolated models, and take time to discuss and critique them, since modeling ability, unlike content, can only be learned through intensive practice. Justi and Gilbert (2002a) put forward 5 pieces of advice to enhance model-based instruction: (1) To have a clear understanding of the nature of model, including what model is, the characteristics of model, and so on. (2) To know when, why and how to apply models in classroom. (3) To develop model-specific instructional strategy to improve students' understanding of model. (4) To encourage students to construct model by themselves. (5) To know how students build their mental model and how to deal with various models.

In fact, almost none of research has investigated specifically how teacher apply models in classroom, for example, what factors teachers give priority to when select models, their application pattern, which are the main questions this research attempted to answer. Even though there has been hardly directly relevant research, these mentioned research, such as Glynn's 6 steps model, could provide some reference when we analyzed teachers' descriptions of their model application and categorized their application pattern.

Some researches with respect to teachers' application of model for teaching are available. Glynn (1991) put forward the Teaching-with-Analogies Model, and it includes 6 steps: (1) introduce the target concept; (2) remind students of what they know of the analog concept; (3) identify relevant features of the concept and analogy: (4) connect the relevant features; (5) indicate where the analogy breaks down; (6) draw a conclusion about the concept. Research indicated that the teaching model was effective, for example, it improves students' conceptual change and helps them reach a good understanding of the advantages and disadvantages of models.

Recent studies have revealed that teachers, both experienced and beginning, fail to have a good understanding of model and they meet various difficulties applying models in teaching (van Driel and Verloop 1999a; Harrison 2001; Crawford and Cullin 2002; Justi and Gilbert 2002a, b, 2003).

Actually, some scholars have already paid attention to improve this issue. Harrison and Treagust (2000b, Harrison and Treagust 2001) recommended that teachers should teach modeling skill, encourage students to use multiple analogical models rather than isolated models, and take the time to discuss and critique them, since modeling ability, unlike content, can only be learned through intensive practice.

Justi and Gilbert (2002a) put forward 5 pieces of advice to enhance model-based instruction: (1) To have a clear understanding of the nature of model, including what model is, the characteristics of model, and so on. (2) To know when, why and how to apply models in classroom. (3) To develop model-specific instructional strategy to improve students' understanding of model. (4) To encourage students to construct model by themselves. (5) To know how students build their mental model and how to deal with various models.

Model Construction

To construct model is, to a certain degree, a process of problem solving, which starts from posing questions, and goes through solving the problem (constructing the model). Scholars have developed various ideas concerning pattern of model construction.

Justi and Gilbert (2002a) put forward a model of modeling framework (Fig. 1).

Model of modeling involves six phrases: (1) decide on the purpose of modeling,

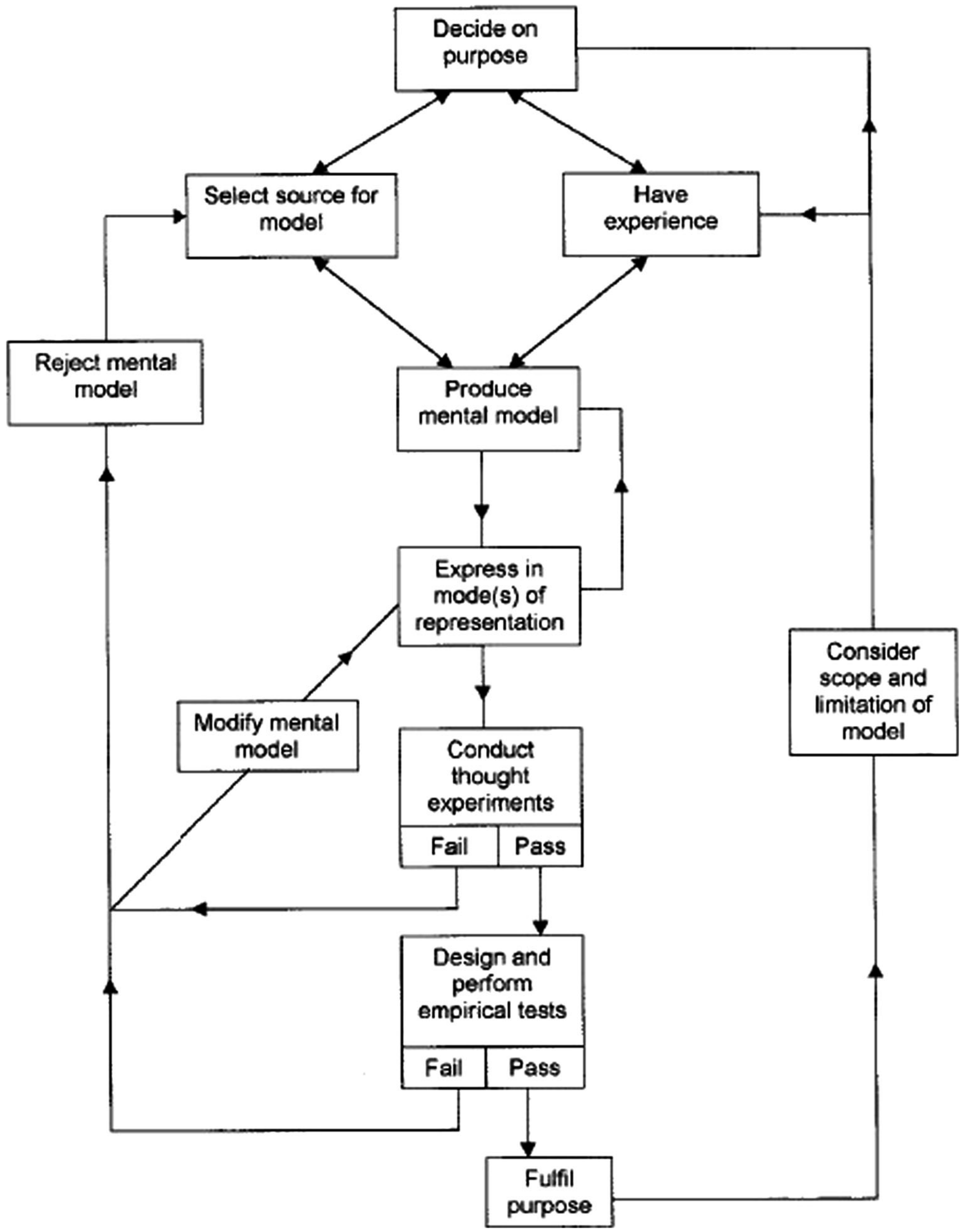
whether it be to describe the behavior of a phenomenon, to establish the entities of which it is thought to consist, to ascribe the reason for the causes and effects of—that behavior, to predict how it will behave under other circumstances, or several or all of these; (2) make observation of phenomenon modeled, select relevant source and obtain some initial, direct or indirect, qualitative or quantitative, experience to form a mental model; (3) express the model by an appropriate mode of representation: material, visual, verbal, mathematical; (4) conduct thought experimentation in mind, if the model fails to produce predictions that are confirmed in the thought experimental test phrase, and then an attempt will have to be made to modify it and to reenter the cycle, if it passes the test phrase, it can go on to the next phrase; (5) design and perform empirical test, which includes design and conduct practical work, followed by the collection and analysis of data, and finally by the evaluation of the results against the model, if the model fails at this phrase, an attempt also has to be made to modify it and reenter the cycle, if it passes the test, the purpose for which it was constructed for has been fulfilled;(6) communicate with others about the model, not only should its value be persuaded, the scope and limitation also should be elaborated, which leads to a reconsideration of the earlier elements in the modeling cycle. On the other hand, if the sub-cycle of model modification and thought and/or empirical test is repeatedly unsuccessful, then the model will have to be rejected.

Lesh and Lehrer (2003) developed a modeling cycle (Fig. 2). It can be seen, the process of modeling involves three elements, that is, purpose, underlying conceptual systems, and media in which the conceptual system is expressed. Models are purposeful description or explanations. Their purpose often involves constructing, manipulating, or predicting the system modeled; and the process of developing scientific models usually involves a series of iterative test and revision cycle.

From the above two patterns, model construction follows a common basic pattern: (1) to decide on the purpose of model construction; (2) to collect data, and apply different approaches to construct a tentative model; (3) to test the model, if it passes the test, complete the model, otherwise, revise or reject it.

During this process, there is interaction among three factors, e.g., purpose, conjecture, and test. Generally, people analyze the question comprehensively and make a plan, and then test the plan based on certain criterion, such as

Fig. 1 A model of modeling



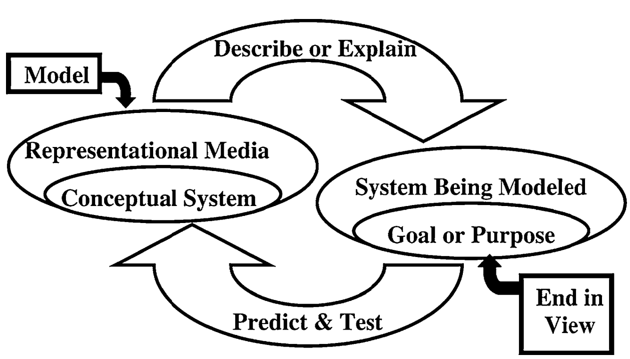


Fig. 2 A modeling cycle

logic and practice. Usually, the result is negative, or partly negative. Consequently, people need adjust the purpose, pose a new conjecture; therefore, a cycle is formed.

If the cycle is positive, it can drive people to and fulfill the purpose gradually, otherwise, people may get further and further from the purpose. Generally, people adopt two kinds of strategies to go through the cycle, one is multiple-choice trial, the other is one-way revision.

(1) Multiple-choice trial

Multiple-choice trial follows the rule-cast a wide net, search selectively. People, on the basic of conceptual sys-

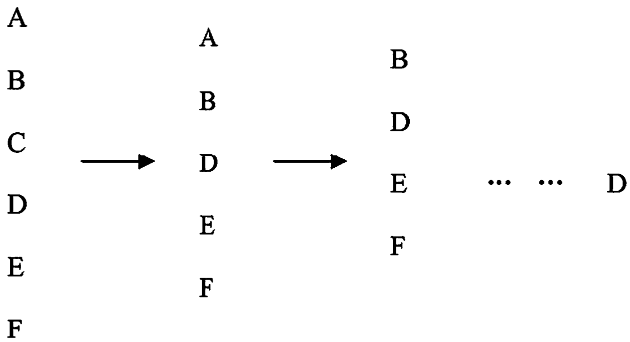


Fig. 3 Multiple-choice trial

Fig. 4 One-way revision

tem, apply imagination, analogy, and other modes of thinking, put forward all possible approach, and choose the most potentially successful to examine. If it fails to pass the test, people then choose the best one among the remaining approaches, and test it (Fig. 3). In a word, people test possible approaches successively according to probability of success until find the effective one.

(2) One-way revision

When the conjecture is denied by facts, one way is to abandon it completely and seek for another one; another way is to revise it with such approaches as adjusting former structure and introducing new auxiliary resource, in order to close the gap between model and fact. Such is one-way revision.

Firstly, people develop a conjecture A to explain a certain kind of phenomenon. However, it is found to be effective within scope; once it goes beyond the scope, inescapable error takes places. Therefore, people attempt to revise A and develop A_1, if A1 is proved to be more effective and general, but still has some flaws, keep revising and, obtain A_2, A_3, until A_n, which reaches the purpose of model construction (Fig. 4).

Based on the above research, RQ3, e.g., how chemistry teachers construct models are analyzed. To be more specific, we arranged teachers' answer according to the phrases of modeling, categorized them into different patterns and discussed their pattern based on the two strategies.

Methods

Under the guidance of the four questions, basing on a synthesis of literature (Lesh and Lehrer 2003; Justi 2005; Van Driel and Verloop 2002), and combining characteristics of the chemistry subject, we developed the test. To establish the validity of the test, we sent the test to some experts in the field of science education, as well as experienced chemistry teachers for advice. According to their feedbacks, we revised and improved the test. The final vision of the test was composed of three items, and each item's test purpose is shown in Table 1.

Table 1 Structure of the Test

Item	Testing purpose
1	Chemistry teachers' knowledge of known models and application
2	How do chemistry teachers select model
3	How do chemistry teachers construct model

Item 1 was based on the model of graphite and examined chemistry teachers' knowledge of known models; Item 1 also examined how teachers would apply the graphite model in teaching; Item 2 presented four different pictures to reflect the principle of the petroleum fractionation in order to explore teachers' criteria in selecting models; Item 3 was based on the atomic planetary model to investigate how chemistry teachers construct models.

The test was administrated for 60 min to ensure that teachers had plenty of time to answer. The subject consisted of 50 chemistry teachers, who participated in a provincial teacher training program. According to seniority (the number of years in teaching), they were divided into three groups: <10 years ($n = 12$, 24 %), 10–20 years ($n = 25$, 50 %) and 20–30 years ($n = 13$, 26 %). Because some teachers' responses were incomplete or irrelevant, which were invalid, we eliminated them and selected 39 teachers' valid responses to analyze.

Results

Question 1: What is Chemistry Teachers' Knowledge of Known Model?

Item 1 asked teachers to draw the structure of graphite in order to know the teachers' mental models of the structure of graphite. We divided the structure of graphite drew by teachers into three categories: (1) It reflected planar structure, but failed to show the three-dimensional layered structure (Fig. 5); (2) It drew three-dimensional-layered structure, but reflected the layers within graphite were completely symmetrically connected and lacked the important characteristic that carbon layers are made-up of a superimposed planar structure (Fig. 6); (3) It provided view of layer stacking (Fig. 7). Statistical data are shown in Table 2.

Fig. 5 The first category of structure

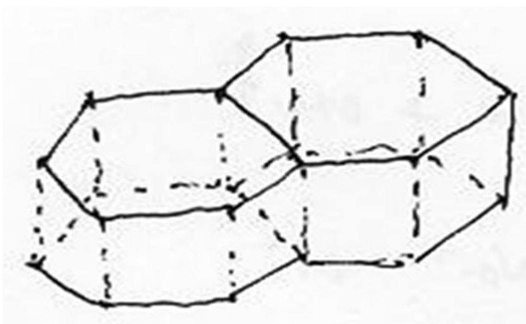


Fig. 6 The second category of structure

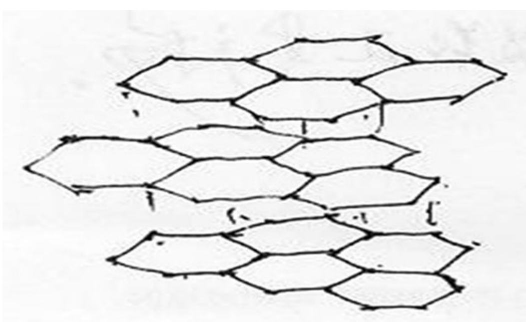


Fig. 7 The third category of structure

The planar structure and the layered structure are the essential structural characteristics of graphite. The two structural characteristics can account for some important properties of graphite. Since the connection way between layers does not play a great role in the process of explaining properties of graphite, it can be viewed as non-essential characteristics. It was found that all chemistry teachers could sketch out the planar structure; moreover, 30 out of 39 the majority of teachers (76.92 %) knew layered structure existed; however, 27 most teachers believed the layers within graphite were completely symmetrically connected. To sum up, the chemistry teachers had a relatively comprehensive understanding of the characteristics of the model of graphite. Although teachers tent to ignore "the nonessential characteristics of graphite," which indicated their knowledge was, to some degree, limited, it did not affect the teachers' understanding of the structure of graphite and their application of the model of graphite in teaching.

Question 2: How do Chemistry Teachers Select Models?

Item 2 showed four pictures or models (A, B, C and D) of petroleum fractionation (shown in Fig. 8) and required teachers to describe both advantages and disadvantages of each model and choose the most appropriate one for teaching. Picture A was a molecular analogy model, which directly represented molecular size by length to reflect the fractionation sequence, that is, small molecules were separated from the top layer, and big molecules were separated from the lower layer. Picture B was a combination of a chart model and symbolic model, using a table and chemical symbols to express the product category and composition of fractionating. Picture C was a structure model with molecular chain length to represent the molecular size, which focused on the structure of device; moreover, it listed the function of each part alongside. Picture D was a chart model, which used a frame structure to illustrate the process of petroleum fractionation.

Teachers' answers were analyzed and classified, which are shown in Table 3.

As for model A, 28 out of 39 teachers (71.79 %) mentioned "molecular size is explicit" in model A, and 15 teachers (38.46 %) thought it "reflected the essential principle." On the other hand, 32 teachers (82.05 %) mentioned "category is not clear, information scanty" and 25 teachers (64.10 %) thought model A was "abstract, distant from reality." In total, only 2 out of 39 teachers (5.13 %) chose model A.

As for model B, the major advantages are "clear data" and "detailed information," referred to by 27 teachers (69.23 %) and 17 teachers (43.59 %), relatively; the major disadvantage is "not graphic, abstract" ($n = 25, 64.10$ %). The majority of teachers thought model B was too hard for students to understand, and only 3 teacher (7.69 %) chose model B.

As for model C, nearly all teachers noticed its greatest advantage ($n = 37$, 94.87 %), that is, "direct-viewing, vivid," and 24 out of 39 teachers (61.54 %) mentioned "familiar in daily life, and reflected the production process." Moreover, just a few teachers noticed model C "reflected molecular size" ($n = 4$, 10.26 %). On the other hand, teachers noticed model C's disadvantages, 20 teachers (51.28 %) thought "information is insufficient and does not explain the specific material and boiling point," 7 teachers (17.95 %) referred to "lack a theory or knowledge of a good system." On a whole, participants tent to prefer model C, chose by 31 teachers (79.49 %).

Table 2 Classification and statistics of the model of graphite drew by teachers

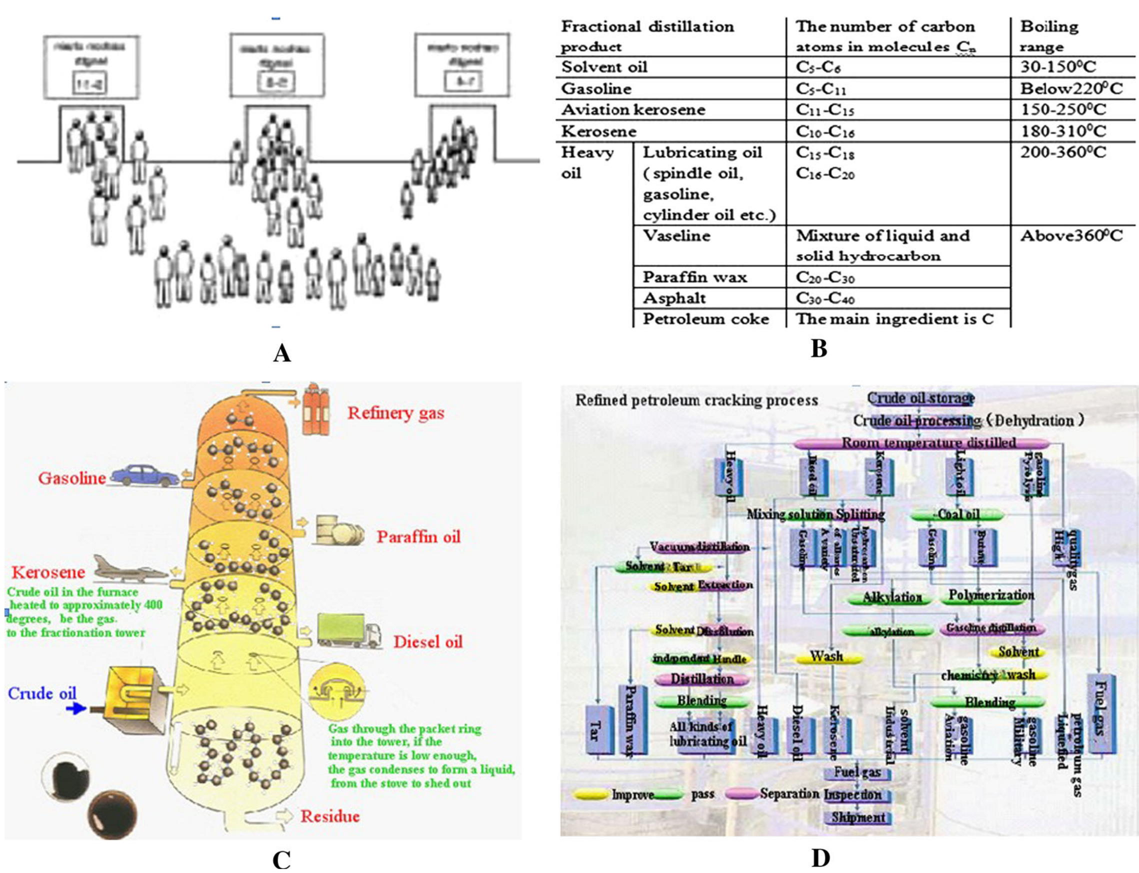
Category	1 Planar structure	2 Three-dimensional layered structure	3 Three-dimensional structure of interlayer sliding
The number of teachers	9	27	3
Percentage	23.07	69.23	7.69

As for model D, 32 out of 39 teachers (82.05 %) mentioned "fractionation process is very detailed, clear context," and 14 teachers (35.90 %) "knowledge is complete and systemic." On the other hand, 33 teachers (84.62 %) cited "the chart is too abstract," and 21 teachers (53.85 %) mentioned "not easy for students to understand; easy to get students weary of studying." Only 3 teachers (7.69 %) chose model D.

From the statistics, teachers paid more attention to the product category, and the range of boiling point and the structure of the device, while ignoring the principle. A total of 31 teachers chose picture C, but just a few of the teachers had mention of one of its advantages ($n = 4$), that is, it reflected the molecular size clearly with molecular chain length. From the presentation mode of content, teachers preferred models that were vivid and direct viewing. As for pictures A, B and D, teachers generally listed their disadvantages: "abstract," and "not easy for students to understand," many teachers selected the picture C ($n = 31, 79.49$ %), for it is more interesting and imaginative than the others (37 teachers mentioned C's advantage ①, see Table 3).

The results showed that chemistry teachers' criteria of model selection were that the model is visual, concise, familiar in everyday life, and easy to arouse students' interest; However, teachers tent to overlook the principles of a model. Model reflects the essential characteristics of things; in chemistry education, model should be used to help students understand and master knowledge. It is important to pay greater attention to help students master the application of petroleum fraction, above all, i.e., understanding the principle of petroleum fraction. Therefore, teachers, when selecting and applying models in teaching, need to focus on models reflecting the principle and improve students' understanding of the principle.

Teachers' answers were analyzed and classified, which are shown in Table 3. From the statistics, teachers paid more attention to the product category, and the range of



Fractional distillation product		The number of carbon atoms in molecules C_n	Boiling range
Solvent oil		C_5-C_6	30-150°C
Gasoline		C_5-C_{11}	Below220°C
Aviation kerosene		C_{11}-C_{15}	150-250°C
Kerosene		C_{10}-C_{16}	180-310°C
Heavy oil	Lubricating oil (spindle oil, gasoline, cylinder oil etc.)	C_{15}-C_{18} C_{16}-C_{20}	200-360°C
	Vaseline	Mixture of liquid and solid hydrocarbon	Above360°C
	Paraffin wax	C_{20}-C_{30}	
	Asphalt	C_{30}-C_{40}	
	Petroleum coke	The main ingredient is C	

Fig. 8 Various model of petroleum fractionation

Table 3 Descriptive statistics on teachers' answers to Item 2

	Advantage	Number of teacher referring to	Disadvantage	Number of teacher referring to	Number of teachers choosing	%
A	① Molecular size is explicit	28	① Category is not clear, information scanty	32	2	5.13
	② Reflect the essential principle	15	② Abstract, distant from reality	25		
			③ Students can't understand easily	8		
B	① Clear data	27	① Not graphic, abstract	25	3	7.69
	② Detailed information	17	② Complex	10		
			③ Separate from the production process	6		
C	① Direct-viewing, vivid	37	① Information is insufficient, and does not explain the specific material and boiling point	20	31	79.49
	② Familiar in daily life, reflect the production process	24	② Lack a theory; no knowledge of a good system	7		
	③Reflect the molecular size	4				
D	① Fractionation process is very detailed, context clear	32	① Chart is complicated, too abstract	33	3	7.69
	② Knowledge is complete and Systemic	14	② Not easy for students to understand; easy to get students weary of studying	21		

boiling point and the structure of device, when teaching petroleum fractionation, while ignored the principle. A total of 31 teachers chose picture C, but most of the teachers had no mention of one of its advantages, that is, it reflected the molecular size clearly with molecular chain length. From the presentation mode of content, teachers preferred models that were vivid and direct viewing. As for pictures A, B and D, teacher generally listed their disadvantages: "abstract," and "not easy for students to understand." Most of the teachers selected the picture C, for "it is more interesting and imaginative than the others."

The results showed that chemistry teachers' standard of model selection was the model that is visual, concise, close to life, easy to arouse students' interest; however, teachers tent to overlook the principles of model. Model reflects the essential characteristics of things; in chemistry education, model should be used to help students understand and master knowledge. It is significant to pay great attention to help students master the application of petroleum fraction, above all, understand the principle of petroleum fraction. Therefore, teachers, when selecting and applying models in teaching, need to focus on models reflecting the principle and improve students' understanding of the principle.

Question 3: How Chemistry Teachers Construct Models?

Item 3 dealt with the atomic planetary model, which aimed at exploring chemistry teachers' model construction. According to teachers' responses, we grouped teachers' answers into "No answer" "Wrong answer" and "Correct answer"; the results are shown in Table 4.

From Table 4, we know that 7 out of 39 teachers did not answer (17.95 %), and seven teachers' answers were wrong (17.95 %), which was an astonishing number.

Among the seven teachers who did not answer, some teachers wrote "I have never taught it in class", "I've not paid much attention to this content basically," "this part is not important and rarely taught" and so on.

Moreover, seven teachers' answers were wrong, and they mentioned:

> "…… alpha particle scattering experiment shows atoms do irregular movement, when they collide with one another, some will rebound, some will go forward….."
> "…alpha particles can pass through gold foil, and the majority of them have low mass; when they pass through gold foil, a few of those with relatively large mass bounce back…. there is space among particles of gold foil…"
> "… alpha particle scattering experiment shows atoms can be divided further, and there is space among atoms……"

Thus it can be seen that alpha particle scattering experiment were strange to some chemistry teachers, who, to different degrees, held misconceptions about the important principles that reveal the structure of the atom.

We referred to relevant materials and found that alpha particle scattering experiment was supplementary knowledge (knowledge is not required in chemistry curriculum standard, in other words, whether it is learned depends on students' needs). Therefore, teachers tent to pay less attention to it, and some teachers were not familiar with it, which prevented them solving this problem successfully. However, alpha particle scattering experiment is one of classical experiments during the development of the atomic theory, which can not only improve students' understanding of the structure of the atom, but also enhance their scientific literacy. Teacher kept the structure of the atom by heart, but ignored the process of putting forward the structure model. The results, on one hand, indicated that teachers' knowledge of the atomic planetary model needed to be improved; on the other hand, teachers likely failed to take advantages of this model in their teaching.

Patterns of Model Construction

We analyzed the 25 teachers' design ideas whose answers to Item 3 were correct, that is, how to help students to establish the atomic planetary model. According to teachers' responses, 4 patterns were found, which we called Pattern A, Pattern B, Pattern C and Pattern D, respectively. The four patterns are discussed in Figs. 9, 10, 11 and 12.

Table 4 Teachers' answer to Item 3

Answer	Number	Percentage
No answer	7	17.95
Wrong answer	7	17.95
Correct answer	25	64.10

In Pattern A, teachers usually used things in daily life (we call them "known similar system") to create an analogy, to guide students to build the atomic structure model, and then describe and explain relevant experiments with models to help students understand the structure of the atom (Fig. 9).

Case A:

A teacher took "the bullet hits obstacles" as an example, analyzed the characteristics of the motion and explained the different motion trajectories owing to different forces between the bullet and obstacles. The teacher, by means of a relevant graph, related the familiar phenomenon with alpha particle scattering experiment, to explain alpha particles' different characteristics of motion when they passed through atoms, and inferred that the atom is not a homogeneous entity, but one with a heavy nucleus at the center.

From the above case, we can see this teacher made use of daily things students were familiar with to make up an analogy. This analogy built a bridge between the microscopic and the macroscopic, thus promoted students' understanding of alpha particle scattering phenomenon, and then helped them construct atomic planetary model.

In patter B, teachers usually described in words or graphs to represent experimental facts, from which teachers developed the atomic planetary model by abstracting and

Fig. 9 Pattern A

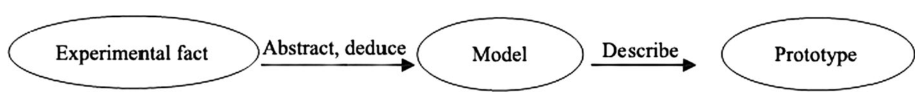


Fig. 10 Pattern B

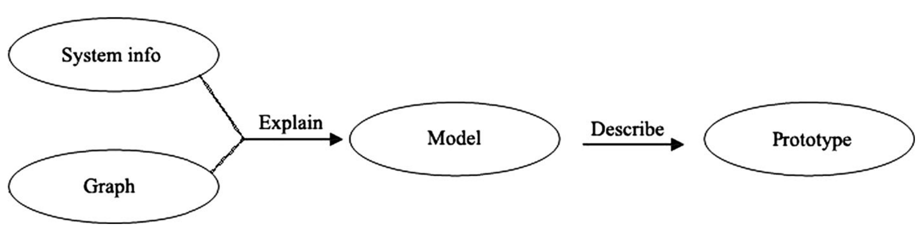


Fig. 11 Pattern C

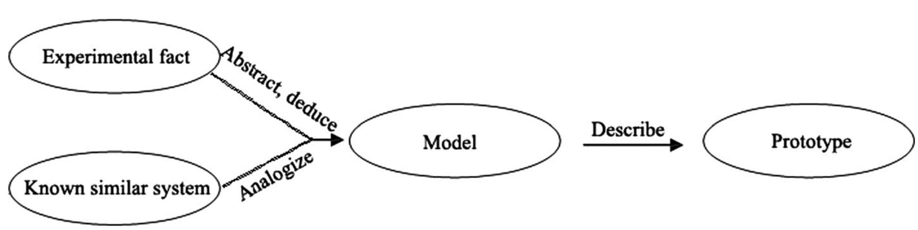


Fig. 12 Pattern D

deducing, and then used the planetary model to describe the structure of the atom (Fig. 10).

Case B:

Introduction to alpha particle scattering experiment

① A large number of alpha particles pass through gold foil} There is relatively large space within atoms
② A small number of alpha particles deflect
③ A small number of alpha particles deflect greatly, that is, the angle can even reach 180° → There exists a small and positively charged nucleus within an atom.

Comprehensive analysis: Atom has a small with positively charge nucleus. Because metal is electrically neutral, there should be electrons outside a nucleus. To conclude, an atom consists of a nucleus and electrons.

From case B, we can see that teachers, relying on experimental phenomena, invented the reason behind phenomena and then constructed the atomic planetary model. This pattern follows a widely used process of scientific exploration, that is, exploring from the phenomenon to the fundamental.

In Pattern C, teachers directly introduced various views of model of the atomic structure, and elaborated the movement characteristics of extranuclear electron, and then represented the theoretical knowledge by a graph to strengthen student's understanding, and finally helped students construct the atomic planetary model. In this pattern, teachers used a deductive reasoning method (Fig. 11).

Case C:

1. Tell students about various views during the development of the atomic theory.
2. Focus on Rutherford's atomic planetary model
3. Display or demonstrate the alpha particle scattering experiment
4. Guide students to discuss the advantages and disadvantages of the atomic planetary model.

As seen from case C, from the perspective of language features, this teacher's idea was clear, and the answer was concise; from the perspective of thinking traits, the teacher, starting with the atomic theory, applied "from the nature to the phenomenon" teaching method instead of "from the phenomenon to the fundamental" one.

In Pattern D, teachers used experimental phenomena to create an abstract model, at the same time, used analogy with similar system to deepen students' understanding of the structure of the atom and then constructed the model (Fig. 12).

Case D:

1. Analyze the Rutherford alpha particle scattering experiment
2. Draw a conclusion: The atom is made-up of a tiny positively charged nucleus at the center and electrons around; electrons are running outside the nucleus at a high speed.
3. Compare the relationship between the movement of the earth and movement of the moon.
4. Construct the atomic planetary model.

In case D, firstly the teachers also inferred the atomic planetary model by explaining relevant phenomenon, and then turned to familiar objects in real life to draw an analogy, and ultimately helped students build the targeted model.

Table 5 shows how chemistry teachers used the four different patterns when they constructed models. As seen from Table 5, teachers preferred using Pattern B ($n = 9$, 36 %). These teachers tent to show the experimental phenomenon (by means of multimedia, pictures, etc.), make an inference, and then constructed the atomic planetary model. This process accorded with general cognitive laws, which students were more likely to follow, so most teachers chose this pattern.

When constructing a model, people usually make a comprehensive analysis of the problem first, and put forward a tentative solution scheme, and then use logic and practice standards to test it. If the test results are negative, or partly negative, people need to adjust the target and develop new hypothesis; in this way, a new cycle is formed.

The above 4 patterns indicated that chemistry teachers generally adopted "one-way revision" patterns; their pattern lacked the step "test;" In other word, teachers ignored model evaluation and correction. Teaching practice usually involves established knowledge; therefore, teachers may often ignore this step. However, model construction requires a continuous cyclic process; the step "test" can help students to cultivate an error correction ability. As a result, teachers should pay more attention to the step "test" when applying models in teaching.

Question 4: How do Chemistry Teachers Use the Model in Teaching?

Application Purpose

Using the structure of graphite is to explain three types of properties of graphite: I structural stability and

Table 5 Teachers' use of fourpatterns

Model	Used number	Percentage
A	6	24
B	9	36
C	6	24
D	4	16

Table 6 Knowledge analysis of Item 1

Property	A (structure description)	B (nature explanation)	C (application listing)
I (structural stability and thermostability)	IA	IB	IC
II (electrical conductivity and thermal conductivity)	IIA	IIB	IIC
III (soft and slippery)	IIIA	IIIB	IIIC

We analyzed teachers' answers statistically according to the above nine aspects. The statistical results of tests are shown in Table 7

Table 7 Results of teachers' answers ($n = 39$)

knowledge	Mentioned and correct		Mentioned but error		Not mentioned	
IA	22	56.41 %	0	0 %	2	5.13 %
IB	9	23.07 %	1	2.56 %	14	35.89 %
IC	4	10.25 %	0	0 %	20	51.28 %
IIA	2	5.13 %	10	25.41 %	12	30.77 %
IIB	10	25.41 %	2	5.13 %	12	30.77 %
IIC	1	2.56 %	0	0 %	23	58.97 %
IIIA	24	61.53 %	0	0 %	0	0 %
IIIB	12	30.77 %	0	0 %	12	30.77 %
IIIC	7	17.95 %	0	0 %	17	43.59 %

thermostability (high temperature resistance); II electrical conductivity and thermal conductivity; and III a soft and slippery feel, thus is used as a dry lubricant. Therefore, this problem investigated whether teachers could describe the structure of graphite to students from the above three types of properties (A), explain its properties based on its structure (B), and list its application (C). The process goes from the easy to the complicated, from the abstract to the concrete. As shown in Table 6, we defined the three types of properties asI,II and III, the three aspects we investigated as A, B and C.

From the above statistical results, we could find that as follows:

(1) As for structure description, teachers could correctly describe the two properties of "stability and thermostability" and "soft and lubricity" (The C atom on the plane and the adjoining three C atoms form three covalent bonds together, and layers are connected by van der Waals force). The number of IA and IIIA was 22 (56.41 %) and 24 (61.53 %), respectively. However, there was a big problem in teachers' cognition on the structure characteristics which allowed graphite to conduct electricity and heat. Only two teachers successfully described IIA, and 10 (35.41 %) teachers' explanations were wrong. Almost all teachers who gave answers mentioned that as follows: "The three electrons of C atom participate in bonding, the remaining one is a free electron, it can move freely between the layers, so that graphite has electrical conductivity and thermal conductivity."

However, the correct reason should be "Those C atoms of each layer are combined with the adjoining three C atoms by σ bonds with sp^2 hybrid orbits and form a hexagonal symmetry plane lamellar structure which is infinite. The distance between the adjacent C atoms in the layer is 142 pm. Each C atom has a p orbit which is perpendicular to the plane but parallel with each other. These p orbits overlap with each other and group into big π bond, and π electrons (delocalized electrons) of the π bond can move freely on the whole carbon atom plane. So, graphite has conductivity similar to metal."

The teachers thought that it was the existence of free electrons that yielded the electrical conductivity and thermal conductivity of graphite, rather than the formation of big π bond; and there were 12 out of 39 teachers (30.77 %) teachers that did not mention the reason of the graphite's conductivity. Three of them mentioned that they analyzed the interatomic bond length and the distance between layers of graphite crystal structure in teaching and calculated the average carbon atom numbers of each hexagon. In middle school, teachers should put particular emphasize on the phenomenon explanation, knowledge acquisition and creative imagination development when applying this model, rather than focusing on rote memorization of such knowledge as concrete numerical value.

(2) As for property explanation, only few teachers used structure to explain properties after the description of the structure. There were 9 (23.07 %) and 12 (30.76 %) teachers who mentioned IB and IIIB, respectively, compared to IA and IIIA, and the number reduced to a great degree. Moreover, the number of teachers who mentioned the conductive properties of graphite (IIB) was 10 (25.64 %), which was higher than the number of teachers describing reason of conductivity previously. So, the teachers knew the conductivity of graphite, but lacked clear understanding of the reason.

(3) As for application listing, only a small part of teachers stated the use of the graphite at the end of the design. There were only 4 (10.26 %), 1 (2.56 %) and 7 (17.95 %) teachers that mentioned IC, IIC and IIIC, respectively.

Application Pattern

The results showed that most of the chemistry teachers (69.23 %) used models in teaching and generally followed a set of process, as shown in Fig. 13.

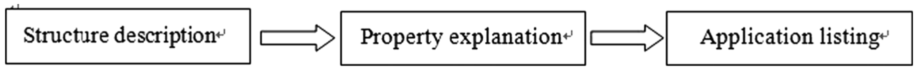


Fig. 13 The basic process of teachers' using model in teaching

Firstly, teachers usually described the existing model of graphite in teaching, and then inferred the main chemical properties and physical properties of graphite from its structure, and finally, illustrated the application determined by properties. This process conformed to both the students' basic cognitive process and teachers' thinking habits.

Case E:

① Each carbon atom in graphite is covalently bonded to the three other carbon atoms, and they form flat layers of hexagon called grapheme sheets.	→	High-melting point	→	Crucible, refractory material
② Between the layers the bond is weak, which is Van der Waals force.	→	Soft and satiny	→	Pencil, lubricant
③ Each carbon atom has an free electron, which can move freely between layers.	→	Electrical conductivity	→	Conductive material

Analysis: The teacher's design was clear; the use of serial numbers and arrows made the design more coherent and logical; it went through a process from the structure to the fundamental, and then to the application, which indicated the teacher's ideas were clear.

In addition, some teachers' responses lacked accuracy, and their answer ideas were disorganized.

Case F:

In the process of teaching, teachers point out the formation of covalent bonds among each carbon atom and three other carbon atoms of graphite by using the solid lines in the picture. While the dashed lines represent the intermolecular forces bonding the layers, the other unpaired electrons can move freely. Graphite can be conductive. The layers can easily slide over each other making graphite soft and slippery and an excellent lubricant. It has high-melting point and boiling point.

Analysis: The teacher described the structure model and properties of graphite in isolation from each other, and the casual relationship between them was unclear which was just like the accumulation of knowledge.

For most teachers, their basic process of using models was similar, but the ways were different. Some teachers' designs and responses were coherent and clear, while others' designs lacked a sense of order. Some teachers lacked certain methodology guidance (it reflected their thinking of teaching design and language organization).

The analysis above showed that chemistry teachers mostly followed a similar process in the teaching using models, that is, explained "properties" with "structure," and illustrated "application" based on "properties." Chemistry teachers tent to apply models in causal explanation of phenomenon and paid more attention to using models to help students to understand contents, rather than the deeper role of models, such as developing students' knowledge of scientific methods and ability of thinking. The result, on the other hand, indicated that teachers' understanding of the role of models was not comprehensive.

Discussion and Conclusions

In this research study, we investigated chemistry teachers' knowledge and application of models, including the understanding of existing models, model selection, model construction and model application. According to the above results, we can draw the following conclusions:

1. The models drew by chemistry teachers basically reflected the essential characteristics of things;
2. Chemistry teachers preferred models that were vivid when they selected models for teaching;
3. Chemistry teachers ignored the step test in the process of model construction;
4. Chemistry teachers tent to apply models by following a "structure-property-use" pattern.

The research results showed that the chemistry teachers' knowledge of models was incomplete, and their application of models in teaching needed to be improved. To improve teachers' understanding of models and promote application of models, teacher training programs need to pay attention to application of model in teaching. Efforts can be made according to the following suggestions:

(1) Promote chemistry teachers to construct models

The process of constructing scientific models can cultivate the ability of students to think and solve problems. Teachers themselves have to have a comprehensive understanding of the modeling process, especially the "test" and "correction" steps, which can cultivate students ability of reflection and error correction. Teachers also should understand the conditions of establishing scientific models, that is, basing on precise experiments and rich observation materials, integrating imagination and creation and applying various logical thinking methods under the guidance of scientific theories.

(2) Help chemistry teachers to master methods of model application

Only by integrating models and teaching effectively, can teachers make full use of models in teaching. In chemistry teaching, model application involves two aspects: model selection and teaching procedure. Selecting an appropriate model rests on the premise that theoretical knowledge can be well represented. Moreover, an appropriate model needs to be used effectively through reasonable teaching procedures; otherwise, students may be misled or put more learning pressure. The interaction of the two aspects can help students understand knowledge and master methods correctly. In order to succeed in selecting models and representing models, teachers need accumulate teaching experience and enhance communication with students.

(3) Develop various methods to improve teachers' understanding and application of models

Besides traditional methods, such as educational training programs, conferences, more innovative methods should be applied. For example, Justi and van Driel (2005) suggested action research, that is, teachers themselves are researchers and subjects, for one thing, they design and practice model-based instruction, for another thing, they reflect on their teaching process as well as their understanding of models. This two-way process has been proved to effectively improve teachers' understanding and application of models. What's more, the component of technology can be integrated, such software as Model—it can give teachers interesting experience of model-based instruction as well as change their ideas of models and teaching.

This research only involved 50 teachers. The limited number of participants and the chosen research design (i.e., survey) prevent broad generalizations to be made; more participants should be involved and an observation sheet should be used to introduce a clear aspect for teaching performances through chemistry teachers. Moreover, we are not able to probe into more details on teachers' knowledge and application of models. The influence of teachers' understanding of model on teachers' application of model also needs further research.

References

Bell PH (1995) How far does light go? Individual and collaborative sense-making of scientific evidence. Presented at the Annual Conference of the American Educational Research Association, San Francisco, CA

Buckley BC, Gobert JD et al (2004) Model-based teaching and learning With Biological TM: what do they learn? How do they learn? How do we know? J Sci Educ Technol 13(1):23–41

Crawford BA, Cullin MJ (2002) Engaging prospective science teachers in building, testing, and teaching about models Paper presented at the annual meeting of International Conference of the National Association for Research in Science Teaching, New Orleans, LA

Crawford B, Cullin M (2004) Supporting prospective teachers' conceptions of modeling in science. Int J Sci Educ 26:1379–1401

Crawford B, Cullin M (2005) Dynamic assessments of preservice teachers knowledge *Research* and the Quality of Science Education. Springer, Dordrecht, pp 309–324

De Jong O, Van Driel JH (2001) Developing pre-service teachers' content knowledge and PCK of models and modelling. Paper presented at the National Association for Research in Science Teaching Annual Meeting, St. Louis, MO

Dori YJ, Barak M (2001) Virtual and physical molecular modeling: fostering model perception and spatial understanding. Educ Technol Soc 4:61–74

Duit R, Glynn S (1996) Mental modelling. In: Welford G, Osborne J, Scott P (eds) Research in science education in Europe: current issues and themes. Falmer, London, pp 166–176

Erduran S, Duschl R (2004) Interdisciplinary characterization of models and the nature of chemical knowledge in the classroom. Stud Sci Educ 40:111–144

Fang L (2003) Research on the development of teachers' pedagogical knowledge. East China Normal University Press, Shanghai, p 26

Francoeur E (1997) The forgotten tool: the design and use of molecular models. Soc Stud Sci 27:7–40

Gilbert SW (1991) Model building and a definition of science. J Res Sci Teach 28:73–80

Gilbert JK, Boulter C, Rutherford M (1998) Modeling explanations part I: horses for courses. Int J Sci Educ 20(1):83–97

Gilbert JK, Boulter CJ, Elmer R (2000) Positioning models in science education and in design and technology education. In: Gilbert JK, Boulter CJ (eds) Developing models in science education. Kluwer, Dordrecht, pp 3–18

Glynn SM (1991) Explaining science concepts: a Teaching-with-Analogies Model. In: The psychology of learning science. Lawrence Erlbaum, New Jersey, pp 219–240

Grosslight L, Unger C, Jay E, Smith CL (1991) Understanding models and their use in science: conceptions of middle and high school students and experts. J Res Sci Teach 28(9):799–822

Halloum I (1996) Schematic modeling for meaningful learning of physics. J Res Sci Teach 33(9):1019–1041

Harrison AG (2000) A typology of school science models. Int J Sci Educ 22:1011–1026

Harrison A (2001) Models and PCK: their relevance for practicing and pre-service teachers. Paper presented at the annual meeting of the National Association for Research in Science Teaching, St. Louis, MO

Harrison AG, Treagust DF (1996) Secondary students' mental models of atom and molecules: implications for teaching chemistry. Sci Educ 80:509–534

Harrison AG, Treagust DF (1998) Modeling in science lessons: are there better ways to learn with models? Sch Sci Math 98(8):420–429

Harrison AG, Treagust DF (2000) Learning about atoms, molecules, and chemical bonds: a case study of multiple model use in grade 11 chemistry. Sci Educ 84:352–381

Harrison AG, Treagust DF (2001) Conceptual change using multiple interpretive perspectives: two case studies in secondary school chemistry. Instr Sci 29:45–85

Henze I, Van Driel J, Verloop N (2007) Science teachers' knowledge about teaching models and modeling in the context of a new syllabus on public understanding of science. Res Sci Educ 37(2):99–122

Hodson D (1993) Re-thinking old ways: toward a more critical approach to practical work in school science. Stud Sci Educ 22:85–142

Ingham AM, Gilbert JK (1991) The use of analogue models by students of chemistry at higher education level. Int J Sci Educ 22(9):1011–1026

Jungck J, Calley J (1985) Strategic simulations and post-socratic pedagogy: constructing computer software to develop long-term inference through experimental inquiry. Am Biol Teach 47:11–15

Justi RS (2005) The development of science teachers' knowledge on models and modeling : promoting, characterizing and understanding the process. Int J Sci Educ 27(5):549–573

Justi RS, Gilbert JK (2001) Teachers'views on models and modelling in science education. Paper presented at the Annual Meeting of the National Association of Research in Science Teaching, St. Louis, MI

Justi R, Gilbert J (2002a) Modelling, teachers' views on the nature of modelling, implications for the education of modellers. Int J Sci Educ 24(4):369–387

Justi RS, Gilbert JK (2002b) Science teachers' knowledge about and attitudes towards the use of models and modelling in learning science. Int J Sci Educ 24:1273–1292

Justi R, Gilbert JK (2003) Investigating teachers' ideas about models and modelling: Some issues of authenticity. Paper presented at the Fourth International Conference of the European Science Education Research Association, Noordwijkerhout, The Netherlands

Justi R, van Driel J (2005) A case study of the development of a beginning chemistry teacher's knowledge about models and modelling. Res Sci Educ 35:197–219

Klein U (2003) Experiments, models, paper tools. Culture of organic chemistry in the nineteen century. Stanford University Press, Stanford, Lehrer

Koponen I (2007) Models and modelling in physics education: a critical re-analysis of philosophical underpinnings and suggestions for revisions. Sci Educ 16(7–8):751–773

Kozma R, Russell J (2005) Students becoming Chemists: Developing representational competence. In: Gilbert JK (ed) Visualization in science education. Springer, Dordrecht, pp 121–145

Lehrer R, Schauble L (2000) Modeling in mathematics and science. In: Glaser R (ed) Advances in instructional psychology: Volume 5: Educational design and cognitive science Mahwah. Erlbaum, NJ, pp 101–159

Lesh R, Doerr HM (2003) Foundations of models and modeling perspective on mathematics teaching, learning, and problem solving. In: Lesh R, Doerr HM (eds) Beyond constructivism:Models and modeling perspectives on mathematics problem solving, learning, and teaching Mahwah. Erlbaum, NJ, pp 3–33

Lesh R, Lehrer R (2003) Models and modeling perspectives on the development of students and teachers. Math Think Learn 5(2–3):109–129

Linn M (2003) Technology and science education: starting points, research programs, and trends. Int J Sci Educ 25(6):727–758

Louca L, Zacharia Z, Constantinou C (2011) In quest of productive modeling-based learning discourse in elementary school science. J Res Sci Teach 48(8):919–951

National Research Councile (2012) A Framework for K-12 science education: Practices, crosscutting concepts, and core ideas. The National Academy Press, Washington, D.C

Pluta JW, Chinn AC, Duncan GR (2011) Learners' epistemic criteria for good scientific models. J Res Sci Teach 48(5):486–511

Schwarz C (2009) Developing preservice elementary teachers' knowledge and practices through modeling-centered scientific inquiry. Sci Educ 93:720–744

Schwarz C, White B (1998) Fostering middle school students' understanding of scientific modelling. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA, pp 13–17

Schwarz C, White B (2005) Meta-modeling knowledge: developing students' understanding of scientific modeling. Cogn Instr 23(2):165–205

Schwarz CV, Reiser BJ et al (2009) Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. J Res Sci Teach 46(6):632–654

Sins PHM, Savelsbergh ER, van Joolingen WR, van Hout-Wolters BHAM (2009) The relation between students' epistemological understanding of computer models and their cognitive processing on a modelling task. Int J Sci Educ 31(9):1205–1229

Smit JJ, Finegold M (1995) Models in Physics: perceptions held by final-year Prospective Physical Science Teachers Studying at South African Universities. Int J Sci Educ 19:621–634

Stratford S (1996) Investigating processes and products of secondary science students using dynamic modelling software. Unpublished doctoral dissertation. University of Michigan

Suckling CJ, Suckling KE, Suckling CW (1978) Chemistry through models. concepts and applications of modeling in chemical science, technology and industry. Cambridge University Press, Cambridge

Talanquer V (2011) School chemistry: The need for transgression. Science & Education, published online 17th September

Van Driel Jan H, Verloop Nico (1999a) Teachers' knowledge of models and modeling in science. Int J Sci Educ 21(11):1141–1153

Van Driel JH, Verloop N (1999b) Teachers' knowledge of models and modeling in science. Int J Sci Educ 21(11):1141–1153

Van Driel JH, Verloop N (2002) Experienced teachers' knowledge of teaching and learning of models and modeling in science education. Int J Sci Educ 24(12):1255–1272

Wei S (2011) Using Rasch modeling to develop computer modeling-based instruments for measuring middle school students' conceptual understanding of structure of matter. Unpublished dissertation, East China Normal University

Windschitl M, Thompson J (2006) Transcending simple forms of school science investigations: can preservice instruction foster teachers' understandings of model-based inquiry? Am Educ Res J 43(4):783–835

Windschitl M, Thompson J, Braaten M (2008) Beyond the scientific method: model-based inquiry as a new paradigm of preference for school science investigations. Sci Educ 92:941–967

Wisnudel-Spitulnik M, Kracjik J, Soloway E (1999) Construction of models to promote scientific understanding. In: Feurzeig W, Roberts N (eds) Modelling and simulation in science and mathematics. Springer-Verlag, New York, pp 70–94

Yang Y (2012) Research on measurement and assessment of the competence in discipline of chemistry. Unpublished dissertation, East China Normal University

Zhang B, Liu X, Krajcik J (2006) Expert models and modeling processes associated with a computer-modeling tool. Sci Educ 90(4):579–604

J Sci Educ Technol (2014) 23:227–237
DOI 10.1007/s10956-013-9484-2

Students' Perceptions of Their Science Teachers' Pedagogical Content Knowledge

Lilia Halim · Sharifah Intan Sharina Syed Abdullah · T. Subahan Mohd Meerah

Published online: 20 December 2013

Abstract Pedagogical content knowledge (PCK) is a type of teacher knowledge to be developed by a teacher. PCK is said to contribute to effective teaching. Most studies investigated the development of PCK and its influence on students' learning from the teachers' perspectives. Only a limited number of studies have investigated the components of science teachers' PCK that helped students' learning from the perspective of students. Thus, it is the aim of this study to investigate the level of science teachers' PCK from students' perspective, in particular whether or not students of different achieving ability had different views of teachers' PCK in assisting their learning and understanding. Based on the PCK research literature, six components of PCK have been identified, which were as follows: (1) subject matter knowledge, (2) knowledge of teaching strategies, (3) knowledge of concept representation, (4) knowledge of teaching context, (5) knowledge of students, and (6) knowledge of assessment in learning science. A questionnaire consisting of 56 items on a five-point Likert-type scale were used for data collection from 316 Form Four students (16 years old). One-way analysis of variance revealed that the differences in science teachers' PCK identified by students of different achieving abilities were statistically significant. Overall, students of various academic achieving abilities considered all the components of PCK as important. The low-achieving students viewed all the components of PCK as being less important compared to the high and moderate achievers. In particular, low-achieving students do not view 'knowledge of concept representation' as important for effective teaching. They valued the fact that teachers should be alert to their needs, such as being sensitive to students' reactions and preparing additional learning materials. This study has revealed that PCK of science teachers should be different for high and low-achieving students and knowledge of students' understanding plays a critical role in shaping teachers PCK.

Keywords Pedagogical content knowledge · Science teaching · Secondary students · Different abilities · Students' needs · Effective science teaching

L. Halim (✉) · S. I. S. S. Abdullah · T. S. M. Meerah
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia
e-mail: profdrlilia@gmail.com; lilia@ukm.my

Introduction

Educational problems are too complex to be attributed to a single factor or a small number of factors (Ingersoll 1999). Yet, it is generally agreed that effective teachers are central to effective science teaching. However, aspects of the quality of science teachers are very extensive which can be described in a variety of features, making it difficult to measure (Rockoff 2004). However, one of the characteristics of effective teachers is the pedagogical content knowledge (PCK) and it is seen as the core of teachers' knowledge in developing effective teachers (Loughran et al. 2004; Abell and Lederman 2007). Nargund-Joshi et al. (2011) opined that PCK could be regarded as a special knowledge program acquired by teachers to facilitate their transformation in subject matter knowledge in order to help in student learning process. The trend among science educators nowadays, has shifted to PCK researches, indicating the importance of how PCK has become a convergence of teachers pedagogy and understanding of content (Abell and Lederman 2007), as the gateway for promoting quality teaching and useful learning.

Students' Perceptions of PCK

While PCK is shown to be integral to effective science teaching, limited researches have been conducted on assessing science teachers' PCK and how their PCK affects students' learning (Abell 2008). As argued by Park and Oliver (2008), in order for the concept of PCK to be more useful, the assumption that PCK is highly related to students' learning should be further investigated. One reason for the lack of investigating the assumption is that often, the common methods such as concept mapping, checklists, classroom observations, and structured interviews that have been used to assess teachers' PCK and its impact on learning as shown in the work by Loughran et al. (2004), is time consuming and laborious. At the same time, there appear to be lack of research on how to develop tools for assessing teachers' PCK and its link to learning (Manizade and Mason 2011).

One way to test the assumption that PCK is highly related to students' learning is to gauge from the students' perspectives of what sort of teacher's knowledge that will help them learn science. Rudduck et al. (1996) noted that one way to improve teaching is to encourage pupils to talk about what makes learning difficult for them as well as asked them how teaching can be improved. The researchers further assert that teachers could learn from consultation from the students.

As argued by Jang (2010), teachers are the authority in the traditional classroom and can easily have self-centered thinking. Hence, teachers might have difficulty to reflect their teachings as well as their PCK as they themselves are being the subject of evaluation. Using students' perceptions will enable teachers to appreciate the perceived instructional influences on students' learning processes (Senocak 2009). It could also help teachers improve their teaching based on the students' perceptions. Students' views on what are needed of their teachers in promoting science learning could also provide information on the quality of their teachers' PCK. Thus, it can be argued that one can understand effective PCK from students' perspectives which in turn can also help teachers to develop and refine their PCK for students' learning.

Tuan et al. (2000) pointed out that students' viewpoints of their teachers might not be consistent with the reality generated by outside observers; however, students' perceptions could present the range of reality for themselves and their peers in the classroom. When students' perceptions of teachers' knowledge (SPOTK) is taken into account in a study, the assumption is absolutely dependent on the fact that they have been taught by the teachers and their minds are already pre-occupied with memories and reactions that inventory for data collection will measure (Adediwura and Tayo 2007).

Tuan et al. (2000) and Jang (2010) assessed students' perceptions of teachers' PCK using a survey questionnaire, with Tuan et al. focusing on secondary students' views while Jang investigated college students' views of their professors PCK. Both studies assumed that there is a link between teachers' PCK and student learning. The outcome of their research work argued that teachers and researchers came to appreciate the perception of the students regarding their learning processes as influenced by both the environment and instructional strategies used. Knowledge of students understanding, instructional repertoire, subject matter knowledge, and representational repertoire were the components used in their study. These components of PCK, which were identified from the literature that draws from expert teachers' practices, can still be considered from the perspectives of teachers which only need to be endorsed or substantiated by the students. As argued by Moustafa et al. (2013), students might be unable to recognize indicators of such practices in teachers, especially if it is only just measured through a survey with the students.

The current study draws on both studies and also acknowledges the fact that the effect of teachers' knowledge on students' learning might be inconsistent due to the diversity of the learners (Brophy and Good 1986; Shulman 1987; Prime and Miranda 2006). Thus, this study argues that it is important to draw from the students' perspectives of what constitutes of an effective PCK based on the identified components of PCK derived from the literature. In particular, this research considers perceptions of varied achieving abilities of students as a form to determine what sort of effective teachers knowledge required of these diverse learners. In summary, this research is directed toward assessing science teachers' PCK in Malaysia while considering the students' perspectives and adopting a quantitative research approach.

Theoretical Framework

Review of studies on students' perceptions of teachers' effectiveness (Tuan et al. 2000; Hills et al. 2005; Shadreck and Issac 2012) has revealed that students expect teachers to have strong content knowledge, effective pedagogical skills, and social competence. These characteristics and dimensions of teachers are similar to the teachers' knowledge base for effective teaching (Shulman 1987), whereby PCK is a part of the knowledge base. More importantly, PCK is different from the other knowledge bases in that it is knowledge of teaching that is domain specific; it is what teachers know about their subject matter and how to make it comprehensible to the students (Shulman 1987; De Jong 2009; Schneider and Plasman 2011). PCK's most basic constituents (De Jong 2009) are (1) knowledge of students' conceptions of specific topics including knowledge of students' difficulties in understanding these topics, (2) knowledge of instructional strategies including knowledge of representations (e.g., models, metaphors) and activities (e.g., explications, experiments) for teaching specific topics, and (3) knowledge of subject matter. These core elements of PCK for science teachers have been extensively and frequently investigated.

Several examples are summarized below in a De Jong (2009) review of basic notions of PCK. For example, Grossman (1990) expanded Shulman's (1987) definition and proposed the following four elements comprising: (1) knowledge of purposes for teaching specific topics at different grade levels, (2) knowledge of students' understanding and (mis)conceptions, (3) knowledge of the curriculum and curriculum materials available for teaching specific topics, and (4) knowledge of instructional strategies and representations for specific topics. Tamir (1991) proposed four PCK elements specified for science (laboratory) lessons as follows: (1) knowledge of students: the specific common (mis)conceptions of specific topics, and how to diagnose students' difficulties in understanding specific topics, (2) knowledge of curricula: the pre-requisite concepts needed for understanding specific topics, and how to design an inquiry-oriented laboratory lesson, (3) knowledge of instruction (teaching and management): the usual phases of (laboratory) lessons, and how to teach students to use laboratory instruments, and (4) knowledge of evaluation: the nature and composition of particular science assessment inventories, and how to evaluate manipulation of laboratory skills.

Magnusson et al. (1999) as well as Park and Oliver (2008) have made distinctions between more elements. They proposed the following five elements: (1) knowledge of purposes and goals for teaching science (at a particular grade level), (2) knowledge of the science curriculum (goals and specific curricular programs), (3) knowledge of students' understanding of specific science topics, (4) knowledge of assessment in science (relevant aspects of students' learning, ways to assess these aspects), and (5) knowledge of strategies for teaching science topics (e.g., use of representations, activities).

Clearly, there is no general accepted meaning and core elements of PCK; this view highlights the need that anyone who studies and discusses PCK should be very clear about his or her conceptualization of PCK. So far, the given examples of constituents of PCK are concerned with the teaching of specific topics. Veal and MaKinster (1999) presented a taxonomy of levels of specificity of PCK. For example, at the bottom level, there is concept PCK: knowledge of teaching and learning specific concepts (e.g., temperature). While, at the highest level, there is discipline PCK: knowledge of teaching and learning specific clusters of domains (e.g., chemistry or science). De Jong (2009) asserts that the development of PCK among students; teachers should focus on the lower levels of PCK, while experienced teachers' PCK should also include the higher levels' of PCK. In this study, the focus is on discipline PCK. It serves to gain an overall view of an effective PCK of experienced science teachers for the science domain.

In this study, the core elements of PCK, namely (1) knowledge of subject matter, (2) knowledge of concept representational or knowledge of strategies for teaching specific to topics (e.g., analogies, activities), and (3) knowledge of students' understanding (e.g., students' difficulties and misunderstanding) form the conceptual framework. This study also includes another three components that reflect discipline PCK level which are (1) knowledge of teaching strategies specific to science discipline (e.g., laboratory and demonstration), (2) knowledge of assessment in learning science, and (3) knowledge of teaching context (e.g., provide an interactive environment). These six components of PCK, derived from the conception of PCK, are used to guide in assessing students' perceptions of what is needed in science teachers that promote their learning. These components of PCK also serve as the framework for the data analysis. In particular, students of different achieving abilities would be able to indicate the components of PCK needed according to their needs. A caveat is in order here. Even though the PCK is regularly interpreted as knowledge, but as argued by Fenstermacher (1994), teacher knowledge is composed of both teachers' teaching performance and thinking of teaching. Therefore in this study, teachers' behavior—in particular, their classroom explanations, representations, and interactions with students' thinking that might affect student outcomes (Hills et al. 2005) are viewed as teacher knowledge.

Objectives and Research Questions

The main objective of this study is to investigate the students' perspectives and their expectations on the components of PCK needed to teach science effectively, and not about the PCK acquired by their teachers from the students' perspective. One of the main differences is students' achieving ability. Thus, this study aims to identify the PCK of effective science teachers from the perspective of students who have various achieving ability levels. In particular, two research questions guided this study and they were (1) what are the components of PCK contributing to students' learning from the perspectives of the students?, and (2) do students with different achieving abilities have different expectations of their teachers' PCK?

Research Methodology

Method and Sample

This study used survey method and data were collected by the use of questionnaire. The researchers chose to focus on quantitative approach to determine students' perceptions of what components of teachers' knowledge that promotes science learning because these tools are more viable to administer with large number of students. The population of this research is made up of Form Four (aged 16) students, which in Malaysia is the first grade upper secondary school and is equivalent to 10th grade and 11th year in the USA and UK, respectively. Students of 16 years of age were selected as the sample for this study on the basis that they are mature enough to provide their perceptions of an effective science teaching. At the age of 16 years, these students have learnt science at the lower secondary level for 3 years.

The sample for this study was chosen from two public secondary schools. These schools would represent a typical secondary school in Malaysia since the secondary school curriculum in Malaysia is centralized and being used in all public secondary schools. The whole population of Form Four science stream students in both schools participated in the study. Each school has four science stream classes and there are about 35–38 students in each class. One teacher, from each school, taught the students Science. Thus, two teachers were involved in this study and both graduated with a bachelor of science in education. The teaching experiences of these teachers were more than 5 years.

Table 1 describes the sample involved in this survey study. A total of 316 respondents were involved and the sample comprised 42.4 % of male respondents and 57.6 % of females. The achieving ability of the students in science is categorized into three groups (high, moderate, and low) achieving ability as shown in Table 1. The achieving ability was determined by the science teachers who taught them Science and they were instructed to identify the ability based on the students' examination results.

Table 1 Background of sample ($N = 316$)

Variable	Subvariable	Frequency (percentage in parentheses)
Gender	Male	134 (42.4)
	Female	182 (57.6)
Achieving ability in science	Low	106 (33.5)
	Moderate	107 (33.9)
	High	103 (32.6)

Instrument

The survey instrument was adapted from the questionnaire developed by Tuan et al. (2000) which consisted of (1) knowledge of subject matter, (2) knowledge of teaching strategies specific to science discipline, (3) knowledge of concept representation or knowledge of strategies for teaching specific topics (e.g., analogies, activities), and (4) knowledge of assessment in science learning. Tuan et al.'s instrument was on SPOTK in relation to their pedagogy and consisted of features of teachers' knowledge from the literature related to instruction, representation, subject matter knowledge, and knowledge of how to assess students' understanding. These components of teacher knowledge are components of PCK identified in the theoretical framework.

However, Tuan et al.'s instrument does not consider two components of PCK as mentioned in the theoretical framework of the study, namely (1) knowledge of teaching context and (2) knowledge of students' understanding. Knowledge of students' understanding is one of the basic constituents of the conceptualization of PCK. All six components possessed relatively high Cronbach's alpha reliability coefficients with each of the constructs having values ranging from 0.62 to 0.76. Table 2 highlights a sample of items for each category of knowledge and the Cronbach's value for each of PCK component.

Procedure of the Survey

The questionnaire was administered to the students and they were required to respond to 56 items using a five-point Likert-type scale (from 1—for 'very unimportant' to 5—for 'very important'). The students were briefed on how to answer the questionnaire—they were asked to provide their opinions based on their perceptions of what is required of a science teacher that would facilitate them to learn science effectively. Thus, the Likert scale items (as shown in Table 2) were worded in the form that asked the students to rate each knowledge statement that they considered important for a science teacher to have or do in order to promote effective science learning. One open-ended question was given at the end of the Likert questionnaire which asked the respondents to write down other characteristics of teachers or teacher knowledge that would help them to learn science effectively.

Statistical Analysis

This study employed both descriptive and inferential statistical analyses. The former is used to describe the importance of categories of teacher knowledge for each category of achieving ability of students. The latter is then

Table 2 Reliability values for each component of PCK and examples of items

Components of PCK	No. of items	Example of an item	Scale means	Cronbach's alpha reliability coefficient
Knowledge of subject matter	10	My teacher needs to know the content he/she is teaching	3.94	0.66
		My teacher needs to know how science is related to technology		
Knowledge of teaching strategies	13	My teacher's teaching methods should keep me interested in science	4.10	0.62
		My teacher should use a variety of teaching approaches to teach different topics		
Knowledge of concept representation	11	My teacher needs to use appropriate diagrams and graphs to explain science concepts	3.71	0.68
		My teacher should use analogies with which I am familiar to help me understand science concepts		
Knowledge of teaching context	7	My teacher must create a conducive environment for learning science	4.17	0.64
		My teacher needs to pay attention to students' reaction during class and adjust his/her teaching approach		
Knowledge of students' understandings	9	My teacher must realize students' prior knowledge before class	4.08	0.65
		My teacher must know students' learning difficulties of subject before class		
Knowledge of assessment in learning science	6	My teacher's tests should allow me to check my understanding of concepts	4.30	0.76
		My teacher needs to use different approaches (questions, discussion, etc.) to find out whether I understand		

used to determine whether there exist differences in the importance of teacher's knowledge according to the types of achieving ability of the students. The analysis involved is mainly analysis of variance (ANOVA). The findings were used to explain the perspectives of students' of varying achieving levels of an effective science teachers' PCK.

For the one open-ended item, content analysis was performed. Various categories of responses were formed based on the keywords given in the responses. An inter-rater reliability check on the categories formed was done between three researchers. A sample of responses ($n = 30$) were selected and checked by the three researchers on the suitability of categories of each response. If there is a discrepancy in identifying a suitable category for the response, a discussion is conducted among the researchers to agree on a consensus. An agreement of 70 % was reached for the categories identified.

Findings and Discussions

We believe that the science teaching quality affects students' answers to the PCK questions. If the students have had different science teachers, they would have understood and experienced the different quality of science teaching. Table 3 presents the descriptive statistics of students' responses to the questionnaire. Overall, the mean values for all six components of PCK studied were above 4.00, indicating every group of the respondents considered that all the components of teacher knowledge are important in enhancing their science learning. Among the PCK components, the knowledge of assessment of learning demonstrated the highest mean value ($M = 4.45$, SD = 0.48). In particular, students indicated that their teachers' assignments should facilitate their understanding of the subject ($M = 4.53$, SD = 0.65). This finding reflects the common notion of learning and teaching science in Malaysia as being that students learn to perform well in examinations. Thus, it is important that students feel they need to understand the subject matter.

On the other hand, knowledge of concept representation which refers to teachers' knowledge in using various means of representations is considered to be the least important to have according to the students' perspectives compared to the other components of PCK; this component had the lowest mean value ($M = 4.23$, SD = 0.60). However, among the 11 items relating to teachers' knowledge of concept representation, students indicated a high need for teachers to use appropriate examples to explain the concepts clearly ($M = 4.63$, SD = 0.55). This suggests that teachers were required to transform the content or abstract

Table 3 Importance of components of PCK as perceived by students

Component	Student's achieving ability	*N*	Mean	SD
Knowledge of subject matter	High	103	4.48	0.40
	Moderate	107	4.37	0.42
	Low	106	4.30	0.48
	Total	316	4.38	0.48
Knowledge of teaching strategies	High	103	4.53	0.56
	Moderate	107	4.46	0.59
	Low	106	4.31	0.40
	Total	316	4.43	0.52
Knowledge of concept representation	High	103	4.28	0.49
	Moderate	106	4.33	0.79
	Low	106	4.09	0.51
	Total	316	4.23	0.60
Knowledge of teaching context	High	103	4.46	0.46
	Moderate	106	4.43	0.51
	Low	106	4.37	0.44
	Total	316	4.42	0.47
Knowledge of students' understanding	High	103	4.43	0.41
	Moderate	107	4.44	0.43
	Low	106	4.30	0.46
	Total	316	4.39	0.43
Knowledge of assessment in learning science	High	103	4.52	0.42
	Moderate	107	4.51	0.48
	Low	106	4.32	0.54
	Total	316	4.45	0.48

concepts so that the concepts are comprehensible and accessible to the students (Shulman 1987).

Overall, the high-achieving ability students had high expectations of their teachers' knowledge. The high-achieving groups had high expectations on two of six of the PCK components required of a science teacher. The two components were knowledge of teaching strategies and knowledge of assessment in science learning. The other two groups of students also had high expectations of their science teachers, but overall their demand seems to be less than the high-achieving group.

The moderate-achieving group indicated a high mean value for knowledge of students' understanding and knowledge of concept representational. It appears that teachers need to think about students' difficulties and to focus on teaching strategies that enable the students to comprehend the content of science through various modes of concept representational. The needs espoused by the students serve as a way to develop effective PCK that may bring an impact on students' learning (Park and Oliver 2008; Schneider and Plasman 2011).

It was found that students from the low-achieving group rated all components of PCK lower than that of other groups of students. In certain cases, studies have shown that lower achieving students are perceived by teachers to demonstrate negative attitudes in learning, which in turn shapes teachers expectations of the students' ability to learn. Low teacher expectations have been shown to reduce the motivation of students to learn (Masters 2011; Bohlmann and Weinstein 2013). Thus, it can be assumed that students do not expect innovative instructions from their teachers. Nevertheless, the low-achieving students in this study still require their teachers to demonstrate high level of competency in all the components of PCK.

As shown in Table 4, overall, there was a statistically significant difference at the $p < 0.05$ level for PCK required by students of their science teachers [$F(2, 313) = 1.73$, $p = 0.004$]. Despite reaching statistical significance, the actual difference in mean scores between the groups is quite small (4.44, 4.42, and 4.38). The effect size, calculated using ETA squared, was 0.04 which is small, while post hoc comparisons using the Tukey's HSD test indicates that the mean score for low-ability group ($M = 4.27$, SD = 0.38) is significantly different from the high-ability group ($M = 4.44$, SD = 0.37) and moderate-achieving group ($M = 4.42$, SD = 0.42). There was no significant difference between the high- and moderate-achieving groups.

One-way ANOVA on each PCK component suggests that there is significant differences in five of the six PCK components, which were knowledge of subject matter [$F(2, 313) = 1.73$, $p = 0.013$], knowledge of teaching strategies [$F(2, 313) = 1.73$, $p = 0.010$], knowledge of concept representational [$F(2, 313) = 1.73$, $p = 0.012$], knowledge of students [$F(2, 313) = 1.73$, $p = 0.038$], and knowledge of assessment in learning science [$F(2, 313) = 1.73$, $p = 0.005$]. A follow-up post hoc Tukey's HSD test revealed significant differences between groups of students for each component. As for knowledge of subject matter and knowledge of teaching strategies, significant differences seem to appear between high- and low-achieving groups only ($p = 0.010$, $p = 0.008$). The perception on the knowledge of concept representation differed significantly between moderate-and low-achieving groups ($p = 0.014$). In addition, for knowledge of students, there was a significant difference between moderate- and low-achieving group ($p = 0.050$). For the knowledge of assessment, there was a significant difference for all three groups of students, namely between the high- and low-achieving groups ($p = 0.013$), and between the moderate- and low-achieving groups ($p = 0.016$). It appears, overall, that low-achieving students constantly have low expectation on the PCK components required by science teachers compared to high- and/or moderate-achieving students.

Students' responses to the one open-ended item are displayed in Table 5. Students were asked to indicate other

Table 4 Comparison of perspectives of teachers' PCK using one-way ANOVA

Component of PCK	No. of items	Mean and standard deviation (SD)			F	Sig.	Tukey's HSD result	
		High ability	Moderate ability	Low ability			Difference between groups	Sig.
Knowledge of subject matter	10	4.48 (.40)	4.37 (.42)	4.38 (.48)	4.37	.013*	High − low	.010*
Knowledge of teaching strategies	13	4.53 (.56)	4.45 (.59)	4.43 (.40)	4.71	.010*	High − low	.008*
Knowledge of concept representational	11	4.28 (.49)	4.32 (.79)	4.23 (.51)	4.45	.012*	Moderate − low	.014*
Knowledge of teaching context	7	4.45 (.46)	4.43 (.51)	4.42 (.44)	0.84	.433	–	–
Knowledge of students	9	4.43 (.41)	4.44 (.43)	4.39 (.46)	3.30	.038*	Moderate − low	.050*
Knowledge of assessment in learning science	6	4.52 (.42)	4.51 (.48)	4.45 (.54)	5.31	.005**	High − low	.013*
							Moderate − low	.016*

* $p < 0.05$; ** $p < 0.01$

Table 5 Frequency of the influence of additional factors on effective science learning

Keywords	Frequencies and percentage (in parentheses)			
	High ($n = 103$)	Moderate ($n = 107$)	Low ($n = 106$)	Total ($n = 316$)
Experimental activities	54 (52.43)	59 (55.14)	43 (40.57)	156 (49.37)
Teachers' personality	23 (22.33)	29 (27.10)	31 (29.25)	83 (26.27)
Learning environment	27 (26.21)	26 (24.30)	19 (17.92)	72 (22.78)
Homework	22 (21.36)	21 (19.63)	18 (16.98)	61 (19.31)
Providing examples	27 (26.21)	20 (18.69)	9 (8.49)	56 (17.72)
Use of information, communication and technology	19 (18.45)	13 (12.15)	11 (10.38)	43 (13.60)
Time tabling	15 (14.56)	9 (8.41)	7 (6.60)	31 (9.82)
Promoting career in science	2 (1.94)	8 (7.48)	6 (5.66)	16 (5.06)
Medium of instruction	1 (0.97)	4 (3.74)	7 (6.60)	12 (3.81)
Teaching and learning facilities	7 (6.80)	2 (1.87)	1 (0.94)	10 (3.17)

factors expected of them of their science teachers that would help and encourage them to learn science effectively.

Based on Table 5, the most important factor that would be able to promote students' interest and contribute to the effective learning of science is providing and conducting science experiments effectively. Almost half of the total number of respondents 49.37 % ($n = 156$) were of the opinion that learning science through experiments can help to promote their interest in science, and hence increases their learning performance in that subject. All three groups of students indicated this opinion as the most important factor. This finding supports the theory of science learning in that hands-on learning activities are not only able to promote students' involvement in learning but also potentially to generate students' thinking through minds-on involvement (Arzi 2003; Hofstein et al. 2004; Ozkan et al. 2006; Halim 2009). Ates and Eryilma (2011) argue that this form of learning would avoid students mastering the scientific knowledge through recitation; instead students have the opportunity to develop knowledge through experience. In this study, respondents from all three categories of achieving group supported such a claim. As shown below, the responses by the respondents are as follows:

> Respondent 39: Do a lot of experiment, students should be encouraged to present experimental results so that students are involved minds-on, hands-on, and encourage attitude that leads to promote science learning—high-achieving ability student
> Respondent 298: Do experiment to understand science better—medium achieving ability
> Respondent 92: Always conduct experiment—low-achieving ability student

Other characteristics or factors deemed to encourage science learning included teachers' personality, ability of teacher to provide clear examples and application of knowledge, well equipped with organizing teaching and learning facilities, including use of ICT, and providing effective time tabling of lesson.

The most interesting finding in this study is related to teachers' personality and their ability to promote science as a career. Teachers' personality or personal quality was recorded as the second highest percentage given by the students. All three groups of students indicated this characteristic as important, namely students of medium and low-achieving abilities. The responses given by the students fit the concept of teachers' personality defined by Marchbanks (2000), which includes being passionate, patient, cooperative, authoritative, and creative. Respondents from this study characterized their teachers' personality as someone who is firm, confident, and who has a good physical appearance and voice control. As shown below, the responses by the respondents are as follows:

> Respondent 232: Teachers need to be authoritative and understand the needs of the students. Also teachers are to be fair and not show any form of favoritism—low-achieving ability student
> Respondent 276: Teachers are to be firm so that students will pay attention during lessons—medium achieving ability student

Approximately a quarter ($n = 83$; 26.27 %) of the total respondents felt that the quality of the teachers' personality plays an important role in teaching and learning science. Even though the focus of this study is on teachers' knowledge, however, students still placed emphasis on the personality aspect of teachers which influenced the effectiveness of their learning. As noted by Rice (2003), personal characteristics are important for a good teacher but it is not usually measured in the previous studies on effective teaching. She further argues that the focus of the study on effective teaching is on aspects of teachers' knowledge and qualifications are inevitable for those features of teacher knowledge can be translated into policy recommendations and being incorporated into teaching practice.

A study carried out on students by Spitzer (2009) clearly demonstrated that they perceived good personal characteristics were far more important than the possession of pedagogical knowledge. In addition, Marchbanks (2000) argues that teacher personality is an important factor to enable teachers to play their role to the maximum, which is to stimulate students' thinking. As indicated by Shadreck and Issac (2012), students value teachers who care and passionate about their students. Report on a study on good practices on addressing low attainment (Dunne et al. 2007) demonstrated that teachers need not only have to adopt differentiated teaching approach but also proper interpersonal skills. Teacher–pupil relations were widely regarded as highly significant to the effective learning of low-attaining pupils.

Another interesting finding is that students would like to know about the relevance of science in the everyday world, the importance of science, the application of science, and the availability of careers in science fields. According to Brodie (2006), teachers' ability to relate science in the everyday situation is important to encourage students to learn science. He further argues that if students' motivation is ignored, even the most cautious preparation and planning by the teachers will be in vain. In other words, the science teacher also needs to have another type of knowledge, which is knowledge of context. This component of PCK is one of the seven types of knowledge proposed by Shulman (1987), but is considered only to a limited extent in the discussions and studies relating to PCK (De Jong 2009).

It is important that science teachers be aware of their students' needs and also be able to address their needs. It might help to overcome low enrollment in science at the school and tertiary levels. According to Salleh et al. (2011), the target set by the Malaysian government for the ratio of Science to the Arts students of 60:40 has not been achieved. In addition, the percentage of upper secondary school students enrolling in the science subjects is decreasing from year to year. As a result, the 60:40 ratio of human resource in science and technology fields in Malaysia has still to be achieved.

In this study, 5.06 % of the respondents agreed that the effective promotion of science will increase their interest in learning science, and most of these respondents were among medium and low-achieving ability students, as shown in their responses:

> Respondent 3: Explain the basic concepts and its benefits, as well as introducing personalities who have excelled in the field—medium achieving ability students.
> Respondent 127: To visit science centers as it will provide additional knowledge about science that will enhance students' understanding—low-achieving ability students.

Another finding that is worth noting is about time tabling. The respondents, regardless of their ability, raised concern about the allocation of time for learning science in schools. A total of 9.82 % ($n = 31$) of the respondents felt that the time allocated for formal teaching and learning science during school hours was inadequate. The students further requested their teachers to hold extra lessons outside of the formal school hours. This finding is consistent with the analysis of TIMSS 2007 study which showed that Malaysia has allocated less time for teaching and learning of science than some developed countries. The lack of time allocated to learn science might hinder the students' ability to learn science effectively. As shown below, the responses by the respondents are as follows:

> Respondents 271 and 285: Teachers to conduct extra classes, activities, or experiments to increase

students' understanding—medium and low-achieving ability student, respectively.

Respondent 291: Having additional classes to improve our knowledge in science—high-achieving ability student

Efforts are being done to rectify this situation as the Malaysian Ministry of Education (2012) has recommended 15 strategies in their recent blue print in improving students' interest in science. Among the strategies was to add on extra time to the teaching of science in school. Teachers are also encouraged to provide study guides that are suitably prepared in electronic or print form. These learning resources act as students' personal tutor that is designed to assist the students with their learning and it is also seen as part of teacher's knowledge of curriculum.

Conclusion and Implications

This study investigated the students' perspectives and their expectations on the components of PCK needed to teach science effectively and not about the PCK acquired by their teachers from the students' perspective. It is noted that students' perspectives of quality science teaching would depend on the science teachers they had. Nevertheless, by analyzing the quality of PCK from perspectives of different abilities of students, it would not only gauge the quality of teachers' they encountered but also how the teachers should address their needs. It appears that as overall the teachers the students encountered have not addressed their needs.

Furthermore, there appear to be differences between types of students regarding the knowledge their teachers should have in order to help them learn science. Low-achieving students have low expectations of their teachers compared to high- and moderate-achieving students. This study suggests that science teachers should have differentiated form of PCK to facilitate science learning among diverse students. The students' viewpoints of what effective science teaching is involved can help to impact the development of effective PCK. According to Park and Oliver (2008) when science teachers are sensitize to the needs of students, only then can the teachers develop or think of improving their practice.

On the other hand, as argued by Park and Oliver (2008), the development of PCK is said to be also dependent on the quality of the students. Their study on gifted students showed that students who are good can provoke the development of an effective PCK. It is a commonly accepted view that Malaysian students tend to be passive during the teaching and learning process. Therefore, for an effective development of PCK of the teachers, it is important for the students to be given the opportunities to raise questions and involve in discussions. The lack of opportunities in doing experiments as shown in the study reinforces the passive learning environment commonly seen in the Malaysian classrooms. Perhaps, this might be also a reason for the high expectation of the students of their teachers' PCK.

The findings of this study have implications on the professional training of the teachers. Student teachers need to develop effective specific and discipline level of PCK in science. While possessing the different components of PCK is important, it is also important that teachers integrate the components as they plan and carry out teaching (Abell 2008) and the crucial factor in this development is teaching experience (De Jong et al. 2005; Loughran et al. 2004; Van Driel et al. 2002). For student teachers who have not had extensive teaching experience, the development of PCK could be enhanced by encouraging them to be reflective by conducting action research during their teaching practice. Halim et al. (2010) demonstrated that student teachers manage to develop effective PCK as an outcome of going through the cycles of action research during teaching practice. Again, it was demonstrated in that study and shown in the current study that knowledge of students' understanding had influence the student teachers' effective development of PCK.

For practicing teachers, their professional development course should also assist them to integrate the components of PCK. For both student and practising teachers, they need to be helped and sensitized to develop the necessary types of PCK, one addressing the different abilities of students and the other to help students learn. The survey instrument in this study can assists practicing science teachers to be aware of the needs of the different ability of their students.

In addition to teachers' professional knowledge, i.e., PCK, effective science teachers from the students' perspective teachers should have a good personality or social competence. Thus, pre-service and in-service courses need to provide equal emphasis on the development of teachers' knowledge and social competence. In addition, teachers also need to be made aware of their beliefs and expectations on their students so that teachers would provide relevant learning experiences appropriate to the diverse needs of their students.

Further studies need to be looked into why the low-achieving ability students have low expectations of the teachers. Another study is to investigate both the students' and teachers' perceptions of the components of PCK and to see the gap between them toward better understanding of the needs and the quality of PCK of the teachers. The components of PCK identified by students could be further validated through Delphi methodology, thus allowing measures of PCK with greater validity.

References

Abell S (2008) Twenty years later: does pedagogical content knowledge remain a useful idea? Int J Sci Educ 30(10):1405–1416

Abell SK, Lederman NG (2007) Handbook of research on science education. Lawrence Erlbaum Associates, Mahwah, NJ

Adediwura AA, Tayo B (2007) Perception of teachers' knowledge attitude and teaching skills as predictor of academic performance in Nigerian secondary school. J Educ Res Rev 2(7): 165–171

Arzi H (2003) Enhancing science education laboratory environments: more than walls, benches and widgets. In: Fraser BJ, Tobin KG (eds) International handbook of science education, vol 1. Kluwer Academic Publishers, Netherlands, pp 595–608

Ates O, Eryilma Z (2011) Effectiveness of hands-on activities on students' ability and attitudes towards physics. Asia Pac Forum Sci Learn Teach 12(1):Article 6

Bohlmann NL, Weinstein RS (2013) Classroom context, teacher expectations and cognitive level: predicting children's math ability judgments. J Appl Dev Psychol 34:288–298

Brodie M (2006) Promoting science and motivating students in the 21st century. J Sci Sch 2:38–41

Brophy JE, Good T (1986) Teacher behavior and student ability. In: Wittrock MC (ed) Handbook of research on teaching, 3rd edn. Macmillan, New York, pp 340–370

De Jong O (2009) Exploring and changing teachers' pedagogical content knowledge: An overview. In: De Jong O, Halim L (eds) Teachers' professional knowledge in science and mathematics education: views from Malaysia and abroad. Faculty of Education, The National University of Malaysia, Malaysia, pp 1–25

De Jong O, Van Driel JH, Veerloop N (2005) Preservice teachers' pedagogical content knowledge of using practical models in teaching chemistry. J Res Sci Teach 48(2):947–964

Dunne M, Humphreys S, Sebba J, Dyson A, Gallannaugh F, Muijs D (2007) Effective teaching and learning for pupils in low attaining groups. Research Report DCSF-RR011. University of Sussex, UK

Fenstermacher GD (1994) The knower and the known: the nature of knowledge in research in teaching. In: Darling-Hammond L (ed) Review of research in education. American Educational Research Association, Washington, DC, pp 3–56

Grossman PL (1990) The making of a teacher: teacher knowledge and teacher education. Teachers College Press, New York

Halim L (2009) Improving Science literacy through a conducive laboratory learning environment: a proposed model. Paper presented at international conference on science and mathematics education, Malaysia

Halim L, Meerah TSM, Buang NA (2010) Developing preservice science teachers' pedagogical content knowledge through action research. Proc-Soc Behav Sci 9:507–511

Hills HC, Rowan B, Ball DL (2005) Effects of teachers' mathematical knowledge for teaching on student ability. Am Educ Res J 42(2):371–406

Hofstein A, Shore R, Kipnis M (2004) Providing high school chemistry students with opportunities to develop learning skills in an inquiry-type laboratory: a case study. Int J Sci Educ 26(1):47–62

Ingersoll RM (1999) The problem of underqualified teachers in American secondary schools. Educ Res 28:26–37

Jang SJ (2010) Assessing college students' perceptions of a case teacher's pedagogical content knowledge using a newly developed instrument. J High Educ 61(6):663–678

Loughran J, Muhall P, Berry A (2004) In search of pedagogical knowledge in science: developing ways of articulating and documenting professional practice. J Res Sci Teach 41:370–391

Magnusson S, Krajcik T, Borko H (1999) Nature, sources and development of pedagogical content knowledge. In: Gess-Newsome J, Lederman NG (eds) Examining pedagogical content knowledge. Kluwer, Dordrecht, pp 93–132

Manizade AG, Mason MM (2011) Using Delphi methodology to design assessments of teachers' pedagogical content knowledge. Educ Stud Math 76:183–207

Marchbanks P (2000) Personality traits: elementary school teachers-prospective elementary school teachers: are they out there? http://www.unc.edu/courses/2004spring/engl/012/054/PDF%20files/Sample%20Papers/Unit%202%20Project/Unit%202–teacher.pdf. Accessed 23 Feb 2012

Masters GN (2011) The power of expectation. ACER occasional essay. Australia Council for Educational Research, Australia

Ministry of Education (2012) Strategies to improve science education in Malaysia. Ministry of Education, Putrajaya

Moustafa A, Ben-Zvi-Assraf O, Eshach H (2013) Do junior high school students perceive their learning environment as constructivist? J Sci Educ Technol 22(4):418–431

Nargund-Joshi V, Park Rogers MA, Akerson VA (2011) Exploring Indian secondary teachers' orientations and practice for teaching science in an era of reform. J Res Sci Teach 48(6): 624–647

Ozkan S, Carikoglu J, Tekkaya C (2006) Students' perceptions of the science laboratory environment. In: Sunal DW, Wright EL, Sundberg C (eds) The impact of the laboratory and technology on learning and teaching science K-16. Information Age Publishing, North Carolina, pp 111–134

Park S, Oliver JS (2008) Revisiting the conceptualization of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. Res Sci Educ 38: 261–284

Prime GM, Miranda RL (2006) Urban public high school teachers' beliefs about science learner characteristics: implication for curriculum. Urban Educ 41(5):506–532

Rice JK (2003) Teacher quality: understanding the effectiveness of teacher attributes. Economic Policy Institute, Washington, DC

Rockoff J (2004) The impact of individual teachers on student ability: evidence from panel data. Am Econ Rev 94(2):247–252

Rudduck J, Chaplain R, Wallace G (1996) School improvement: what can pupils tell us?. Fulton, London

Salleh A, Mohamad BA, Fatin APA, Salmiza S (2011) Compilation and reviews of research findings related to science and mathematics education in Malaysia. Research report for science and mathematics cluster. National Council of Professors, Putrajaya

Schneider R, Plasman K (2011) Science teacher learning progressions: a review of science teachers' pedagogical content knowledge development. Rev Educ Res 81(4):530–565

Senocak E (2009) Development of an Instrument for assessing undergraduate science students' perceptions: the problem-based learning environment inventory. J Sci Educ Technol 18(6): 560–569

Shadreck M, Issac M (2012) Science teacher quality and effectiveness: Gweru urban junior secondary school students' points of view. Asian Soc Sci 8(8):160–165

Shulman LS (1987) Knowledge and teaching: foundations of the new reform. Harv Educ Rev 57(1):1–21

Spitzer L (2009) Personality or Pedagogy: which personal characteristics are necessary for ESL teachers to possess and what role do they play compared to formal pedagogical training in ESL teaching? Stud Learn Eval Innov Dev 6(3):80–92

Tamir P (1991) Professional and personal knowledge of teachers and teacher educators. Teach Teach Educ 7:263–268

Tuan HL, Chang HP, Wang KH, Treagust DF (2000) The development of an instrument for assessing students' perceptions of teachers' knowledge. Int J Sci Educ 22(4):385–398

Van Driel JH, De Jong O, Veerloop N (2002) The development of preservice chemistry teachers' PCK. Sci Educ 86:572–590

Veal WR, MaKinster J (1999) Pedagogical content knowledge taxonomies. Electro J Sci Educ 3 www.unr.edu/homepage/crowther/ejse/vealmark.html. Accessed 20 Feb 2012

J Sci Educ Technol (2014) 23:238–251
DOI 10.1007/s10956-013-9472-6

Secondary Students' Stable and Unstable Optics Conceptions Using Contextualized Questions

Hye-Eun Chu · David F. Treagust

Published online: 1 December 2013

Abstract This study focuses on elucidating and explaining reasons for the stability of and interrelationships between students' conceptions about *Light Propagation* and *Visibility of Objects* using contextualized questions across 3 years of secondary schooling from Years 7 to 9. In a large-scale quantitative study involving 1,233 Korean students and 1,149 Singaporean students, data were analyzed from responses to the *Light Propagation Diagnostic Instrument* consisting of four pairs of items, each of which evaluated the same concept in two different problem situations. Findings show that only about 10–45 % of students could apply their conceptions of basic optics in contextualized problem situations giving rise to both stable and unstable alternative conceptions. Students' understanding of *Light Propagation* concepts compared with *Visibility of Objects* concepts was more stable in different problem situations. The concepts of *Light Propagation* and *Visibility of Objects* were only moderately correlated. School grade was not a strong predictive variable, but students' school achievement correlated strongly with their conceptual understanding in optics. The teaching and learning approach and education systems in the two countries may have had some influence on students' conceptual understanding.

H.-E. Chu (✉)
Natural Sciences and Science Education, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616, Singapore
e-mail: hyeeun.z@nie.edu.sg

D. F. Treagust
Science and Mathematics Education Centre, School of Science, Curtin University, GPO BOX U1987, Perth, WA 6845, Australia
e-mail: D.F.Treagust@curtin.edu.au

Keywords Optics concepts · Stable alternative conceptions · Unstable alternative conceptions · Contextualized questions

Introduction

Several studies have highlighted the assessment of students' alternative conceptions about optics concepts (Chang et al. 2007; Driver et al. 1994; Duit and Treagust 1998; Scott et al. 2007; Shapiro 1989). This study expands on previous research and goes a step further and focuses, in particular, on elucidating and explaining reasons for the stability or lack of stability of understanding of the related concepts and the interrelationships between students' conceptions about *Light Propagation* and *Visibility of Objects* in different problem situations across 3 years of secondary schooling in Singapore and Korea. In the past decade, the introduction of PISA has resulted in students responding to contextualized items (Fensham 2009). However, the PISA items do not assess students' learning of the same science concepts in different problem situations across several years of schooling using specially designed contextualized questions. Consequently, the main purpose of this study was to assess students' stable and unstable conceptions in optics using contextualized two-tier multiple-choice questions in different problem situations. In this study, the contextualized questions involved the application of science concepts to real-world situations that are familiar to students, often concerning a short scenario. In this case, two different problem situations are shown in pairs of items such as light propagation at night and during the daytime.

Korea and Singapore have different educational systems and national testing systems. In Korea, a national test is administered in Year 9, mostly to identify low achievers who will receive a more optimal education best suited to their level of academic attainment. Most students are assigned to a school near their home. However, if they have high enough test scores, they can apply for their preferred school (e.g., academic, technical, or commercial school) in another town. Recently, however, this Year 9 test has been discontinued in some areas in Korea. In Year 12, Korean students take their most important test that determines their admission to a university. In Singapore, students take the national test in Year 6 to progress to a differentiated secondary school education system. Depending on their test scores, students can choose their secondary school. Most students take the General Certificate of Education Ordinary level (GCE O-level) examination at the end of Year 10 and the corresponding Advanced level (A-level) examination at the end of Year 12 prior to continuing their studies at university.

Other main differences in the education system between the two countries are that the Korean science education curricula have integrated science/convergent science programs such as Science-Technology-Society-Environment (STSE) curriculum and the Science, Technology, Engineering, Art, and Mathematics (STEAM) curriculum until Year 10. On the other hand, the Singapore science education curriculum is more subject oriented with most students choosing one, two, or three science subjects such as chemistry, physics, or biology in Year 8. Secondary schools in Korea are standardized, not streamed according to academic achievement, but in Singapore students are streamed based on their national test scores in Year 6 into normal technical, normal academic, and express groups. In Korea, students speak Korean at home and are taught in the Korean language. However, in Singapore with its multilingual population, students' mother tongue may be Chinese (Mandarin), Malay, Tamil, or English, while the language of school instruction for everyone is English.

Theoretical Background

The Need for Research in Different Contexts for Students' Conceptual Understanding

For almost three decades, researchers have identified the existence of students' alternative conceptions in science education, and this research is well documented in books (e.g., Driver et al. 1985), reviews (e.g., Wandersee et al. 1994), and bibliographies (e.g., Duit 2009). Given the consistent evidence of students' alternative conceptions across age groups and across nationalities, science educators have sought to investigate the nature and process of how students change/sustain their conceptions. In the *Handbook of research on science education* (Abell and Lederman 2007), Anderson (2007) considered three traditions of research on student learning in science which he has labeled the conceptual change tradition, the sociocultural tradition, and the critical tradition. Similarly, Scott et al. (2007) summarized students' conceptions and conceptual learning in science, focusing on science learning as acquisition and participation. Interestingly, in these two major reviews of research in science learning, no mention is made of contextualized questions, while in the *Second international handbook of science education* (Fraser et al. 2012), only one chapter examines conceptual learning across contexts (King and Ritchie 2012).

Clough and Driver (1986) and Tao and Gunstone (1999) have investigated students' conceptual understanding in different problem situations. Even though the questions in the student interviews were not everyday scenarios with graphs or reports from authoritative organizations, students were required to make connections between scientific concepts and real-word situations. Clough and Driver (1986) found those students' conceptions to be contextually different but not necessarily contextually dependent. Students were required to apply scientific concepts—pressure in fluids, atmospheric pressure, conductivity, some biology concepts such as genetics and adaptation—in two to three different real-world situations, for example, pressure on goldfish at different depths in a tank of water and pressure on a submarine lying on the seabed. Students' responses to questions were varied illustrating their context-dependent conceptions when the task in different situations had perceptual dissimilarity such as considering conductivity of metal spoons in hot water and of a metal chair and a plastic chair in cold weather. On the other hand, when the tasks in the different situations were related to students' sensory experiences in the real world, such as vacuum and suction as an active pulling agent, their responses illustrated context-independent conceptions. In other words, their conceptions were stable.

Tao and Gunstone (1999) showed that students' understanding was contextually based and that they may acquire scientific conceptions in some contexts, but may retain or revert to their alternative conceptions in other contexts. Context-independent and stable conceptual change was rarely found in their research on force and motion concepts when they provided students with three different contexts during computer-supported physics instruction using a model car, a wooden box, and skydivers. After instruction, students had to show how they had applied the same or similar physics concepts—terminal speed related to balanced force—in these three different contexts. Only a few students were able to achieve stable correct conceptions by perceiving the commonalities and accepting the generality of scientific conceptions across contexts.

Recent research studies on students' conceptual learning have highlighted the relevance of students' development of representational resources in everyday contexts through negotiation with teachers and peers (Hubber et al. 2010). It would be helpful for students to view the task of learning as an investigation of individual meaning (e.g., as concept, explanation, and idea) using real-world phenomena. Tytler and his colleagues (Tytler 1998; Tytler and Peterson 2004) also indicated that students constructed their explanations inconsistently across different contexts. Tytler emphasized the valuable function of these naive inconsistent conceptions, which may help students extend their ideas when challenged with questions in different situations/contexts. In responding to these challenges, students are able to achieve a generalizable conception that is coherent and personally convincing and becomes stable overtime, so that students can actively apply the concept in new contexts.

Students' Conceptual Understanding in Fundamental Optics

The concepts of *Light Propagation* and *Visibility of Objects* were selected as the domain for the investigation of students' conceptual understanding in two different contexts using paired diagnostic test items. Phenomena involving light propagation and visibility of objects in different problem situations are prevalent in everyday life and students are consistently aware of these phenomena from an early age.

Previous research findings over the last 20 years have identified students' difficulties related to light propagation involving a variety of light resources and optical systems. Jung (1987) found that students had difficulties interpreting their experience of vision because they could not distinguish between seeing an object and receiving light from it. Ramadas and Driver (1989) describe the concept of light as a physical entity that is far removed from the concept in everyday language. The everyday concept of light is more psychological rather than physical in nature because what we call "light" in colloquial speech is mediated by a person's visual system (Anderson and Karrqvist 1981). Moreover, Galili and Hazan (2000) indicated how the optics concepts, including vision and light propagation, are difficult for students and teachers because there are obstacles in the construction of scientific knowledge about the optical phenomena; for example, the interpretation of optical phenomena based on elementary optics is far from students' direct everyday perceptions, and the process of seeing operates subconsciously even though the observer is an inherent part of optical system. Although there is a large body of research exploring students' understanding about light, vision, and optical phenomena, none of these research studies focused on students' understanding in different contexts or across educational systems, and most of research findings were based on interview data.

Diagnostic Assessment

Research has shown that items in two-tier multiple-choice instruments (Treagust 1988, 1995) are useful for analyzing students' understanding of the concepts across a wide range of topics in the secondary science curriculum. The design and development of these instruments have been used in biology (for example, diffusion and osmosis—Odom and Barrow 1995), in chemistry (for example covalent bonding—Birk and Kurtz 1999; Peterson et al. 1989), and in physics (for example, several key concepts in physics—Chang et al. 2007). These two-tier multiple-choice tests are more readily administered and scored than the other methods of ascertaining students' understanding such as interviews or Predict–Observe–Explain tasks, and thus are particularly useful for classroom teachers (Peterson et al. 1989; Tan and Treagust 1999), enabling them to use the findings of research to inform their teaching (Treagust 1995). Two-tier test items have been used by the National Science Council in Taiwan as the central part of their national assessment project (Treagust et al. 2007; Treagust and Chandrasegaran 2007).

The two-tier multiple-choice diagnostic items that were used in this study were designed based on prior research findings of students' understanding of fundamental optics; especially, students' understandings of contextualized questions were emphasized to validate the items and investigate students' understanding in different problem situations. The interconnection between context and concepts was considered in addition to students' reasoning about the phenomena in the item. The research efforts build on the work of Tao and Gunstone (1999) and Clough and Driver (1986) were discussed earlier.

Purpose of the Study

The purpose of this research was to assess whether or not contextualized diagnostic assessment items can be used to investigate the factors that influence students' conceptual understanding across different problem situations. The research questions are as follows: (1) Do students apply scientific concepts consistently in different problem situations? (2) Do students show stable alternative conceptions or unstable alternative conceptions in the two different problem situations? and (3) What are the factors that influence students' conceptual understanding of optics concepts?

Methods

Participants

The investigation was a large-scale quantitative study involving 1,233 students from three Korean schools and 1,149 students from three Singapore schools in Years 7–9 (13–15 years old). Students in Korean schools are not streamed, but students in Singaporean schools are streamed according to their achievement. The Singaporean students in school SA were from a high achieving group, and schools SB and SC were from middle and low middle achievement groups (see Table 1). Even though the Korean schools are not streamed according to academic achievement, there are reports that students showed different achievement depending on their school locations because of socioeconomic status, the school environment, and the community environment (Byun and Kim 2010; Kim 2010; Lee 1998). Consequently, the above reasons were considered in selecting schools for collecting Korean students' responses in various achievement levels on the basis of school KA from Southern Seoul, KB from a fringe area - near Seoul, and KC from Northern Seoul.

Students in both countries learn these relevant fundamental optics concepts in the LPDI questionnaire during primary school (Korea Years 3 and 6, Singapore Year 4) as well as during secondary school years (Korea year 2, Singapore Years 2 and 3). The concepts in these school years are as follows: the properties of light propagation, reflection (including image formed by plane mirror), and refraction (including image formed by lenses) (CIE 2009; MOE and HRD 2007; MOE 2004, 2007).

Questionnaire

Data were obtained at the end of 2007 in Korea and at the end of 2008 in Singapore by administering the *Light Propagation Diagnostic Instrument* (*LPDI*) consisting of eight two-tier multiple-choice items that were developed by a team of four researchers including the authors from studies reported in the research literature (Fetherstonhaugh and Treagust 1992; Langley et al. 1997; La Rosa et al. 1984; Saxena 1991; Shapiro 1989).

Table 1 Number of participants across Years 7–9

Country	School		Year 7	Year 8	Year 9
Singapore	SA	(High achieving level)	218	233	228
	SB	(Middle achieving level)	120	38	80
	SC	(Middle-low achieving level)	77	77	78
	Total		415	348	386
Korea	KA	(In southern Seoul)	139	148	77
	KB	(Near Seoul)	134	148	142
	KC	(In northern Seoul)	135	162	148
	Total		408	458	367

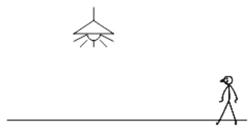
Table 2 The item situations of each pair of items in the *Light Propagation Diagnostic Instrument*

Contextualized two-tier multiple-choice diagnostic questions	
Light Propagation	Visibility of Objects
Item pair 1:	Item pair 2:
Item 1—Light propagation during the day	Item 3—Visibility of non-luminous object
Item 2—Light propagation at night	Item 4—Visibility of luminous object
Item pair 3:	Item pair 4:
Item 5—Observing lighted lampfrom window above an obstruction	Item 7—Vision of cats in complete darkness
Item 6—Observing light propagation to illuminated windows above an obstruction from the lighted lamp	Item 8—Human vision in complete darkness

Three major aspects (Treagust 2006) were considered in the process of developing the items: (1) the content was defined and represented in a concept map that accommodates the propositional statements, (2) information about students' conceptions of fundamental optics concepts on *Light Propagation* and *Visibility of Objects* was identified from the extant research studies, and (3) the two-tier multiple-choice diagnostic items were developed. Students' understanding of two concepts, *Light Propagation* and *Visibility of Object,* has been studied in the past 30 years in science education. The students' conceptions that were identified from mainly interview research studies (Fetherstonhaugh and Treagust 1992; Galili 1996; Galili and Hazan 2000; Langley et al. 1997; Saxena 1991; Shapiro 1989; Shelley 1996) were used for developing the second tier choices that are the reasons for choosing the first tier option. Face validation was conducted by two science educators to ensure that the items were included in the appropriate item groups. Also, the correct statement of each concept and the equal possibility of misconceptions being selected in each item were validated.

Each of four pairs of items investigated students' understanding of a particular concept in different situations in everyday contexts. The multiple-choice options in the first and second tiers were the same in the paired items to investigate the effects of the two different given situations. The eight items have been categorized in two concept groups: *Light Propagation* and the *Visibility of Objects* (see Table 2). Items in *Light Propagation* group asked students

Item1



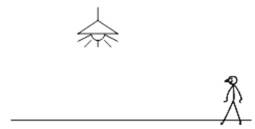
You have the light on during the day. The light from the bulb:

A. stays on the light bulb.
B. comes out about halfway towards you.
C. comes out as far as you are but no farther.
D. comes out until it hits something.

The reason I chose my answer is that:

1. light travels in all directions from the bulb.
2. light does not travel at all during the day.
3. light travels farther at night than during the day.
4. light travels about 100 to 300 m during the day.
5. light rays travel in a preferential way towards an object.

Item2



You have the light on during the night. The light from the bulb:

A. stays on the light bulb.
B. comes out about halfway towards you.
C. comes out as far as you are but no farther.
D. comes out until it hits something.

The reason I chose my answer is that:

1. light travels in all directions from the bulb.
2. light does not travel at all at night.
3. light travels farther at night than during the day.
4. light travels about 100 to 300 m at night.
5. light rays travel in a preferential way towards an object.

Fig. 1 An example of a pair of items in *LPDI*

to apply the "light propagation" concepts to two different situations, at night and during the day for the pair of items 1 and 2 (how light travels during the day/at night?) and observing light propagation from the lighted lamp to illuminated windows above an obstruction and observing the lighted lamp from the window above the obstruction for the pair of items 5 and 6 (which window can one see the lamp/which window are illuminated by the light of the lamp?). Items in *Visibility of Objects* asked students to apply the "visibility of objects" concept to two different situations, seeing a non-luminous object and a luminous object (how the boy is able to see a flower/candle flame) and the vision of cats and humans in complete darkness for the pair of items 7 and 8 (Felix, the cat/Bill the boy would see the box in the completely dark room). An example of a pair of items 1 and 2 in two different situations, light propagation during the day and at night, is shown in Fig. 1.

The diagnostic test items were translated into Korean for Korean students, and it was back translated into English to confirm the meanings of items (Brislin 1970). Also, an Australian science educator and two authors discussed the back-translated questionnaires; the translated questionnaire was considered acceptable for use by Korean students. The expressions in the questionnaire were acceptable for Singaporean students in Years 7–9, based on the comments of two science educators and two science teachers who checked the English in the questionnaire.

The Cronbach's alpha reliability value of the eight items was 0.70 in both countries (Singapore: $n = 1{,}149$, Korea: $n = 1{,}233$). According to Nunally and Bernstein (1994), in cognitive tests, a Cronbach's alpha reliability coefficient greater than 0.7 indicates a high reliability, while values in the range 0.5–0.7 indicate moderate reliability.

Data Analysis

SPSS version 18+ was used for the data analysis. The percentages of students' correct answers and distracters and also the mean scores of the conceptual categories were computed. The kappa measure of agreement was used to show the consistency of students' correct responses in the two different situations in each item pair. To investigate students' stable and unstable conceptualization in two different situations, the percentage of students' consistent incorrect conceptions in pair of items was calculated. Also, to identify factors that influenced students' conceptual understanding, one-way ANOVA was conducted for each variable in each country. Scores of students' conceptual understanding were the dependent variable while school and school years were the independent variables. Also, the country variable was considered to compare students' conceptual understanding scores.

Findings and Discussion

Item Analysis

Students' correct answers and agreement in different contexts: The paired items (1 and 2; 3 and 4; 5 and 6; 7 and 8) each involved the same optics concept in two different situations (see Table 3). The kappa measure of agreement

Table 3 The frequency of students' consistent correct answers in combined tiers, one pair, and two pairs in different problem situations in the *LPDI* (% in parentheses)

Concept categories	Item no	Frequency of students' correct answers							
		Singapore (n = 1,149)				Korea (n = 1,233)			
		Combined tiers	One pair of items		Two pair of items	Combined tiers	One pair of items		Two pair of items
Light Propagation	1	591 (51)	538 (47)	$K = 0.8$*	193 (17)	542 (44)	467 (38)	$K = 0.7$*	104 (8)
	2	601 (52)				569 (46)			
	5	620 (54)	301 (26)	$K = 0.5$*	$K = 0.2$*	302 (25)	171 (14)	$K = 0.5$*	$K = 0.2$*
	6	340 (30)				280 (23)			
Visibility of Objects	3	254 (22)	171 (15)	$K = 0.3$*	106 (9)	256 (21)	139 (11)	$K = 0.4$*	53 (4)
	4	443 (39)				303 (25)			
	7	532 (46)	506 (44)	$K = 0.3$*	$K = 0.1$*	285 (23)	261 (21)	$K = 0.3$*	$K = 0.1$*
	8	897 (78)				707 (57)			

K kappa agreement value

* $p \leq 0.05$

was calculated to show the consistency of students' responses in the two different situations in each item pair. Students' correct answers were coded as 1 and wrong answers as 0. Therefore, the kappa values indicated the students' consistency in suggesting correct and wrong answers in the paired items. A kappa value of 0.5 represents moderate agreement, while a value above 0.7 represents good agreement (Peat 2001). The paired items (1 and 2, and 5 and 6) in the *Light Propagation* concept group involved the concept that "light travels in straight lines in all directions until it strikes an object." The item pairs showed kappa values in the ranges 0.7–0.8 (Items 1 and 2) and 0.5 (Items 5 and 6). The paired items (3 and 4, and 7 and 8) in the *Visibility of Objects* concept group involved the concept that "an object is visible because light is reflected from the object to the eyes." These item pairs had kappa values of 0.3–0.4 (Items 3 and 4) and 0.3 (Items 7 and 8).

Students' percentage of correct responses to the combined tiers of each item was higher than for each item pair for all items in both concept categories (see Table 3). Also, the percentage of correct answers in conceptual categories, for example, students' consistent correct answers in items 1 and 2 and items 5 and 6, decreased dramatically ($K = 0.2$). This trend suggests that students were unable to apply the same scientific concept in combined tiers and paired items in two different contexts.

In brief, students' responses showed moderate to good agreement between the items in the paired items relating to the *Light Propagation* concept group (*K:* 0.5–0.8). However, there was lower agreement between the items in the paired items relating to the *Visibility of Objects* concept group (*K:* 0.4–0.5), suggesting that students' understanding about the concept of visibility is more highly dependent on the situations than was the case with students' understanding of light propagation.

As shown in Table 4, students' pseudo-longitudinal conceptual understanding has been investigated using the percentages of students' correct answers across the school years. Most items showed higher percentages of correct answers across the school years for the Singaporean sample. However, Korean students in Year 8 provided a higher percentage of correct answers in most items except items 7 and 8 related to visibility of objects in a completely darkroom. This difference may be due to the different curricula in the two countries. Korean students learn about fundamental optics concepts when they are in Year 8, but Singaporean students learn about fundamental optics concepts when they are in Years 8 and 9. The influence of school years is discussed again in the section on *Variables Analysis*.

Stable and Unstable Alternative Conceptions

The contextualized two-tier diagnostic items were able to identify students' stable and unstable alternative conceptions in different problem situations. From both countries, 12–25 % of students held stable alternative conceptions for the *Light Propagation* concept groups but not always for the *Visibility of Objects concept groups*. However, the consistent percentage of students' answers in two different problem situations in *the Visibility of Objects concept groups* was lower than 50 % for the stable alternative conceptions, while the consistent percentage of students' answers in two different situations in the *Light Propagation* concept groups was higher than 50 %.

Table 4 The percentages of correct answers to items in the *LPDI* across school years

Conceptual groups	Item no	Correct answer	Singaporean students (n = 1,149)				Korean students (n = 1,233)			
			Y7	Y8	Y9	Total	Y7	Y8	Y9	Total
Light Propagation	1	Light travels in straight lines in all directions from the bulb during the day	44	52	59	51	46	47	38	44
	2	Light travels in straight lines in all directions from the bulb at night	48	50	59	52	47	50	41	46
	5	Light travels in straight lines in all directions from the lamp and is received by the observer at the windows	54	47	60	54	25	25	23	25
	6	Light travels in straight lines in all directions from the lamp and lights up the windows	28	29	32	30	23	25	20	23
Visibility of Objects	3	Light is shown emanating from the non-illuminated object and being received by the eye	20	23	24	22	15	28	19	21
	4	Light is shown emanating from the illuminated object and being received by the eye	38	38	40	39	21	27	25	25
	7	Light is not shown emanating from objects. No light is reflected from the book to be received by the cat's eyes in a completely dark room	40	46	54	46	18	22	31	23
	8	Light is not shown emanating from objects. No light is reflected from the book to be received by human eyes in a completely dark room	72	82	82	78	56	57	59	57

Singaporean students: Year 7 (n = 415), Year 8 (n = 348), and Year 9 (n = 389)

Korean students: Year 7 (n = 408), Year 8 (n = 458), and Year 9 (n = 367)

Students' Alternative Conceptions About Light Propagation

Two alternative conceptions were identified that were held by 10–30 % of students in all school years from Singapore and Korea as shown in Table 5. Only more than 10 % of students' conceptions were considered as alternative conceptions in this paper as using a higher value may result in losing certain alternative conceptions (Tan et al. 2002). Furthermore, most alternative conceptions about the *Light Propagation* concept were stable, i.e., students displayed the same alternative conceptions in the two different situations in each pair of items. For example, 11–17 % of students displayed the alternative conception in Item 1 about light propagation by day, and 12–24 % of students displayed the same alternative conception in Item 2 about light propagation at night. They suggested that "the light from a bulb comes out until it hits something, because light rays travel in a preferential way toward an object." Moreover, 64–88 % of students who showed the alternative conception in item 1 had the same conceptions consistently in item 2 (see the row written in italics in Table 5).

Also, in Items 5 and 6 in the two different situations of observing a lighted lamp from a window and observing light that illuminates windows, 24–32 % of students displayed the alternative conception that "we can see all windows (or a lamp from a window) above an obstructing wall because light from the lamp is visible at all points above the obstruction" in Item 5, and 20–30 % of students displayed the same conception in Item 6. In these specific problem situations in items 5 and 6, more than half of the students, 51–60 %, who displayed the alternative conception consistently held the same conceptions in Item 6(see the row written in italics in Table5). These *Light Propagation* conceptual groups are related to students' sensory experiences with exposure to everyday phenomena. For example, when the light turns on, students experience that the light fills up the room, spreads to illuminate the space or a surface and is perceived to move in preferential ways toward the observer (Galili and Hazan 2000).

Table 5 Students' alternative conceptions about *Light propagation* in *LPDI* (% in parentheses)

Contexts	Alternative conceptions	Item (choice)	Country	School year 7	8	9	Total
Day and night	In the daytime, the light from a bulb comes out until it hits something, because light rays travel in a preferential way toward an object	Item 1 (D5)	Singapore	56 (14)	49 (14)	42 (11)	147 (13)
			Korea	59 (15)	61 (13)	61 (17)	181 (15)
	At night, the light from a bulb comes out until it hits something, because light rays travel in a preferential way toward an object	Item 2 (D5)	Singapore	49 (12)	53 (15)	14 (12)	146 (13)
			Korea	87 (21)	79 (17)	89 (24)	253 (21)
	In the day/at night, the light from a bulb comes out until it hits something, because light rays travel in a preferential way toward an object	*Items 1 and 2 (D5)*	*Singapore*	*36/56 (64)*	*39/49 (78)*	*37/42 (88)*	*112/147 (76)*
			Korea	*40/59 (68)*	*40/61 (66)*	*49/61 (80)*	*129/181 (71)*
Observing lighted lamp from window and observing light to illuminate windows	We can see a lamp from a window above an obstructing wall because light from the lamp is visible at all points above the obstruction	Item5 (A1)	Singapore	97 (24)	114 (32)	91 (24)	302 (27)
			Korea	115 (28)	137 (30)	90 (25)	342 (28)
	All windows above an obstructing wall are illuminated by the light of a lamp because light from the lamp is visible at all points above the obstruction	Item6 (A1)	Singapore	80 (20)	85 (25)	82 (22)	247 (22)
			Korea	122 (30)	122 (27)	94 (26)	337 (27)
	We can see a lamp from a window above an obstructing wall, and all windows above an obstructing wall are illuminated by the light of a lamp	*Items 5 and 6 (A1)*	*Singapore*	*53/96 (55)*	*58/114 (51)*	*50/91 (55)*	*161/302 (53)*
			Korea	*74/115 (64)*	*78/138 (57)*	*52/90 (58)*	*204/342 (60)*
	This is because light from the lamp is visible at all points above the obstruction						

Singaporean samples (n = 1,149): Year 7 (n = 415), Year 8 (n = 348), and Year 9 (n = 386)

Korean samples (n = 1,233): Year 7 (n = 408), Year 8 (n = 458), and Year 9 (n = 367)

Item choice: choice combination

Italics: frequency of students displaying consistent alternative conceptions

Students' Alternative Conceptions About the Visibility of Objects

Several alternative conceptions were held by 10–35 % of students from Korea and Singapore as shown in Table 6. Most alternative conceptions appeared in only one specific situation in each pair of items among the Korean students—identifying unstable alternative conceptions—but appeared in both situations among the Singaporean students—identifying stable alternative conceptions.

Based on the contexts in items 3 and 4, a non-illuminated object/illuminated object, Singaporean students showed the stable alternative conception that "objects are visible because of bundle of rays" without considering the observer's eyesight (Item 3: 10–13 %, Item 4: 14–23 %). However, 38–46 % of the students who showed the alternative conception in Item 3 displayed the same conceptions in Item 4. For example, a total of 137 Singaporean students showed the conception, and among them, 57 (42 %) students showed the same conception in both items (see the row written in italics in Table 6).

One reason for this alternative conception could be the wording "bundle of rays" in the second tier of options. The Singaporean students learn all subjects in English at school, but individual students also have their own mother tongue such as Chinese (Mandarin), Malay, or Tamil. During secondary school education, in science class the key scientific terms are emphasized for students to have clear understanding about the terms in English. On the other hand, Korean students showed the same alternative conception but it appeared only in item 4 (21–23 %). Many previous research studies have reported that students can consider the bundle of rays for the illuminated objects only (Langley et al. 1997; La Rosa et al. 1984).

Table 6 Students' alternative conceptions of the *Visibility of objects* in *LPDI* (% in parentheses)

Contexts	Alternative conceptions	Item (choice)	Country	School year			Total
				7	8	9	
Visibility of non-illuminated objects and illuminated objects	A non-illuminated object (e.g., flower) is visible because of bundles of rays from the object	Item 3 (A1)	Singapore	41 (10)	47 (14)	49 (13)	137 (12)
			Korea	–	–	–	–
	An illuminated object (e.g., candle flame) is visible because of bundles of rays from the object	Item 4 (A1)	Singapore	56 (14)	58 (17)	51 (13)	165 (15)
			Korea	93 (23)	107(23)	77 (21)	277 (23)
	A non-illuminated (e.g., flower)/an illuminated object (e.g., candle flame) is visible because of bundles of rays from the object	*Items 3 and 4 (A1)*	*Singapore*	*19/41 (46)*	*18/47 (38)*	*20/49 (41)*	*57/137 (42)*
			Korea	–	–	–	–
	An illuminated object (e.g., candle flame) is visible because light is present around the object	Item 4 (D4)	Singapore	41 (10)	38 (11)	68 (18)	147 (13)
			Korea	46 (11)	51 (11)	46 (13)	143 (12)
	A non-illuminated object (e.g., flower) is visible because light is present around the object (misunderstanding on the diagram of light propagation from the object)	Item 3 (D3)	Singapore	–	–	–	–
			Korea	77 (19)	97 (21)	57 (16)	231 (19)
Cat and human eyesight in completely dark room	Cats are able to see the object after adjusting their eyes to the darkness	Item 7 (B4)	Singapore	52 (13)	54 (16)	44 (12)	150 (13)
			Korea	–	–	–	–
	People are able to see the object after adjusting their eyes to the darkness	Item 8 (B4)	Singapore	56 (14)	–	40 (10)	123 (11)
			Korea	58 (14)	74 (16)	58 (16)	190 (15)
	Cats/people are able to see the object after adjusting their eyes to the darkness	*Items 7 and 8 (B4)*	*Singapore*	*21/52 (40)*	*10/54 (18)*	*19/44 (43)*	*50/150 (33)*
			Korea	–	–	–	–
	Cats see the object very clearly after adjusting their eyes to the darkness	Item 7 (C4)	Singapore	68 (17)	48 (14)	55 (14)	171 (15)
			Korea	–	–	–	–
	Cats see the object very clearly because they can see in the dark	Item 7 (C2/C3)	Singapore	68 (17)	40 (12)	45 (12)	153 (13)
			Korea	134 (33)	155 (34)	97 (26)	386 (31)

Singaporean samples ($n = 1{,}149$): Year 7 ($n = 415$), Year 8 ($n = 348$), and Year 9 ($n = 389$)

Korean samples ($n = 1{,}233$): Year 7 ($n = 408$), Year 8 ($n = 458$), and Year 9 ($n = 367$)

Item choice: Choice combination

Italics: frequency of students displaying consistent alternative conceptions

Also, there was an unstable alternative conception that "light is present around the illuminated objects" in item 4 from both countries (Singapore 10–18 %, Korea 11–13 %). Previous research (Featherstonehaugh and Treagust 1992) has shown that students' conceptions were strongly influenced by diagrams, pictures, and cartoons from their childhood and everyday life experiences. Korean students (16–21 %) showed misunderstandings of the diagram of light propagation from the objects in Item 3. Based on the context of cats' eye vision in Item 7, Singaporean students mainly showed two alternative conceptions—cats can see objects after adjusting their eyes (Option B4 13–12, Option C4 14–17 %) and cats can see clearly in a completely dark room (12–17 %). On the other hand, Korean students showed the conception that "cats can see clearly in a completely dark room" in only Item 7. Featherstonehaugh and Treagust (1992) mentioned that most students believe that cats' eyes are shiny so they have a special ability to see objects in a completely dark room. Also, they emphasized that students' beliefs about cats' eye vision were difficult to change.

In item 8 of the context about humans' eye vision, students from Korea showed the alternative conception that "people are able to see the objects after adjusting its eyes to the darkness" across Years 7–9 (14–16 %), but Singaporean students showed this alternative conception only in Years 7 and 10. Singaporean students also showed the same alternative conception in cat's eye vision across all the school years, but this was an unstable alternative conception for Korean students. However, among Singaporean students who showed the alternative conception that "cats are able to see the objects after adjusting their eyes to the darkness", less than half of them displayed the same

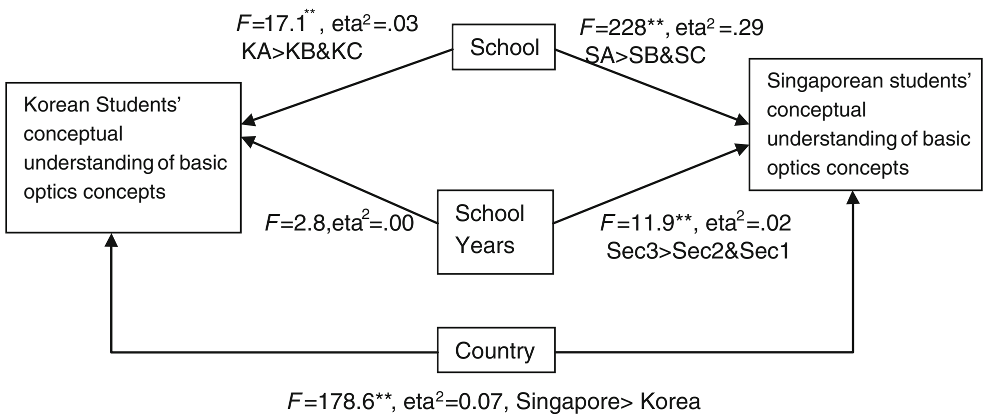


Fig. 2 Variables that influence students' conceptual understanding of basic optics concepts. $^{**}p \leq 0.005$

conception in item 8 (see the row written in italics in Table 6).

Previous research (Featherstonehaugh and Treagust 1992) has indicated that most students do not have experience in complete darkness, resulting in their belief that people can see objects in complete darkness. In this study, a similar finding was obtained with students believing that people as well as cats are able to see in complete darkness after adjusting their eyes to the darkness. The *Visibility of light* conceptual group included the two different situations in each paired item that had perceptual dissimilarities (Clough and Driver 1986) and different students' beliefs. Also, our eyes are a part of the optical system, but the light is reflected to our eyes from objects without perceptible muscular effort (Galili and Hazan 2000). These conceptual characteristics could cause the varied students' responses in the two different situations.

The Correlation between the Two Concept Groups

The Pearson correlation coefficient value of $r = 0.4$ in both countries indicates that there was a significant correlation of medium strength between the two concept groups (Cohen 1988). Reasons for the limited correlation between the two concept groups are likely to be the different characteristics of the two concepts. Both scientific concepts are far from students' everyday language, but the light propagation concept can be readily experienced by students (Galili and Hazan 2000; Ramadas and Driver 1989). On the other hand, students have limited experience of being in complete darkness as well as understanding the distinction between seeing an object and receiving light from it (Jung 1987; Galili and Hazan 2000). Also, this medium correlation supports the kappa agreement values for the correct answers in Table 3. The *Light Propagation* concept group showed higher/moderate agreement values, but the *Visibility of Objects* concept group showed low agreement values.

Table 7 Influence of school on students' understanding in Singapore and Korea

Country	School	Mean STD	F	η^2
Singapore ($n = 1{,}149$)	SA ($n = 679$)[b]	4.6 ± 1.8	228	.29**
	SB ($n = 238$)[a]	2.5 ± 1.5		
	SC ($n = 232$)[a]	2.3 ± 1.7		
Korea ($n = 1{,}233$)	KA ($n = 364$)[c]	3.0 ± 2.1	17.1	.03**
	KB ($n = 445$)[b]	2.7 ± 1.9		
	KC ($n = 424$)[a]	2.2 ± 1.8		

Different superscripts indicate that there are significant differences between year levels

$^{**}\ p \leq 0.005$, $^{*}\ p \leq 0.05$

Variables Analysis

Two variables, school and school years, in each country and country variable were considered when investigating the main factors that influenced students' conceptual difficulties in fundamental optics concepts through one-way ANOVA. The interrelations between the various factors are summarized in Fig. 2, and the detailed analysis findings were presented and discussed below.

Schools

In Table 7, among the Singaporean students in the study, the students' school was found to have significant and strong influence on their understanding ($F = 228$, $\eta^2 = 0.29$; SA > SB and SC). Students from the high achieving school (KA) showed a higher mean score (4.6 ± 1.3) in understanding of optic concepts in two different contexts than students from the medium and medium–low achieving schools (2.5 ± 1.5, 2.3 ± 1.7). Even though Korean schools are standardized, students showed

Table 8 Influence of school year on students' understanding in Singapore and Korea

Country	School year	Mean STD	F	η^2
Singapore ($n = 1{,}149$)	Year 7 ($n = 415$)	$3.4^{a} \pm 1.9$	11.9^{**}	.02
	Year 8 ($n = 348$)	$3.7^{a} \pm 2.1$		
	Year 9 ($n = 386$)	$4.1^{c} \pm 2.0$		
Korea ($n = 1{,}233$)	Year 7 ($n = 408$)	2.5 ± 1.9	2.8	.00
	Year 8 ($n = 458$)	2.8 ± 1.9		
	Year 9 ($n = 367$)	2.6 ± 2.0		

Different superscripts indicate that there are significant differences between year levels

** $p \leq 0.005$, * $p \leq 0.05$

significantly different conceptual understanding depending on the school. However, the eta-squared value is very small (KA: 3.03 ± 2.05, KB: and KC: 2.7 ± 1.9–2.2 ± 1.8, $F = 17.3$, $\eta^2 = 0.03$, KA > KB and KC).

Students from school SA the high achieving group in Singapore significantly showed the highest scores among schools, Korean school KA in Southern Seoul significantly showed the second highest scores, and the other schools were in the same score group (SA: 4.6 ± 1.8, KA: 3.0 ± 2.1, Other schools: 2.2 ± 1.8–2.7 ± 1.9, $F = 132$, $\eta^2 = 0.22$).

School Years

School years significantly influenced Singapore students' conceptual understanding ($F = 11.9$, $p < 0.001$), but the strength of the variable was small ($\eta^2 = 0.02$). Students in Year 9 in Singapore showed higher scores (4.1 ± 2.0) than those in Years 7 and 8 (Year 7 3.4 ± 1.9, Year 8 3.7 ± 2.1). There are no significant differences in students' conceptual understanding scores across school grades in Korea (see Table 8). Korean students in Year 8 showed higher scores (2.8 ± 1.9) than students in the other school years (Year 7: 2.5 ± 1.9, Year 9: 2.6 ± 2.0).

Country

The country variable was an effective variable for comparing students' conceptual understanding ($F = 178.6$, $p < 0.001$) with Singaporean students' achieving higher mean scores than Korean students on the LPDI (Singapore 3.7 ± 2.04, Korea 2.6 ± 1.9); this variable was of medium strength ($\eta^2 = 0.07$). As the school ability difference was reported in *School* variable analysis above, school SA showed the highest score in this research. One reason for this difference could be that many high achieving students from school SA in Singapore volunteered in this research.

Conclusions

This research was designed to investigate how well the developed contextualized diagnostic test items measured students' understanding in two different situations of paired items. The study also investigated the factors that may possibly influence students' fundamental optics conceptual understanding. Several conclusions were drawn based on the findings.

With respect to the first research question, "Do students apply scientific concepts consistently in different problem situations?", students' choice of responses indicated that generally they did not understand the scientific concepts fully; many students could not provide the appropriate reasons for the correct answer in the combined tier and could not apply the scientific concept in different situations within the same conceptual groups of everyday contexts.

Regarding the second research question, "'Do students show stable alternative conceptions or unstable alternative conceptions in the two different problem situations?", students from both countries held stable alternative conceptions for the *Light Propagation* concept groups but not always for the *Visibility of Objects concept groups*. Most Korean students showed unstable alternative conceptions in two different situations, but Singaporean students sometimes showed stable alternative conceptions in the *Visibility of Objects* conceptual group. Although there is little difference in curriculum of the two countries, the reason for these differences could be the language diversity in Singapore. Due to this language diversity in Singapore, the key words are emphasized during lessons to help students' scientific conceptual understanding in English, e.g., bundle of rays. Moreover, the concepts of *Light Propagation* and *Visibility of Object* were only moderately correlated; also, the *Visibility of Objects* concept group showed lower reliability compared with the *Light Propagation* concept group. One reason could be that in the *Light Propagation* concept group, students showed stable alternative conceptions, and in the *Visibility of Objects,* students showed mainly unstable alternative conceptions-even though Singaporean students showed stable alternative conceptions but the consistency of their conceptions were lower than 50 %. These findings actually support the above conclusions.

Referring to the third research question, "'What are the factors that influence students' conceptual understanding of optics concepts?", this study showed that the type of schools/location influenced students' understanding of the optics concepts involved in this research, but it was a strong variable in Singapore only where schools are streamed and where a greater number of students volunteering to respond to the LPDI were in the high ability group. The socioeconomic status, school, and community environment could have an influence on Korean students' national test scores, but the variable for this LPDI

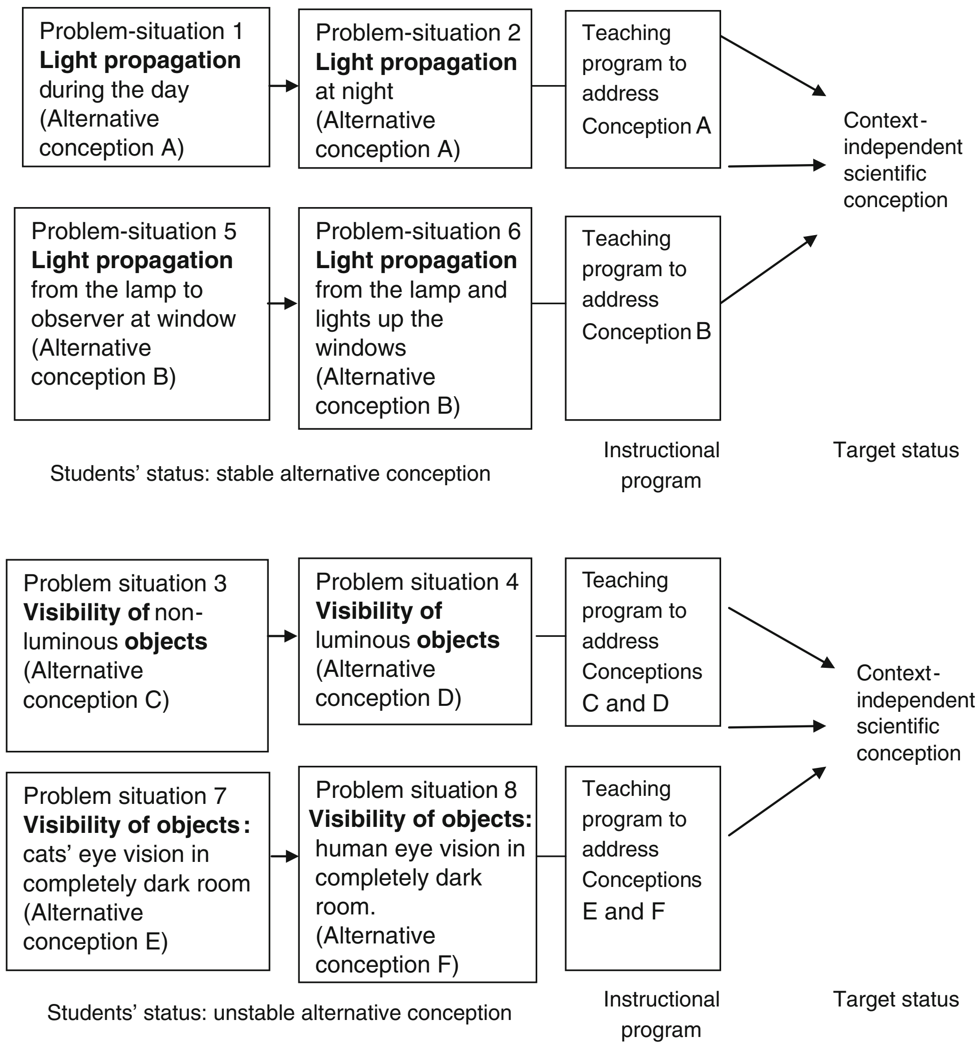


Fig. 3 Pattern of students' alternative conceptions in the contextualized LPDI

conceptual test was weak. School year was not an important variable that influenced students' conceptual understanding, thus supporting the first conclusion. Even though, because of the national testing system, students in Singapore start their review of learning science knowledge/concepts from Year 9, the school grades did not strongly influence students' conceptual understanding.

Implications

This study has several pedagogical implications. In the teaching of optics, students should be provided with opportunities to compare concepts in real-world contexts in order to facilitate their conceptual development of the fundamental underlying concepts.

First, these research findings show that some students' alternative conceptions in different situations in real-life contexts are stable. It means that the conceptions exist in two different situations, for example, during the day and at night. On the other hand, some students' alternative conceptions in different situations in real-life contexts are unstable. It means those students' alternative conceptions were influenced by the two different problem situations, for example, illuminated objects and non-illuminated objects. Consequently, students' learning should be overtly context based by providing a wide range of learning opportunities in different situations to enhance students' learning and reduce any conceptual conflict in the different situations. Experience to solve questions in different situations may help students understand contextualized scientific conceptions from which they may generalize their understanding across the different situations.

Second, previous research has shown that students have difficulties transferring their learning across different situations. Rather than simply responding to questions and determining the correct answers and reasons, students should have opportunities to think and reflect on their own understandings in different contexts. Singley and Anderson (1989) indicated that when a new context is provided for students to solve questions, they need to be reminded of earlier analogous contexts, and the commonalities across the previous and new contexts should be emphasized. It is recommended that teachers should provide a greater variety of contextualized diagnostic items that focus on different situations.

Third, it is essential that basic optics concepts are not taught only in one grade in secondary school but should be progressively developed from Years 7 to 9. Further, the teaching program could be designed based on the alternative conceptions identified from the results of administering this diagnostic instrument to enable students to conceptualize their stable and unstable alternative conceptions into scientifically acceptable conceptions (see Fig. 3). The problem situations of real-life contexts for the conceptual group of *Visibility of Object*s itself can be used to plan a teaching program for students to realize their own understanding and discuss their own answers with peers and teachers. Furthermore, the teacher might need additional visual materials to facilitate students' discussions. In this research, for the conceptual group of *Light Propagation*, the teaching program could include ICT resources and demonstration kits for students' discussions instead of having classroom discussion based only on the questions of the *Light Propagation* conceptual group.

The major contribution of this contextualized two-tier LPDI diagnostic instrument for teaching practice in science teacher education is that it can be used to identify the stable and unstable alternative conceptions in different problem situations of optics from test takers. Therefore, the results can help teachers to focus on the students' specific alternative conceptions in the classroom and prepare their teaching programs to help their students to reach a state where they can apply the scientific conceptions appropriately in a variety of everyday contexts.

References

Abell KS, Lederman GN (eds) (2007) Handbook of research on science education. Routledge, New York

Anderson CW (2007) Perspective in science learning. In: Abell SK, Lederman NG (eds) Handbook of research on science education. Lawrence Erlbaum, Mahwah, pp 3–56

Anderson B, Karrqvist C (1981) Light and its properties. In: Driver R, Squires A, Rushworth P, Wood-Robinson V (eds) Making sense of secondary science. Routledge Falmer, London, pp 41–47

Birk JP, Kurtz MJ (1999) Effect of experience on retention and elimination of misconceptions about molecular structure and bonding. J Chem Educ 76:124–128

Brislin WR (1970) Back-translation for cross-cultural research. J Cross Cultural Psychol 1(3):185–216

Byun S-Y, Kim K–K (2010) Educational inequality in South Korea: the widening socioeconomic gap in student achievement. Res Sociol Educ 17:155–182

Chang H-P, Chen J-Y, Guo C-J, Chen C–C, Chang C-Y, Tseng Y-T (2007) Investigating primary and secondary students' learning of physics concepts in Taiwan. Int J Sci Educ 29(4):465–482

Clough EE, Driver R (1986) A study of consistency in the use of students' conceptual frameworks across different task contexts. Sci Educ 70(4):473–496

Cohen J (1988) Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum

Driver R, Guesne E, Tiberghien A (1985) Children's ideas and the learning of science. In: Driver R et al (eds) Children's ideas in science. Open University Press, Philadelphia, pp 10–32

Driver R, Squires A, Rushworth P, Wood-Robinson V (1994) Making sense of secondary science: research into children's ideas. Rutledge, London

Duit R (2009) Bibliography: students' and teachers' conceptions and science education. http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html. Accessed 1 July 2012

Duit R, Treagust DF (1998) Learning in science—from behaviorism towards social constructivism and beyond. In: Fraser B, Tobin K (eds) International handbook of science education, part 1. Kluwer, Dordrecht, pp 3–25

Fensham JP (2009) Real world contexts in PISA science: implications for context-based science education. J Res Sci Teach 46(8):884–896

Fetherstonhaugh T, Treagust DF (1992) Students' understanding of light and its properties: teaching to engender conceptual change. Sci Educ 76(6):653–672

Fraser BJ, Tobin K, McRobbie CJ (eds) (2012) Second international handbook of science education. Springer, New York

Galili I (1996) Students' conceptual change in geometrical optics. Int J Sci Educ 18(7):847–868

Galili I, Hazan A (2000) Learner's knowledge in optics: interpretation, structure and analysis. Int J Sci Educ 22(1):57–88

Hubber P, Tytler R, Haskam F (2010) Teaching and learning about force with a representational focus: pedagogy and teacher changes. Res Sci Educ 40(1):5–28

Jung W (1987) Understanding students' understanding: the case of elementary optics. In: Novak JD (ed) Proceedings of second international seminar: misconceptions and educational strategies in science and mathematics, vol III. Cornell University, Ithaca, pp 268–277

Kim SS (2010) 2009 assessment analysis of academic achievement: Regional differences and school characteristics. The 5th education curriculum and assessment policy forum, ORM 2010–2014. Korean Institute of Curriculum and Evaluation (KICE), Seoul

King D, Ritchie S (2012) Learning science through real-world contexts. In: Fraser BJ, Tobin K, McRobbie CJ (eds) Second international handbook of science education. Springer, New York, pp 69–79

La Rosa C, Mayer M, Patrizi P, Vicentini-Missoni M (1984) Commonsense knowledge in optics: preliminary results of an investigation into the properties of light. Eur J Sci Educ 6(4):387–397

Langley D, Ronen M, Eylon B-S (1997) Light propagation and visual patterns: preinstruction learners' conceptions. J Res Sci Teach 34(4):399–424

Lee H-M (1998) The impact of regional differences on secondary school students' academic achievement. Ment Cultural Stud 21(3):109–136

Ministry of Education (MOE) (2004) Science syllabus—primary. Curriculum Planning and Development Division, MOE, Singapore

Ministry of Education (MOE) (2007) Science syllabus—lower secondary. Curriculum Planning and Development Division, MOE, Singapore

Ministry of Education and Human Resource Development (MOE and HRD) (2007) Science and curriculum. MOE and HRD, Seoul

Nunally JC, Bernstein IH (1994) Psychometric theory, 3rd edn. McGraw Hill, New York

Odom AL, Barrow LH (1995) Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. J Res Sci Teach 32(1):45–61

Peat J (2001) Health science research: a handbook of quantitative methods. Allen and Unwin, Sydney

Peterson RF, Treagust DF, Garnett P (1989) Development and application of a diagnostic instrument to evaluate grade 11 and 12 students' concepts of covalent bonding and structure following a course of instruction. J Res Sci Teach 26:301–314

Ramadas J, Driver R (1989) Aspects of secondary students' ideas about light. Center for Studies in Science and Mathematics Education, University of Leeds, Leeds

Saxena AB (1991) The understanding of the properties of light by students in India. Int J Sci Educ 13(3):283–289

Scott P, Asoko H, Leach J (2007) Student conceptions and conceptual learning in science. In: Abell S, Lederman N (eds) Handbook of research on science education. Erlbaum, Mahwah, pp 31–56

Shapiro BL (1989) What children bring to light: giving high status to learners' views and actions in science. Sci Educ 73(6):711–733

Shelley NJ (1996) Towards a phenomenography of light and vision. Int J Sci Educ 18(7):837–846

Singley MK, Anderson JR (1989) The transfer of cognitive skill. Harvard University Press, Cambridge

Tan DK-C, Treagust DF (1999) Evaluating students' understanding of chemical bonding. School Sci Rev 81(294):75–83

Tan KCD, Goh NK, Chia LS, Treagust DF (2002) Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. J Res Sci Teach 39(4):283–301

Tao P-K, Gunstone RF (1999) The process of conceptual change in force and motion during computer-supported physics instruction. J Res Sci Teach 36(7):859–882

Treagust DF (1995) Diagnostic assessment of students' science concepts. In: Glynn S, Duit R (eds) Learning science in the schools: research reforming practice. Lawrence Erlbaum Associates, Mahwah, pp 327–346

Treagust DF (1988) The development and use of diagnostic instruments to evaluate students' misconceptions in science. Int J Sci Educ 10:159–169

Treagust DF (2006) Diagnostic assessment in science as a means to improving teaching, learning and retention. UniServe science assessment symposium proceedings. http://science.uniserve.edu.au/pubs/procs/2006/treagust.pdf. Accessed 28 May 2013

Treagust DF, Chandrasegaran AL (2007) The Taiwan national science concept learning study in an international perspective. Int J Sci Educ 29(4):391–403

Treagust DF, Chiu M-H, Guo C-G (Eds) (2007) Taiwan National science concept learning study: a large scale assessment project using two-tier diagnostic tests. International Journal of Science Education (special issue) 29(4)

Tytler R (1998) The nature of students' informal science conceptions. Int J Sci Educ 20(8):901–927

Tytler R, Peterson S (2004) From "try it and see" to strategic exploration: characterizing young children's scientific reasoning. J Res Sci Teach 41(1):84–118

University of Cambridge International Examinations (CIE) (2009) Cambridge O level physics: syllabus code 5054. http://www.cie.org.uk/qualifications/academic/middlesec/olevel/overview. Accessed 28 May 2011

Wandersee JH, Mintzes JJ, Novak JD (1994) Research on alternative conceptions in science. In: Gabel DL (ed) Handbook of research on science teaching and learning. Macmillan, New York, pp 177–210

J Sci Educ Technol (2014) 23:252–266
DOI 10.1007/s10956-013-9454-8

On the Evolution of a Lesson: Group Preparation for Teaching Contest as Teacher Professional Development Activity for Chinese Elementary Science Teachers

Xiaowei Tang · Faxian Shao

Published online: 6 June 2013

Abstract Group preparation for teaching contest, or lesson polishing, is a teacher professional development activity unique to China. Through participant observation and discourse analysis of a typical case, this study explores how a science lesson evolved through lesson-polishing process and how such process influenced individual learning and the development of local teaching community. Our work illustrates both the values and the issues of lesson polishing as a type of teacher professional development activity. On one hand, combining professional interactions and trial lessons, lesson-polishing activity opens up space for critical yet cooperative professional interactions and tryouts of different designs and teaching strategies, providing opportunities for individual learning and development of practical rationalities within local community. On the other hand, the functions of such activities are greatly limited by the tendency of refining every detail in lesson design, the existence of overriding dispositions and authorities with overriding power, as well as the focus on practical suggestions that can be directly implemented. Suggestions for improvement are made in the final discussion.

Keywords Professional development · Lesson polishing · Practical rationalities

X. Tang (✉)
Research Center of Basic Education, College of Education, Southwest University, Beibei, Chongqing 400700, People's Republic of China
e-mail: xiaowei.tang@gmail.com

F. Shao
Chongqing Educational Science Academy, Yuzhong, Chongqing, People's Republic of China
e-mail: shaofaxian@qq.com

Professional Development Within an Alternative Culture

It is widely acknowledged that teacher professional development (PD) plays an essential role in improving teaching quality and achieving reform goals regarding student learning (Borko 2004; Darling-hammond 1997). During the past decades, consensus started to form regarding how to conduct successful PD programs. From the perspective of individual learning, accumulated evidence suggested that helping teachers develop better understandings of subject matters and guiding teachers to attend to and understand student thinking can both lead to positive changes in teaching practices (Franke et al. 2001; Smith and Neale 1991; Levin et al. 2009). From a sociocultural perspective, researchers suggested the necessity of establishing professional communities that can afford critical yet cooperative examination of teaching (Grossman et al. 2001;Lave and Wenger 1991). Difficulties have been reported in developing such community. Many argued that cultural emphasis on individualism, privacy, and autonomy obstacle teachers from challenging each other's practices and actively addressing their conflicts and differences (Bryk and Schneider 2002; Grossman et al. 2001). While teachers in the same school or district may feel comfortable enough to share teaching stories and give advice "when asked and only when asked," they usually practice in isolation and have no initiative to be part of interdependency when it comes to teaching (Little 1990).

Most studies in this area were done in American and European countries, but the lessons learnt and the issues discovered are often considered universal. As a field, we know much less about PD in countries outside the broad western culture. For example, though it has been suggested that routine PD activities in China include collaborative work on "designing curriculum, polishing lessons, observing one another's teaching, participating in study groups, and conducting research on teaching" (Darling-hammond 2005), and such routines may partially account for Chinese students' outstanding performances in international tests of student achievements (Stigler and Stevenson 1991; Huang and Bao 2006), the way such activities work remains in black box: Do Chinese teachers share similar criteria of teaching practices and therefore pursue similar goals in PD? How do the PD approaches Chinese teachers find effective work? As a country bearing the label of "collectivism," that is, where individuals are supposed to work in an interdependent way and accept the risks and rewards group work may bring, would it then be easier to build effective professional communities among teachers?

For sure such questions cannot be fully answered in any single study. It is reasonable, we think, to start with in-depth case studies on various types of PD activities popular with Chinese teachers, revealing the criteria, goals, and norms held by their communities of practice. Studies of this kind, when set in comparison with what existing literature suggested about PD in the western world, can generate a more holistic picture of the field and afford reflections on the impacts of cultural differences.

In this paper, we look into a case of "lesson polishing," a type of common Chinese PD activity in which a group of teachers work together, preparing a lesson for a teaching contest. The case is selected for its uniqueness, its popularity with teachers, and its richness in professional interactions. As far as we know, China is the only country holding regular, multi-level teaching contests and using them as PD opportunities. The tradition has been in existence for decades,[1] yet many young teachers still find it the most effective way for them to grow in teaching. In a teaching contest, each contestant teacher should perform a well-polished lesson with an unacquainted group of students in front of audience and judges. The lesson-polishing process goes through iterative cycles, each consisting of contestant teacher's trial lesson and a school-based teacher group discussion about the trial lessons. While the discussions open space for intensive professional interactions focusing on understanding, evaluating, and improving the lesson, the trial lessons provide opportunities for implementing modifications and testing out different designs. Examining such activity therefore allows insights into not only the teachers' perspectives and ideas on teaching practices, but also the process of constructing shared understanding of good teaching exemplified by a polished lesson.

In the following pages, we first introduce the theoretical lens adopted and the methods employed for collecting and analyzing data. Then, we present the case—a group of elementary science teachers worked together polishing a lesson for a province-level teaching contest. Our analysis illustrates how this lesson evolved and how the lesson-polishing process contributes to such evolution, identifying the group's shared criteria of teaching practices, summarizing the working mechanism of lesson polishing as a PD approach, and revealing the norms and values speaking to the nature of this professional community. In discussion, we comment on the advantages and issues of lesson polishing as a PD activity, reflect on the effects of cultural contexts, and make suggestions for improvements.

Community of Practice and Practical Rationality

Developments in a Community of Practice

A community of practice (COP), as Wenger (1998) defined, is any group of people bound together by a joint enterprise. While group members learn from each other, producing communal knowledge and solutions through collaboration, negotiation, and idea sharing, they also develop their own norms, relationships, and social identities. In this study, the school-based teacher group can be considered as a COP, with the long-term joint enterprise being developing and teaching to shared vision of science teaching. During the contest preparation period, it is this COP's routine activity to observe the contestant teacher's trial lessons, discuss about the observations, and share ideas for modifications. Suggestions accepted would then be implemented by the contestant teacher in later trial lessons.

Via this process, developments took place on two levels: on one hand, through professional interactions on a particular lesson, the COP negotiated their understandings of science teaching and developed an example of their shared vision; on the other hand, as the contestant teacher modified this lesson to the community's criteria, her science teaching practice also developed. The polished lesson is both a representation of the group's vision on science teaching and the contestant teacher's personal learning product.

The definition of COP would suggest development in a third dimension, that is, interactive norms and social relationships. However, it should be noted here that the school-based COP under study existed long before the contest. The participating teachers also observed and commented on

[1] We did not come across any valid source mentioning the origin of teaching contest. A teacher instructor told us in an interview that this tradition started before the Cultural Revolution when China built its educational system in alignment with that of the former Soviet Union, but this was only what he heard from an older teacher instructor.

each other's teaching on a regular basis; the school principal and the teacher instructor who sometimes joint the lesson-polishing process were also part of their regular lesson studies or teaching research activities. Through observing the lesson-polishing process, we may detect norms and relationships in function and conjecture about their origins, but our data can hardly suggest how such factors get established.

The Framework of Practical Rationality

In order to interpret the developments on both community and personal level, and to identify the influencing norms and relationships, we need to construct a thorough description of how professional interactions within the COP contributed to the evolution of the lesson. The framework of *practical rationality* (Herbst and Chazan 2002; Herbst et al. 2011) provides a way to comb through the process and a language to talk about it.

While most literature on teacher thinking and action focused on relating individual teachers' actions to their general goals, knowledge, or beliefs, Herbst and his colleagues attended to what teacher community see as reasonable or unreasonable actions in specific situations and why. Building on Bourdieu's (1998) notion of practical reason, they used the term practical rationality to denote "categories of perception and appreciation with which teachers talk about how they handle the demands of their work, and the dispositions that, as a result, observers ascribe to teachers' action (Herbst et al. 2011, p. 224)."

In an instructional situation, categories of perception refer to what members of a practice give prominence to in terms of the moments, actions, people, and objects; categories of appreciation point at the principles or values they use to judge whether an action is reasonable or not. Various dispositions regarding these two categories can be activated when different teachers were confronted with similar situations, which allow teachers to construct different practices "against the backdrop of their personal commitments and the demands of the institutional contexts where they work (Herbst and Chazan 2002, p. 2)."

When a group of teachers gather to discuss about specific practices, they have experienced together as audience—either through watching video episodes and scenario animations (as in Herbst and Colleagues' studies) or through direct classroom observations (as in our case), competing yet acceptable dispositions can be hypothesized and communicated. Such communications open up space for teachers to confirm, refine, or refute different dispositions, not only revealing but also continuously shaping the particular group's practical rationality. Herst and colleagues' studies did not go further than identifying the dispositions available to a group of teachers. As in their case, the teachers were only invited by researchers to watch the video episodes or animation and make comments; they were under no pressure to reach consensus and practice to it. In our case, however, the teaching contest created a strong momentum for the group of teachers to pursue consensus on a detailed lesson design and a way to implement such design. Shared dispositions were formed not only through persuasions and argumentations, but also through conformations and compromises, especially when power relationships were involved.

Our adoption of the analytical lens therefore goes beyond identifying different dispositions, but extend to evidencing local changes in practical rationality and exploring the mechanism underneath such dynamics. With this goal in mind, we rephrase our research questions as follows:

- What were the common dispositions regarding science teaching revealed in the process of lesson polishing? (characterizing the shared vision of science teaching).
- How did different dispositions contribute to the evolution of the lesson? (characterizing the general way lesson polishing works)
- When competing dispositions emerged, what factors influenced how the conflicts got solved? (characterizing norms and relationships pertaining to the COP)

Methodology

Data Sources and Data Collection

The case we studied took place in an urban elementary school in southwest China. Within 22 days, the contestant teacher K and her colleagues prepared a fifth-grade lesson called "who will welcome the daybreak first?" for a province-level elementary science teaching contest annually organized by local Educational Science Academy (ESA). We consider the case as a typically successful one, as the polished lesson won a first prize in the contest, indicating that this school-based group's vision of science teaching has also been approved in a much larger community.

The first author observed and videotaped all six of K's trial lessons (all with different classes) and the group discussions following each trial lesson, which serve as primary data source for the study. Besides, we conducted three semi-structured interviews with the contestant teachers, one before the contest preparation to detect her understandings of science teaching, then one during the preparation and one after the contest to detect changes. We also interviewed a teacher instructor and some other science teachers who have participated in these discussions, mainly about their understandings of science teaching and lesson-polishing activity. Triangulating data from different sources helps us interpret dispositions and disposition changes with more precision.

Analytical Methods

First, we go through the videotapes and transcripts of group discussions, identifying all kinds of comments and negotiations between comments. Then, we code the comments both by their dominating dispositions and the operational dimensions of the practical suggestions that followed. The coding process involves four steps: firstly, the two authors code the first lesson-polishing discussion together, generating in vivo codes and further clustering them into seven disposition categories and three operational dimensions; secondly, the two authors use these categories to separately code the other four discussions and identify comments that cannot be put in any of the categories; Thirdly, the two authors share thoughts about the comments they cannot code and generate two more disposition codes; finally, the two authors compare their codes and negotiate about where they code differently. The inter-rater reliability before negotiation is 83 %.

In our final codes, dispositions are divided into nine categories, five focusing on perceptions and four focusing on appreciations. Comments in perception categories are primarily derived from observations, while comments in appreciation categories are primarily based on the teachers' external understandings of principles, values, and requirements in various perspectives. Below we list these nine categories:

1. *Disciplinary knowledge (DK).* Comments attending to issues with the content knowledge involved. For example, in trial lesson 1, the contestant teacher put up signs of four directions on the four walls of classroom, and the students were confused about what it meant for the simulating earth to "turn from east to west." A *DK* comment on this situation then pointed out, since they were simulating the earth, they should looked at the direction "as if they were on a sphere," with "north up in the ceiling and south down on the floor."
2. *Student participation (SP).* Comments attending to the quantity, quality, and opportunities of student participation. For instance, in the discussion after trial lesson 2, a SP comment suggested that the students sitting outside the earth-simulating circle are not well engaged, since no clear participating role was assigned to them.
3. *Student thinking (ST).* Comments attending to the substances of particular student ideas. For instance, in the discussion after trial lesson 2, the school principal made a ST comment, emphasizing the value of a non-canonical student idea dismissed by the teacher.
4. *Lesson goal (LG).* Comments attending to whether proper learning goals have been set and pursued during the trial lesson. In the discussion after trial lesson 3, for example, the teacher instructor made a LG comment, suggesting that the lesson should be "one about simulated experiment."
5. *Classroom management (CM).* Comments attending to whether the students can well follow the teachers' instructions. In trial lesson 1, students on the central circle walked out of pace and created a mess when simulating earth's self-rotation. A CM comment addressed this issue and attributed it to that the teachers have not given clear instructions regarding how these students should move.
6. *Understandings of Inquiry (UI).* Comments based on the teachers' understandings of what inquiry is and how inquiry should be implemented. For example, the contestant teacher made a UI comment explaining why she chose to prepare this lesson for the teaching contest: "it (this lesson) emphasized evidence-based thinking and the links between lessons in this unit." In earlier interview, she suggested that evidence-based thinking was what she saw as "the core of inquiry."
7. *Student cognitive needs (CN).* Comments based on the teachers' understandings of students' cognitive development stages and derived needs. For example, in the discussion after trial lesson 1, several teachers made CN comments, stating that it is necessary to create activity for students to directly experience relative motion, since this concept is "too abstract for elementary students, who are concrete thinkers in nature."
8. *Lesson structure requirements (LSR).*Comments based on what teachers consider as principles regulating the general structure of a standard lesson. A typical LSR comment was made in the discussion after trial lesson 2: "to make the lesson complete, you need to summarize what you have learnt in the end."
9. *Contest requirements (CR).* Comments based on what teacher sees as the requirements or preferences of the contest holders. In the discussion after trial lesson 5, a typical CR comment suggested that the teacher should write the title of this lesson before starting to teach, because "they would watch for things like this in the contest."

Practical suggestions are divided into three categories according to their operational orientations. Those focusing on changing the material tools employed are classified into the category of *classroom physical settings*; those focusing on selecting and rearranging activities are classified into the category of *lesson structure*; and those focusing on modifying instructional details are classified into the category of *lesson details design*. Comments without clear expressions of dispositions or disposition-indicated context are only coded for practical suggestions; while comments without practical suggestions are only coded for dispositions.

Second, by comparing each trial lesson with the next one, we identify modifications made each round and code them into the three operational dimensions. Checking the modifications against the dispositions revealed and the practical suggestions made in previous teacher discussions allowed us to track whether certain practical suggestions were taken and whether a modification directly followed a suggestion, took root in a comment, or held no connection with the discussion. Triangulating the major themes of modifications, the dominating dispositions, and what the contestant teacher suggested about her learning in the process, we come up with a description of the group's shared vision of science teaching, which was contextualized by this particular lesson and gradually developed through the lesson-polishing process. By mapping the relationships between the comments and the modifications and seeking for patterns, we also generate some understandings regarding how different dispositions contribute to the evolution of the lesson.

Finally, we conduct close analysis on selected negotiations between competing dispositions, attending to the nature of the discrepancies and possible factors interfering with how the discrepancies got solved. Through this analysis, we reveal some norms and values contextualizing interactions in this group and inform the third research question.

The Contest, the Lesson, and the Group

As we have mentioned above, teaching contests are phenomena unique to China. The closest activity in America is the National Board Certification program, in which teachers hand in videos of their own classrooms in order to get evaluated. Those who pass the evaluation and get certified are often seen as teachers achieving certain level of excellence (according to national boards' teaching standards). Chinese teaching contests have similar significance in this perspective. It is one of those unspoken criteria that an elementary teacher has to win teaching contest prizes of certain level in order to get promoted.

In a typical science teaching contest, lesson performance is evaluated by judges (usually teacher instructors and expert teachers) based on a set of criteria attending to all sides, including lesson contents, classroom managements, student participations, teacher's language, lesson structure, and the embedded visions of scientific inquiry. Earlier contests used more specific rubrics to differentiate prize levels, but in recent years, the contest holders have tried to make it more flexible and put stronger emphasis on designing and implementing classroom inquiries.

Since each school can only enter one teacher into a province-level contest, whether a contestant teacher can get a prize is not only seen as a matter of the teacher's own ability but also as an indicator of the school's overall teaching strength in a content area. Some schools would hold school-level contest to select contestant for city-level or province-level contest, but the school in which we conduct this study is well known for its strength in science teaching, so the teachers usually take turn to enter the contest rather than competing for the right.

K, the contestant teacher, was in her 30th and her fifth year teaching elementary science. She was the youngest and least experienced among all the science teachers within the school and never participated in a teaching contest before. K chose to enter this contest with the lesson "who will welcome the daybreak first?" as she got an idea for modifying the textbook activity. The major textbook activity is a simulated experiment, in which some students form a circle simulating the earth and different places on earth, and one student outside the circle simulating the sun. The "earth" should try rotating from west to east and then from east to west, so as to see which "city" on it will see the "sun" first under different conditions. K thought replacing the small circle with a large one consisting of 24 students would allow more to participate and afford a natural connection with time zone, which is also part of the lesson content suggested by the textbook.

K and three other science teachers (L, I, F[2]) working in this school constituted the core COP working on polishing this lesson. Principal and vice principal of the school, C and U,[3] who used to be science teachers as well, each participated in two out of the five cycles. The teacher instructor (A)[4] participated only in the third cycle. Toward the end, when the lesson polishing was roughly done, the group also invited teachers from other content areas to join the discussion. L, the leader of the school science teacher

[2] L and I are both male teachers; F is a female teacher. L has been a science teacher for eight years, while I and F were both in their seventh year teaching elementary science. They all have bachelor's degrees and were all in their 30th.

[3] C and U were both male and in their 40th. Both of them had taught elementary science for more than eight years before switching to administrative positions.

[4] A was in his late 40th. He has been a teacher instructor for 10 years, before that he used to be a high school geography teacher.

department, described their way of lesson polishing as follows:

> As K's fellow teachers and supporters, we observe her trial lesson and share our ideas about the lesson with her…Together we focus on exploring how to conduct this lesson, how to better perform it. From all of our opinions, K can pick up what she agrees with, and ignore what she disagrees with. It's like brainstorming. She won't take our words as judging her practices.

It turned out to be partially true. The lesson-polishing process was indeed focusing on improving the lesson design and its implementation. But when there were conflicts between different dispositions and practical suggestions, it was not always up to K whether a comment should be taken up or not.

The Evolution of the Lesson

The first trial lesson started with a 8.5 min whole class introductory discussion, starting from earth shape, then shifting to earth motion, to earth rotation period, and finally, to the central topic of the lesson: the direction of the earth's self-rotation. Several students drew on the pre-knowledge that the sun rises in the east and sets in the west, but arrived at conflicting ideas. K solved the conflict by explaining about relative motion. In order to get further evidence, K instructed a circle consisting of 24 students (standing for 24 cities in 24 time zones) to carry out a simulated experiment, that is, to rotate "from west to east" and to observe which "city" welcomed the "sun" (a picture of the sun on the wall labeled "east") first. As part of the classroom settings, K labeled the four walls with four directions and matching pictures of daybreak, noon, dawn, and midnight. When rotating, students did not walk in pace and created a mess. Later they expressed confusions over several issues: first, according to the direction labels, both clockwise and counter-clockwise rotation can be considered as "from west to east"; second, to see the "sun," one had to face the wall labeled "east," but "would not the sun be in the west at dusk?" third, they were not clear where the "daybreak" should be. K was also confused. She briefly ended the class by concluding that cities in the east will welcome daybreak first.

In the second trial lesson, classroom physical settings, lesson structure, and instructional details all changed in significant ways. Gaining understandings of how directions should be set for this simulating experiment, K removed the labels and used a lamp to represent the sun. To further avoid confusions, she matched "clockwise" with "from east to west" and "counter-clockwise" with "from west to east" when talking about rotations and provided an explanation on the meridian plane using a small globe, making sure that the students understood where daybreak or dusk should be in this model. Accepting the suggestions on shortening the introduction and emphasizing the question in the lesson title, K reduced the introductory discussion to five minutes, and then held a discussion on "who will welcome daybreak first" as well as what need to be known to answer this question. To manage the rotation process, K instructed students sitting outside the circle to command students on circle, regulating their pace by counting from 1 to 24; she also asked students on circle to hold hands while walking. To make the simulated experiment more of an evidence-collecting process, K instructed the circle to rotate in both directions and let students reason out the direction of earth's rotation by combining their observations with data on sunrise time in Shanghai and Chongqing. The relative motion explanation was then presented as an additional piece of evidence. The lesson ended with exercises predicting differences in sunrise time through time zone counting, which, as K suggested, was to get students "apply what they have learnt."

Changes in the third trial lesson concentrated on lesson structure and classroom physical settings. To focus on the central topic, K reduced the introductory discussion to a one sentence review, and expanded the explanation of relative motion to a whole class discussion followed by an experiencing activity. To motivate participation, K assigned the role of "astronauts" to the students outside the earth-simulating circle, which came with the task of recording and reporting observations. There were also some small changes in instructional design. For example, to match the way their textbook defined rotation direction, K dropped the term "clockwise/counter-clockwise" and added in an explanation of relative direction. The hand-holding strategy was also withdrawn.

In the fourth and fifth trial lessons, more time was allocated to the simulating experiment. K instructed the students to rotate for a third time, experiencing the relative motions involved in sunrise and sunset; she also cut short the "what need to be known" discussion, saving time for a more detailed explanations regarding the simulating relationships and the goal of experiment. Besides such small adjustments in lesson structure, most modifications were made in terms of instructional design. For instance, since the group emphasized the importance of promoting student thinking by creating and catching discrepancies, K tried her best to push for conflicting ideas in their discussion on "who will welcome the daybreak first." In order to make the topic more relative to students' lives, K designed a problem-solving situation: "someone wants to organize kids around the world to celebrate children's day together, so he would like to know whether cities in the world all welcome daybreak at the same time."

The final lesson performed in the contest carried a few more modifications: instead of sitting on stools, students on the earth-simulating circle all sat on floor with a top-viewed world map in the center, which was a strategy suggested for creating a more relaxed atmosphere and inviting the students in the back to participate; the meridian plane and divisions of 24 time zones were drawn on the floor beforehand, so as to save the explanation time. When conducting the simulated experiment, the students could choose what cities they wanted to observe. By the end of this lesson, K presented a table with daybreak times around the world, so as to create cognitive conflicts that "would drive further exploration into earth motion."

"Appendix" provided a thorough account on the evolutionary route of the lesson, with all modifications coded in the three operational categories.

A Shared Vision of Science Teaching

Table 1 summarized in codes the practical suggestions and their underneath dispositions. The columns starting with "√" also showed how many of each type of suggestions got implemented in the next trial lesson.

According to the last row of the above table, about 33 % of the dispositions leading to modifications belong to the category of student cognitive needs (CN). This proportion is followed by that of lesson goal (LG) and student participation (SP) at 14 %, of classroom management (CM) at 12 %, and of lesson structure requirement (LSR) at 10 %. Dispositions in the other four categories were relatively less influential.

By matching the major dispositions in the five most influential categories with corresponding changes during the lesson evolution, we argued that the group has developed the following practical rationality:

- *Teacher should promote student thinking through creating "cognitive conflicts."* In the first trial lesson, students spontaneously brought up conflicting ideas on the direction of earth's self-rotation. In later trial lessons, K made great effort to duplicate such conflicts with other classes, since it was what the group considered as "necessary precondition for students' cognitive development" and "the momentum behind student thinking" (CN). The final touch of presenting daybreak times differing from students' predictions was crafted for similar reason.

Table 1 How lesson-polishing discussions contribute to the evolution of the lesson

Comment	Discussion 1		Discussion 2		Discussion 3		Discussion 4		Discussion 5	
	Suggested	√	Suggested	√	Suggested	√	Suggested	√	Suggested	√
Classroom physical setting	Total: 10	7	Total: 4	4	Total: 1	1	Total: 2	2	Total: 6	6
	CM: 1	0	CN: 2	2	CM: 1	1	CM: 1	1	CN: 2	2
	CM/SP: 2	2	LG: 2	2			–: 1	1	CM: 2	2
	CN: 6	4							CR: 2	2
	DK: 1	1								
Lesson structure	Total: 4	3	Total: 6	5	Total: 2	2	Total: 1	1	Total: 3	3
	CN: 1	0	CR: 1	1	LG: 1	1	LG: 1	1	C: 1	1
	LSR: 2	2	LG: 2	2	SP: 1	1			LSR: 1	1
	UI: 1	1	LSR: 2	1					CR/UI: 1	1
			UI: 1	1						
Lesson details design	Total: 6	3	Total: 6	5	Total: 9	7	Total: 1	1	Total: 6	4
	CM/SP: 1	1	CM: 1	0	CN: 3	3	CN: 1	1	CN: 4	3
	CN: 2	1	CN: 2	2	DK/CN: 1	0			LSR: 1	1
	LSR: 1	1	CR: 1	1	SP: 3	3			CM: 1	0
	DK: 1	0	LG/CN: 1	1	UI: 1	1				
	UI: 1	0	SP: 1	1	–: 1	0				
Just disposition	Total: 0	0	Total: 2	1	Total: 2	0	Total: 2	0		
			LG: 1	1	ST: 1	0	ST: 1	0		
			ST: 1	0	UI: 1	0	DK: 1	0		
Total suggested	CM: 10; CN: 23; CR: 5; DK:4; LG: 8; LSR: 7; SP: 8; ST: 3; UI: 6; ALL: 74									
Total √	CM: 7; CN: 19; CR: 5; DK:1; LG: 8; LSR: 6; SP: 8; ST: 0; UI:4 ALL: 58									

See p. 7–p. 8 for the disposition codes corresponding to the abbreviations used in the table

- *The experimental design should be clearly presented in an easy-to-follow way.* As the lesson evolved, more and more time was spent on providing detailed explanation on the design of the simulated experiment, including instructions on where the meridian plane should be, what simulated what, how to tell relative directions, and how to move when simulating earth rotation. "Students need to have a clear picture of the design before they can reason with the simulation (CN)" and "you should let students know what they are doing and keep them from walking out of pace (CM)" were both the types of dispositions behind such modifications. Additional props and settings (such as lamp, top-viewed world map, and drawings of meridian planes) were also designed to make the simulation more "concrete," since "the students at this age were concrete thinkers (CN)."
- *A well-designed lesson should follow certain structure.* With the introduction shortened and problem-solving situation crafted, time zone counting exercises arranged, and further cognitive conflicts stirred up, the lesson gradually evolved into a 40-min structured performance, with finely designed opening and ending. The LSR dispositions behind such changes include "an introduction should not be longer than 5 min, better under 3" and "it is the tradition to "draw a circle" in the end, testing how well the students learnt and get at new questions."
- *A well-designed lesson should be organized around a central activity.* As the focus was set on simulated experiment, discussions on prior knowledge of earth motion and on "what need to be known" were both cut short, and the relative motion–experiencing activity was integrated into the simulated experiment. While there are many possible ways to organize a lesson, the group stated that "the question in title needs to be laid out right at the start (LSR)," "this lesson is one about simulated experiment (LG)," and "other activities should be cut short to give prominence to the simulated experiment (LG/LSR)," showing a preference of having a core activity, with all the offshoots trimmed away.
- *Teacher should enhance the opportunities for the whole class to participate.* From the very beginning, the group shared the worries about how students outside the earth-simulating circle would have "left out feelings" (SP). The role of astronauts, the task of recording and reporting observations, was designed to eliminate such feelings and bring these kids in. Similar reason was behind stool removal strategy in the final lesson. This concern about participation was much less about the quality though. For instance, the "astronaut" was asked to record their observations by filling the blanks on their worksheets, and when reporting, they were simply reading aloud the sentence on the worksheet.

The above network of dispositions constitutes a subset of the group's shared vision of science teaching. In our interview with K, we found that what she learnt from the experience was also alignment. Before the lesson-polishing activities, she considered herself as "strong at designing lessons but weak at implementing the designs," since she could "well understand the textbook through text analysis," but when it came to practice she was "not that good at interacting with students." When we asked her what she had learnt in an interview after the third trial lesson, she gave the following response:

> As the group pointed out for me, my instructions are often not clear enough. That's a critical issue for. When the instructions do not explain well, students will be confused—what is the simulated experiment about? What's my role in it? So I added the explanation about what represent what and used several questions to probe their understandings of the design…The greatest progress, for me, is that I found a new direction to pursue: I should pay more attention to the students; I should carefully consider how to bring on cognitive conflicts and guide them to think. In the past my focus was mostly on the design itself.

Here, she suggested the importance of clearly presenting the experimental design and creating cognitive conflicts. In the interview after the contest, her comments also covered the dispositions on participation and central activity:

> The best thing about our final product, I think, is the idea of integration. We integrate all the contents in one activity—relative motion, earth rotation, simulated experiment, the idea of time zone. The students experience all these things together rather than in separate activities. The ideas they develop through such experience would also be in connection… When [the group] shared ideas, you can always get something you never thought about. Like the idea of sitting on the floor, I never thought it would have such effects. The students are relaxed both in terms of body and mind. And the students in the back are in because now they can have eye contacts with you.

Contributions from Professional Interactions: General Features

The section above illustrated the group's shared vision of science teaching, yet did not speak to how this vision got developed through professional interactions. In this section, we dug into the underlying mechanism of lesson-polishing process, focusing on the patterns regarding disposition and modification distributions as well as the connections in between.

The general features identified are as follows:

- *Fix lesson structure first.* As Table 1 shows, most practical suggestions toward lesson structure were made in the first two discussions, and the significant changes in lesson structure tended to concentrate in the first three trial lessons. This has also been confirmed as a common strategy in lesson polishing by teacher instructor A in interview:

> When you select a lesson and start to think, the first practice won't be that good, the second and the third ones will involve many changes. When you reorganize your lesson afterward, general structure and strategies will emerge. In one or two more practices, you will gradually build up feelings and experiences. That's what we call lesson polishing.

- *Increased rigidness in lesson design.* As more and more suggestions based on various dispositions were accepted and implemented, the lesson design became more and more rigid. Though the students were different in each trial lesson, the classroom conversational flow became rather predictable after the fourth trial lesson, as the formats and orders of activities, time arrangements, teaching strategies, and even the exact wording of some key instructional moves were all carefully designed and fixed. In addition, K suggested that through three to four lessons, she largely knew "what ideas could be out there" and "how to avoid or deal with the misconceptions." With her attention preoccupied by implementing a rigid lesson design and her confidence being that there would not be unexpected idea, K hardly responded to emerging student ideas with genuine curiosity. She simply got the ideas out, explained them away, and continued with what she wanted the class to pursue.
- *Opportunities for trials.* Some changes were first made but then withdrawn in later trial lessons. It was usually because the trial did not run very well or new issues were initiated. For example, the hand-holding strategy was suggested for regulating the rotation of the earth-simulating circle. But when implemented in trial lesson 2, both K and other teachers noticed that making fifth-grade boys and girls hold hands could break their social norms and shift their attention away from the scientific topic. Students on the circle took on unease looks, while the students outside the circle started to make face and play jokes. This strategy was withdrawn in the next trial lesson. We can see from such instances that lesson polishing is not a linear process, but one with opportunities to come up with and try out different designs and teaching strategies, sifting out the ones they find most effective. Such opportunities are what many teachers see as the PD value of lesson polishing. One of the participating teachers put it this way:

> When you are teaching three classes in a row, there is hardly any time to think, to reflect. Yes, I may still try to make small adjustments from class to class, but the real work will be left to the next time I teach this part, maybe 4 years later [laughing]! And by then I won't even remember the issues encounter today. …The greatest thing about lesson polishing is that it allows you to concentrate on trying things out and modifying a lesson within a short period. And it is not only you, but your fellow teachers as well. They will come up with ideas and suggestions, and you'll be like, oh, I never think this way. You can then try different things out and see what really works. I think this is the best way for young teachers to grow.

- *Focus on student cognitive needs (CN).* Comments rooted in CN dispositions made the greatest contribution to the lesson evolution. We believe that is not accidental. It was common for these elementary science teachers to base their suggestion on what is required for students to be active in cognition and what may exceed the students' cognitive abilities. For example, many expressed the worry that fifth graders, as concrete thinkers with limited spatial imagination abilities, might find it difficult to think about planetary movements. Such assumption led to the need of initiating thinking by creating cognitive conflicts, and the necessity of lowering cognitive requirements by allowing everyone to "directly experience" the earth's self-rotation through simulated experiment. These suggestions showed limited understandings of cognitive learning theories: thinking cannot be well initiated if the conflicts were simply created but not solved through argumentations and/or experimental efforts; and there were accumulating evidences that elementary students can think abstractly and draw on their previous experience as materials for knowledge constructions (Hammer 2008).
- *Influence of arbitrary requirements.* While most practical suggestions were directed at teaching and learning effects in the trial lessons, there were also suggestions made solely to accommodate established rules, such as *lesson structure requirements* (LSR) and *contest requirements* (CR). For example, the group repeatedly suggested K to cut short the opening

discussion and quickly jump to the central topic, as an introduction "should be within 5 min, better within 3." many also suggested that the lesson should have a special ending, either "showing the audience what the students have learnt" or "making connection with what the student will learn in the future," since "this is for contest, not for your regular class." Such suggestions sounded arbitrary but often get well taken.

- *Comments without practical suggestions result in few changes*. When dispositions were simply revealed in the discussion with no company of practical suggestions, there was little chance that any change in practice could take root in it. For example, in cases identified with ST dispositions, interpretations of student ideas were used to either evidence student learning or evidence that K might not well understand the ideas. On that basis, the participating teachers did suggest the necessity to better attend to the substances of student ideas, but since the ideas would not be the same the next time and there were so many executable suggestions to follow, the lesson never really evolved in this aspect.

In summary, the lesson evolved mainly through selectively accepting and implementing the group's practical suggestions. The suggestions were rooted in various dispositions. Among all the dispositions revealed, *student cognitive needs* (CN) gained most attention and made greatest contribution, while the disposition of *student thinking* (ST) resulted in almost no change, as it could hardly lead to suggestions that can be easily implemented. During the lesson-polishing process, the first few cycles focused on establishing the lesson structure, while more instructive details got worked out in later cycles. As the process went on, the rigidness of lesson design increased, leaving the teacher with less and less space to attend to the real time classroom dynamics. There were opportunities for the teachers to try out different designs and teaching strategies, making choice based on practical effects; yet, the process was also influenced by what the teachers took as preset rules, such as *lesson structure requirements* (LSR) and *contest requirements* (CR).

When Dispositions Ran into Conflicts

In this section, we look into representative episodes in which competing dispositions present, attending to the norms contextualizing the professional interactions within lesson-polishing discussions.

The Unspoken Rule About Structure

The following exchanges took place in the group discussion following K's first trial lesson:

1. L: I can accept and understand the way you conduct the instruction part, the only thing is that it is kind of long.
2. K: I know this introduction is pretty long, basically because they have not had the prior lessons, so I need to-
3. L: I think the format is Okay. It is good for the audience, because they may not know what comes before and after. And many teachers do not know how to locate a lesson in its unit. So you did a great job, showing deep understanding of the textbook. But if I were to teach this, since your introduction is quite long, I may shorten it by using "ask-and-answer." Elaborations on things like "what evidence support that the earth is round" may not be necessary.
4. K: When I taught Grade 3 the other day and asked them this, they could not answer. The earth is round, everyone knows that. What they may not know is the evidence, but we NEED to base claims on evidence, so I decided to talk about that.
5. L: You may give the points and then briefly mention the evidence yourself.

In brief, K started her first trial lesson by posing a series of four questions: "What shape is the earth?" "What causes the day–night alternation?" "How long does one rotation of the earth take?" and finally, "In what direction does the earth rotate?" As students replied to each question, K also probed their answers by asking "what evidence would support that?" or "Because—?"

The "introduction" lasted about eight minute and a half, which, as both K and L (the head of the school's science department) noticed, were "pretty long." L then suggested shortening the "introduction" by changing to "ask-and-answer" style. In response, K shared prior teaching experience and revealed her goal of emphasizing "base claims on evidence," indicating that not having the extended discussion might lead to loss in scientific learning. L further suggested what he saw as a time-saving alternative, again, indicating the need of cutting down the introduction.

This emphasis on keeping the introduction short would seem totally irrational had the reader not heard of Kairov's five-step lesson procedure or not been aware of its status as an unspoken rule in Chinese teacher communities. This widely spread structural paradigm has dominated Chinese teaching practices for decades (Hu 2002). According to it, a standard lesson can be divided into five parts: settle down

(1–2 min.); introduction (3–5 min); teaching new knowledge (30 min.); summary (5 min.); and give homework. Even though current science curriculum standards put such great emphasis on inquiry teaching and even though the teachers did agree that evidence-based thinking is at the core of inquiry, when the goal of inquiry runs into conflict with lesson structure requirements, the later would have the ball at its foot.

This is a Lesson on Simulated Experiment!

In the first group discussion, K explained why she chose this lesson:

> First, the unit was commonly considered as a hard one, so if I can teach it well, it can definitely attract eyeballs. I do have many lessons better prepared than this, but for teaching contest, you need something that can steal the limelight. Second, this unit has an overarching clue, that is, to make explanations for earth movements based on evidence. All the explanations in there are evidence-based. I made a large table, leaving the parts about earth's self-rotation and revolution blank, so we can fill those when we got there. It would then reveal a continuous chain of evidence, and helps students make the connections. And finally, this lesson is hard, but there is something great in it. It emphasized evidence-based thinking and the links between lessons in this unit.

The three reasons she shared had different disposition bases. Stemming from the disposition of contest requirements, she saw in the unit an attractive difficulty coefficient. Attending to the students' cognitive needs, she saw in the unit the chance to comb through their thoughts by linking evidences into chains. Based on her own understanding of scientific inquiry, she saw in the lesson great opportunities to foster evidence-based thinking, which she took as the core of inquiry. Her initial design, therefore, focused on getting students to look for evidence and think about where and how one may look for evidence. In the first discussion, the teacher group also nodded to this design and shared their thoughts on how to engage students in evidence-based thinking.

In the second trial lesson, K used about 10 min to discuss the following questions with students before they got into the simulated experiments: "what do we need to know in order to find out who will welcome daybreak first?" "is there anything you do not know but you want to know?" and finally, "who do you think will welcome daybreak first?" Students clearly laid out their thoughts, and one can tell from their answers that they did see evidence as crucial for their judgments.

In K's later practices, however, much shorter time was assigned to such discussion. The comments from C, the school principal, and A, the teacher instructor, should be responsible for this. According to C, the lesson was "one about simulated experiment" in which "the most important thing was for students to get clear about what represents what." A made similar but more detailed comments in the discussion following the third trial lesson, suggesting the need for students to participate in constructing the simulation by identifying the "what represent what" relationships and clarifying relative directions.

The goal of K's teaching greatly shifted from then on. In the final contest, she spent 4 min discussing about the central question, quickly summarizing the related factors, and suggested the need for a simulated experiment. Then, she used 14 min to provide a detailed explanation of the experimental design, clarifying relative directions and simulating relationships through brief interactions with students.

Reflecting on this shift, we see irrationality in the claim that the lesson was "one about simulated experiment." This lesson does provide opportunity to teach about simulation; but fostering evidence-based thinking is also an important and feasible lesson goal. No strong evidence or rationale suggested that one goal should outweigh the other. While K held the original disposition that this lesson should "emphasize evidence-based thinking" and her fellow teachers seemed to approve this in the first discussion, how come none of them made any argument against the proposal of this major shift in lesson goal?

We therefore suggest that power relationship might have a role here. It is quite possible that the teachers saw their principal and the teacher instructor as authoritative figures, choosing to follow their advices without further thinking. If that was the case, then the COP could not be considered "critical yet cooperative" in a strict sense. A COP truly valued critical thinking should build its practical rationalities on rationales rather than allowing authorities to have overriding power.

Everyone is the First to Welcome the Daybreak

The following idea came from S2, a student in K's second trial lesson, as a response to the question "who will welcome the daybreak first?"

> S2: I think everyone is the first to welcome the daybreak, because when this side faces the sun, that side is in dark, and then when that side faces the sun, this side will be in dark, and, and then it turns again. At the beginning this side is the first, and then that side is also the first. So there is no before or after, everyone is the first to welcome the daybreak.

K froze for a few second, said nothing in reply but directly turned to the next student. In the discussion that followed, L started a conversation on what S2 probably meant:

1. L: I think he is talking about circulation.
2. K: Yeah, circulation. My understanding is that he is saying if you are not facing the sun then it is night, and if you face the sun, it is daytime, but actually it is daybreak, noon, and then dusk. So I think that is because we did not go through the part on day–night alternation.
3. L: everyone welcomes the daybreak first, he is trying to say—
4. C: This kid actually goes beyond the goal of this lesson. His idea is the best part of this lesson, but you did not catch it. You know why when you ask the class about New York and Shanghai later, the class say that New York will welcome the daybreak first. This kid, he sees it right here. He realizes that it is a circulation, turning around and around. Your daybreak is my dusk, your dusk is my daybreak; then, based on what can we say, you reach daybreak before me? He sees it from the space. So his thought is beyond what this lesson requires. Listen to the kids. That is what you need to learn. There is no before and after, it is a circle.

While K saw in S2's idea a lack of deep understanding in day–night alternation, L and C suggested another possible interpretation. C claimed in line 4 that S2 "sees it (the earth) from the space" and "realized that it (earth's self-rotation) is a circulation," which was "beyond what this lesson requires." He justified this interpretation by quoting and rephrasing S2's words, revealing the underneath rationales: the earth simply turns around and around without naturally determined start and end (the students have not learnt about meridian line and dateline); from this perspective, how can we judge who welcomes the daybreak first? C also suggested that this rationale could also explain the confusion the whole class experienced later.[5]

Unlike how A and C talked about the lesson goals in the episode above, when focusing on interpreting certain student ideas, the teachers (including C) automatically justified their claims with evidence and reasoning. What K might learn here is a valuable disposition of attending to the substances of student thinking, which is quite different from the more common disposition (and also the disposition her original interpretation identified with) of attending to how well students' ideas meet the canonical knowledge. Building such disposition into practical rationalities would align with what the literature recommended as one of the most effective PD pathways: guiding teachers to attend to and understand student thinking.

However, since this disposition could not lead to suggestions more practical than learning to "listen to the students," interactions like this neither gained popularity in the lesson-polishing process nor brought significant changes to the lesson. If the teachers could spend more time on interpreting student ideas before fixing the goals or designing lesson details, more flexibility can be expected in the lesson produced and more rationales can be expected in the community's practical rationalities constructed.

Findings in this section can be summarized as follows:

- In the lesson-polishing process, traditional norms regarding lesson structure requirements and power relationships have overriding status. When practices based on other dispositions run in conflict with these norms, compromises or transformations often take place.
- When polishing a lesson, professional interactions may attend to the substances of student ideas sometimes, but since hardly any practical suggestion can be generated on that basis, such interactions often lead to no observable learning or development of practical rationality.

Discussion

Through analysis of a typical case, our study illustrates how a group of Chinese elementary science teachers modified a lesson to represent their shared vision of science teaching through multiple lesson-polishing cycles. Within this illustration, we identify the practical rationality developed that constitute their shared vision, reveal general patterns speaking to the working mechanism of lesson-polishing activities, and explore the underlying norms followed by this COP in their professional interactions.

As a PD activity, lesson-polishing combines professional interactions and trial lessons, opening up space for critical comments and detailed practical suggestions on teaching, affording negotiations as well as opportunities for experimenting with different designs and teaching strategies. Through such activities, the contestant teacher grows individually while the local teaching community further develops their practical rationality. In this sense, it does

[5] When K asked "who will welcome the daybreak earlier, Shanghai or New York?" the student representing New York was sitting right at "the daybreak," while the student representing Shanghai was 12 h apart sitting on the other side. The students were given the question without learning about the meridian line or dateline, thus many made judgments based on relative positions and earth self-rotation direction only, which led to the answer of "New York". Their judgment echoed S2's idea that "there's no before or after."

have great potential to serve as an effective PD approach, as it provides "authentic opportunities to learn from and with colleagues inside the school (Lieberman 1995, p. 591)."

The group's practical rationality consisted of a network of dispositions covering many aspects. On one hand, it stressed the importance of creating cognitive conflicts and enhancing opportunities for student participation, which aligned with the reformative ideas suggested by influential literature on conceptual change (e.g., Posner et al. 1982) and student participation (e.g., Greenwood et al. 1984; Lemke 1990). On the other hand, it also emphasized providing clear experimental design, keeping the lesson on a single track, as well as following certain requirements on lesson structure, which aligned more with the Chinese traditions of teaching, including transmitting canonical knowledge, organizing lesson in a teacher-centered way, and following Kairov's five-step lesson procedure (Hu 2002). Synthesis of these dispositions led to surficial applications of learning theories and compromises in practices. While attention was paid to conflicting ideas, the group concerned much less about how to deal with such conflicts; instead of using them as starting points for students to argue and explore, they would end the discussion there and engage students in conducting prescribed experiment. While the group made great effort creating equal opportunities for all students to participate, they did not look closely at the quality of participation (Lemke 1990).

Literature suggested individualism and autonomy as what obstacle teachers from collaborative yet critical COP (Bryk and Schneider 2002; Darling-hammond 2005; Little 1990). Our case study shows that it was indeed much easier to establish collaborative, and to some extent, critical, communities in China. In Chinese culture, teaching is widely considered a collective enterprise, so our teachers are used to commenting on each other's teaching practice and making practical suggestions (Paine and Ma 1993).

However, as reflected by our case of lesson-polishing activity, the functions of Chinese teachers' critical examinations of teaching were still limited by many factors. First, suggestions rooted in dispositions of lesson structure requirements and suggestions made by authoritative figures both have overriding powers. Their overriding status strongly constraint how a lesson could evolve, making the development of local practical rationality less rational. Second, the focus on preparing a single lesson in a short term makes it difficult for dispositions that cannot lead to immediate modifications to contribute. It is therefore hard to push for changes in more dynamic aspects of teaching practice, such as the way of attending to student thinking. Finally, the routine of fixing lesson structure in the first three cycles and refining every detail afterward led to increased rigidness in lesson design, driving practitioners away from attending to classroom dynamics.

To overcome these limitations and improve the functions of COP consisting of Chinese teachers, we need to reduce negative influences from some cultural traditions and adjust the organization of lesson-polishing activity.

It is characteristic for Chinese teachers to follow experts and authorities, as well as stick to the traditional model of lesson structure (Hu 2002; Paine and Ma 1993). To establish rational basis for professional interactions in a Chinese COP, we suggest that it is necessary to encourage individualism and autonomy in teaching practice to certain extent. Only when teachers are aware of their rights and responsibilities in making rational decisions regarding their own teaching practices, can they construct practical rationality on equal footing, reducing irrational submissions to traditional requirements and authorities.

Another cultural tradition in China is to create examples of practices through intensive preparation. In lesson-polishing activities, this product-oriented tradition leads to the tendency of fixing every detail and focusing on what can be changed in a short term only. To avoid such tendency, it is necessary to reorganize the activity of lesson polishing in the following ways. First, the time for preparing a specific lesson should be shortened, so that the polishing work can focus more on selecting proper goals and roughly outlining the activities, avoiding the loss of innate flexibility in teaching. Second, there should be regular communications between the holders and participants of a teaching contest. Issues such as whether a lesson design should follow the traditional lesson structure requirements and what teachers take as contest requirements can then be put on the table and discussed in depth. Finally, lesson polishing should extend beyond a preparation for teaching contest, but developed into a type of regular school-based teaching research activities. In such long-term efforts, there will also be more space for dispositions regarding student thinking to exert its impacts.

Acknowledgments This work is supported by a grant from the 12th Five-year National Plan on Educational Research (Grant BFA110052, "How sociocultural learning environments influence the effectiveness of classroom scientific inquiry").

Appendix

See Table 2.

Table 2 The evolutionary route of the chosen lesson

	Classroom physical settings	Lesson structure	Lesson details design
Lesson 1	Signs of four directions and matching pictures of daybreak, noon, dawn, and midnight on the four walls of classroom Sign of sun on the east wall A 24-seat circle in the middle, with 24 signs with city names under the seats	8.5 min. discussion on earth shape, earth motion, earth rotation period, and earth's self-rotation direction; K explains relative motion The class carry out simulated experiment to explore who will welcome daybreak first (turn from west to east only) Class discussion of confusions over directions Teacher draw on proofs and suggest that cities in the east will welcome daybreak first Further discussion of confusions	Teacher starts the experiment without prior instruction on relative directions, how to turn, or meridian plane (As a result, students get confused over directions, form a mess when turning, and cannot tell whether they are "welcoming the sun") Teacher selects two cities and asks students outside the circle to watch for the one welcoming the sun first
Lesson 2	Remove direction signs and pictures of noon and midnight Set a lamp to represent the sun Label cities on a flash map Set a large globe in the center of the circle Add a small globe for K to demonstrate the meridian plane	5 min. introduction Discussion on "who will welcome daybreak first" and what need to be known to answer the question Discussion on earth self-rotation direction Simulated experiment (both directions) K explains where the meridian plane is in the model K shows sunrise time data, proving that cities in the east welcomes daybreak first, and deducing out that earth rotates from west to east; K suggests relative motion as additional piece of evidence Time zone counting exercises	Students hold hands while turning Outside students command students in circle to move by counting from 1 to 24 K talks about rotation direction in terms of both clockwise/counter-clockwise and east to west/west to east
Lesson 3	Use stools of different colors to mark the meridian Label the outside students as "astronauts" Worksheets for "astronauts" to record observational results	One sentence introduction Discussion on "who will welcome daybreak first" and what need to be known to answer the question Relative motion discussion and experiencing activity K explains details of simulated experiment, relative directions and meridian plane Simulated experiment K shows sunrise time data, proving that cities in the east welcomes daybreak first Time zone counting practices and textbook reading	Withdraw hand-holding strategy Talk about rotation direction only in terms of east to west/west to east; Detailed explanation about meridian and how to turn in the simulated experiment Replace explanation of relative motion with discussion and direct experience.
Lesson 4	Use larger globe for demonstration Redesigned worksheets with more structured questions	Similar to lesson 3 except for: Remove the part of textbook reading Instead of teacher explanations, let students suggest what represents what in the simulated experiment;	K tries to get different ideas and create cognitive conflict Simplify relative motion–experiencing activity K emphasizes that the experiment is to clarify the relationship between earth rotation and the order of welcoming daybreak.
Lesson 5	Replace city name signs with hangtags Label the outside tables as spaceships	Similar to lesson 4 except for: Let students turn for a third time in simulated experiment, focusing on the relative direction of "sunrise" and "sunset"	Contextualize the question as one for determining whether kids around the world can welcome new year at the same time

Table 2 continued

	Classroom physical settings	Lesson structure	Lesson details design
Lesson 6	Remove stools from the middle circle Draw out meridian and time zones on the floor; Different worksheets for students on the circle and outside the circle	Similar to lesson 5 except for: Present global sunrise time in the end to further create conflicts Replace the relative motion discussion with a video clip	Let students read from PPT their experimental tasks K writes down the central question on blackboard at the beginning

References

Borko H (2004) Professional development and teacher learning: mapping the terrain. Educ Res 33:3–15

Bourdieu P (1998) Practical reason: on the theory of action. Stanford University Press

Bryk AS, Schneider BL (2002) Trust in schools: a core resource for improvement. Russell Sage Foundation Publications

Darling-hammond L (1997) The right to learn: a blueprint for creating schools that work. Jossey-Bass, San Francisco

Darling-Hammond L (2005) Teaching as a profession: lessons in teacher preparation and professional development. Phi delta kappan 87(3):237–240

Franke ML, Carpenter TP, Levi L, Fennema E (2001) Capturing teachers' generative change: a follow-up study of professional development in mathematics. Am Educ Res J 38:653–689

Greenwood CR, Dinwiddie G, Terry B, Wade L, Stanley SO, Thibadeau S, Delquadri JC (1984) Teacher-versus peer-mediated instruction: an ecobehavioral analysis of achievement outcomes. J Appl Behav Anal 17(4):521–538

Grossman P, Wineburg S, Woolworth S (2001) Toward a theory of teacher community. Teach Coll Rec 103:942–1012

Hammer D, Russ R, Mikeska J, Scherr R (2008) Identifying inquiry and conceptualizing students' abilities. In: Duschl RA, Grandy RE (eds) Teaching scientific inquiry: recommendations for research and implementation. Sense Publishers, Rotterdam, pp 138–156

Herbst V, Chazan D (2002) Exploring the practical rationality of mathematics teaching through conversations about videotaped episodes: the case of engaging students in proving. Learn Math 23(1):1–13

Herbst P, Nachlieli T, Chazan D (2011) Studying the practical rationality of mathematics teaching: what goes into "installing" a theorem in geometry? Cogn Instr 29(2):218–255

Hu GW (2002) English language teaching in the People's Republic of China. In: Silver RE, Hu GW, Iino M (eds) English language education in China, Japan, and Singapore. Nanyang Technological University, Singapore

Huang R, Bao J (2006) Towards a model for teacher professional development in China: introducing Keli. J Math Teach Educ 9(3):279–298

Lave J, Wenger E (1991) Situated learning: legitimate peripheral participation. Cambridge University Press, Cambridge

Lemke JL (1990) Talking science: language, learning, and values. Ablex Publishing Corporation, Norwood

Levin DM, Hammer D, Coffey JE (2009) Novice teachers' attention to student thinking. J Teach Educ 60(2):142–154

Lieberman A (1995) Practices that support teacher development: transforming conceptions of professional learning. Phi delta kappan 76(8):591–596

Little J (1990) The persistence of privacy: autonomy and initiative in teachers' professional relations. Teach Coll Rec 91(4):509–536

Paine L, Ma L (1993) Teachers working together: a dialogue on organizational and cultural perspectives of Chinese teachers. Int J Educ Res 19(8):675–697

Posner G, Strike K, Hewson P, Gertzog W (1982) Accommodation of a scientific conception: towards a theory of conceptual change. Sci Educ 66(2):211–227

Smith DC, Neale DC (1991) The construction of subject-matter knowledge in primary science teaching. In: Brophy J (ed) Advances in research on teaching. Teachers' knowledge of subject matter as it relates to their teaching practice, vol 2. JAI Press, Greenwich, pp 187–243

Stigler JW, Stevenson HW (1991) How Asian teachers polish each lesson to perfection. Am Educ 15(1):12–20

Wenger E (1998) Communities of practice: learning as a social system. Syst Think 9(5):2–3

J Sci Educ Technol (2014) 23:267–279
DOI 10.1007/s10956-013-9459-3

Development of an Instrument for Assessing the Effectiveness of Chemistry Classroom Teaching

Changlong Zheng · Lihai Fu · Peng He

Published online: 13 July 2013

Abstract Classroom teaching is a main frontier of the implementation of new curricular ideas in China. The study reported in this article is concerned with the effectiveness of system of classroom teaching (SCT) in chemistry lessons. According to the Systems Science theory, we took a macroscopic view on the SCT, arguing that SCT is a hierarchy of system, which includes class system, plate system, unit system, and primitive system. In this study, we focused on primitive system of classroom teaching (PrS)—the lowest level in a SCT. Using focus group interviews, this study investigated the variables related to the effectiveness of PrS. We found a total of 21 such variables. To identify the main factors underlying the effectiveness of PrS, we further used exploratory factor analysis and confirmatory factor analysis. We found five main factors: rational use of time, quality of teaching behavior chain, match degree, quality of using resource and technology, and rationality of primitive content. Based on these findings, we constructed an evaluation scale for assessing the effectiveness of primitive system of chemistry classroom teaching.

Keywords Chemistry classroom teaching · Primitive system · Exploratory factor analysis · Confirmatory factor analysis

C. Zheng (✉) · L. Fu · P. He
Department of Chemistry, Institute of Chemical Education, Northeast Normal University, Changchun 130024, Jilin, People's Republic of China
e-mail: zhengcl@nenu.edu.cn

L. Fu
e-mail: fulh750@nenu.edu.cn

P. He
e-mail: hep905@nenu.edu.cn

Introduction

Since the new curriculum reform in China in 2001, how to improve the effectiveness of classroom teaching has become an important research topic for education researchers (Sun 2008; Cui 2009) and front-line teachers (Song 2004; He 2007). In a review of a large amount of literature on this topic, we have found that research on effective classroom teaching in the past decade has dealt with both the content and constituent elements of effective teaching. Some researchers have put forward influential factors for effective teaching (Alton-Lee 2003; Song 2004) and effective teaching characteristics (Yao 2004). This body of previous research can be summarized into six principal areas of effective teaching: teaching behavior, teacher–student relationship, teacher quality, environment, the use of technology, and teaching assessment.

Teaching Behavior

Teaching behavior has been recognized as an important variable in science teaching and a necessary part of teaching strategies. Strong classroom management and the incorporation of effective teaching strategies create the strongest environment to improve student achievement (Lahue-McCully 2012). Studies have found that effective teaching strategies in science include participation, inquiry, cooperative learning, assessment, and feedback (Çimer 2007). Research has also found that competitive activities, cooperative activities, social activities, and off-task behaviors can influence students' achievement (Peterson and Fennema 1985). There are researches focusing on teaching strategies fostering effective learning (Smith 1995; Wang 2000).

In short, the above research studies all suggest that teaching behaviors or methods can make a significant impact on student learning outcomes. Because traditional learning models are not highly relevant to students' needs, generative instruction strategies (Jones 1995) have been proposed in order to produce stronger student engagement, such as group work, group thinking, projects, and presentations (Lahue-McCully 2012). Other student interactions, such as peer tutoring, cooperative learning, collaborative learning, and reciprocal teaching (Cameron 2002) may also create engaging learning opportunities (Gurney 2007).

Teacher–Student Relationship

A harmonious teacher–student relationship facilitates the formation of conductive learning environment, in which student can learn without pressure and enjoy the process of learning (Zhao 2002). Studies have found such factors as teacher qualities, teaching methods, and the teacher–student relationship to be influential in motivating underperforming students (Oesterle 2008). It has been reported that teacher–child closeness is positively associated with children's academic performance and school adjustment (Birch and Ladd 1997), and perceptions of a caring and supportive relationship with a teacher and a positive classroom environment are related to school satisfaction (Baker 1999). In addition, positive teacher–child relationships provide children with an emotional security (Pianta 1999) as well as the strategy for creating good relationship (Tang 2003). In summary, creating a good teacher–student relationship is significant in the development of effective teaching.

Teacher Quality

Some researches assert that teacher qualifications are consistently linked to students' achievement (Haycock 1998; Wenglinsky 2000). Some others show that the certified teachers are more effective than non-certified teachers in increasing student achievement (Darling-Hammond 2000; Goldhaber and Anthony 2007). An effective teacher must possess characteristics of a quality teacher, including professional qualities, efficiency, compassion, passion, and context (Howard 2008). Teachers should understand teacher professionalism, such as identity and self-efficiency (Davis et al. 2006). There are some other factors relating teacher quality and increasing student achievement, for example, verbal ability, content knowledge, enthusiasm for learning, and so on (Kaplan and Owings 2001). Research on students with different gender and different grade shows that teacher quality is significant to teaching effectiveness (Rui et al. 2010).

Environment

Creating a good learning environment has been considered as one of the effective ways to improve teaching efficiency. A good learning environment contributed to the cultivation of the students' learning interest, activation of learning motivation (Oesterle 2008), and the establishment of a good relationship between teachers and students (Wolk 2001). Some educators have asserted that teaching is a process of creating and fostering learning environment in which students actively participate in activities for learning (Floden 2001; Seidel and Shavelson 2007). To improve the effectiveness of teaching and learning, we need to take into account of the perceptions teachers have of their teaching context (Prosser and Trigwell 1997). One of the characteristics of quality teaching is the effective link between school and the cultural context of the school (Alton-Lee 2003).

In sum, teaching–learning environment influences the process and outcomes of learning. Thus, teachers need to develop an understanding of learning environments (Davis et al. 2006).

The Use of Technology

Learning outcomes can be enhanced by technology (Dowing and Harland 2001). One of the ways that raising the student learning outcomes is using technology as a tool for learning, communication, and collaboration (Jones 1995). Research has studied the relationship between effective use of technology and teaching strategy, and the factors influencing the use of technology for teaching (Byrom and Bingham 1998). It has also been reported on most cost-effective and appropriate ways to use technology (e.g., computers, video, and telecommunications technologies) (Chickering and Ehrmann 1996).

The lack of professional development has been identified as one of the biggest barriers to effective use of technology in education (Norman 2000). Benefits perceived by teachers participating in in-service technology training must consider the cost of time and energy (Shelly 2000).

In summary, technology should be used as an effective tool to improve teaching methods and strategies to make the process of learning more meaningful.

Teaching Assessment

Assessment plays an important role in the process of learning and enhances the effectiveness of the learning process. Assessment is a part of the learning and not the end (Gurney 2007). Angelo and Cross have identified characteristics of classroom assessment techniques with each including a concise description and step-by-step procedures for administering the technique (Angelo and Cross 1993). Delandshere (2002) has explored the possible uses of inquiry as a way to understand, assess, and learn. According to Skinner, we need to consider the teacher assessment and balance the "summative" assessment and "formative" assessment (Preece and Skinner 1999). We need to understand the effect of assessment in the process of learning and take advantage of assessment to increase the efficiency of learning.

Although researches on teaching effectiveness have been extensive as reviewed above, there are still a number of limitations. The aforementioned literatures were not systematic, they all paid attention to certain aspects of teaching effectiveness. As teaching effectiveness in science was popular in the past decade, few studies were discipline-based (i.e., chemistry) specific (Preece and Skinner 1999; Çimer 2007; Davis et al. 2006). Moreover, almost no existing quantitative studies explored the main factors of teaching effectiveness. Most of them got their results through qualitative research methods such as interview or summarizing literature. Notwithstanding, however, only few empirical studies have been conducted to quantitatively assess effectiveness of chemistry classroom teaching in China.

In sum, the study presented in this article focuses on chemistry classroom teaching which is specific to Chinese culture. Using systems science theory, we have analyzed the classroom teaching and learning chemistry lessons and developed a chemistry classroom teaching system theory called CPUP model (four-hierarchy system model of Class-Plate-Unit-Primitive), which has further enabled us to explore the influential factors of effectiveness of chemistry classroom teaching and confirmed construct validity and reliability of the instrument for assessing the effectiveness of chemistry classroom teaching.

The reasons for developing an instrument of assessing the effectiveness of chemistry classroom teaching are included: firstly, since the new curriculum reform in 2001, chemical education experts have only provided some certain teaching theory knowledge to secondary chemistry teachers. Many teachers responded that the instructions given by experts can be understood accessibly but hard to be implemented in the real teaching practice. Besides, some expert-like teachers have possessed so many valuable experiences of effective teaching that new teachers always learn the tricks of teaching in observation lessons. It is extraordinary meaningful for transforming these teaching experiences to specific operational evaluation tool. So we made an attempt to construct a good reliability and validity of an instrument to conduct chemistry teachers on how to carry out the effective teaching.

Theoretical Framework

Systems Science Theory

Von Bertalanffy (1950) was among first scholars who to use the concept of systems science. The systems science is a subject of study on system phenomenon and system problem (Miao 2010). In general, a system phenomenon exists in all disciplinary fields. Bertalanffy argued that system was the integration of the elements that have affiliation and interaction. This definition can be briefly represented as (Miao 2010; Von Bertalanffy 1950) follows:

A system S exits if the object set S meets the following two conditions:

1. S contains at least two different objects.
2. The objects of S are associated with each other in a certain way.

Here S stands for a system, and the objects of S are called components of the system. The components of the systems can be divided into smaller components. Minimum components to constitute a system, which cannot be further divided or do not need to be further divided, are called system elements. The basic characteristics of system elements are that they have primitive properties. A property is relative to its membership in the system. Leaving the system, the component element itself can be seen as a system made up of smaller components. The collection of all element contacts is called the system structure. Ignoring irrelevant and irregular contacts, the structure is seen as a relative stable summation with certain rules of the contact method between elements. Elements and structure are two integral aspects of the system composition. The system is the unity of the elements and structure. (Miao 2010; Von Bertalanffy 1950).

In a system with many elements and complicated structures, independent elements are grouped in a relative order, which have their own overall characteristics and are more closely linked in some way. This kind of grouping that exists in the system and becomes the system of the group is called sub-system. Therefore, there exists a hierarchy in the system. Hierarchy exists in a larger system that is composed of interconnected sub-systems with different levels (Liu 2008; Miao 2010).

System of Classroom Teaching (SCT)

Classroom teaching and learning is essential for improving students' scientific literacy (Wang 2005; Lv and Wang 2007). Specifically, research on chemistry teaching and learning is significant. Adopting a systems science theory in this study, we analyze the teaching and learning in chemistry classrooms from a perspective of an internal system. Accordingly, we propose the following chemistry classroom teaching system theory.

CPUP Model

Based on the observation data from more than 900 chemistry lessons in classrooms, we propose that the whole Class System (CS) be composed of several Plate Systems (PS). Each Plate System is made up of several Unit Systems (US), which are made of several Primitive Systems (PrS) (Fig. 1). CPUP model is a four-hierarchy system model which provides a means to divide a whole lesson into several tiny sections. Thus, we analyzed chemistry lesson from the smallest system to the largest system.

An Application Example of the CPUP Model

In one high school (named DSFZ) of the study, a chemistry teacher taught the 10th grade students a lesson about "Molar Volume of Gas," which is in Chemistry Compulsory 1 of the new national chemistry curriculum standard in China (Ministry of Education 2003; People's Education Press 2004). The classroom teaching system in this lesson had four plate systems; they were gas, liquid, solid, and material volume rule. Within the system, students in the class explored the reason for the material volume law, developed a concept of the volume of mole vapor, and applied the concept to establish the link between volume and amount of substance. The classroom teaching system started with the material substance and its changes related to a macroscopic phenomenon to gain experience and to discover the law. Following the above, the system moved to a microcosmic perspective by using the particle view and particle function view to explain the experience and law and to form new concepts and laws. Based on the new concepts and laws, students in the class established the relational knowledge, experiencing a transition from the macroscopic to microcosmic stages. From this process, we can identify a relationship among four plate systems in a logical progression (Fig. 2):

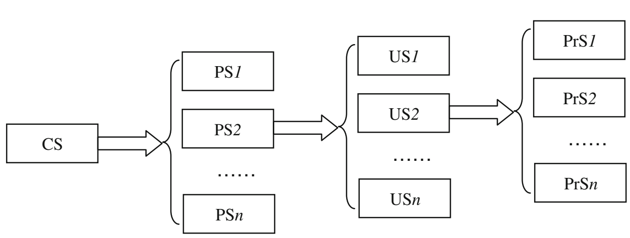


Fig. 1 CPUP model of chemistry classroom teaching system

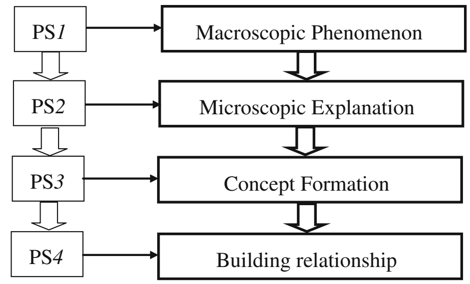


Fig. 2 Structure of plate systems in the CPUP model case

Primitive System (PrS)

Primitive Systems (PrS) are the minimum components of a chemistry classroom teaching system. "Minimum" is relative to the chemistry classroom teaching system; a minimal component may be a certain system composed of other smaller components. A chemistry teaching primitive system has the attributes of its chemistry teaching system, including objectivity, subjectivity, diversity, educational meaning, and other kinds of properties of the chemistry teaching. Thus, a primitive system can satisfy most basic conditions of chemistry teaching.

The following classroom teaching excerpt shows an example of a primitive system:

> [Teacher] OK. Please look at our data for 10 gases. Look carefully through the data tables, what sort of conclusion could you get? Discuss in groups, please.
> (Students were discussing.)
> [Teacher] Well, time is up. Have you got a conclusion? Who can? Come on.
> [Student 1] Most of these numbers ranged from 24 to 25.
> [Teacher] Well. Most of them are located between 24 ml and 25 ml. Do you agree? Guo Shi.
> [Student 2] Although the gas for each number is different, the produced volume is almost the same.
> [Teacher] Almost the same. Answered very well! Anything else?
> [Teacher] Well, for the 10 numbers in the table, most of them are located between 24.0 ml and 25.0 ml, and to the first group and the second group, the two data become partial small, there may be what, with eight other than group there may be some experimental error. So, ignoring the experimental error, we can get such a conclusion: under the same temperature and pressure, the same amount of substance of any gas volume accounts for approximately equal.

In this primitive system case, teacher and students worked together to accomplish the task of discovering the gas law under the same temperature and pressure. In this example, teaching and learning steps were as follows: the teacher put forward a question, and then asked students to discuss. After discussion, two groups were asked to send representatives to report their points of view and to exchange ideas with other groups. And finally, the teacher gave a simple evaluation to students' presentation and then she gave a conclusion.

The purpose of the present study was to identify factors affecting classroom teaching systems and to develop an instrument to measure the quality of primitive systems in chemistry classroom teaching. The specific research questions are the following:

a. How to develop a good reliability and construct validity of evaluation scale?
b. Whether there is a significant difference between well-designed lessons and ordinary lessons or not?
c. How chemistry teachers can use this instrument to improve their everyday lesson planning?

The above research questions are significant both theoretically and practically because developing an instrument is a new way for assessing the effectiveness of chemistry classroom teaching in China. To ensure the scientificalness and accuracy of research results, we use quantitative research methods such as EFA and CFA. Besides, using the instrument, we instructors can give some particular advises to chemistry teachers when we mentoring chemistry classroom teaching. On the basis of instrument results, chemistry teachers may be aware of what they still need to improve and how should they do in their following everyday lesson planning.

Method

Lesson Sampling

In this study, we selected 12 classroom teaching lesson cases from the resource of videotaped classroom lessons. We defined each classroom lesson case as a SCT. The lesson cases are in two groups which contained several lessons of equal quantity, different design styles and content types. The two groups are listed as follows: one group consists of six lessons on Gas Molar Volume, Ionic Reaction (I), Ionization Equilibrium, Neutralization Titration, Mg & Al, and Acetylene; and another group includes six lessons on Ionic Reaction (II), Chemical Property of Metals, Electrolyzation & Electrolytic Application, Ionization Equilibrium, Chemical Reaction, and Chemical Bond and Ethanol.

In order to ensure representativeness of the lessons, we followed a stratified sampling approach based on two aspects: the design quality and the category of chemical knowledge. Six lessons (Gas Molar Volume, Ionic Reaction and Ionization Equilibrium, Ionic Reaction (II), Chemical Property of Metals, and Electrolyzation & Electrolytic Application) were well designed during the national competitions organized by Chemistry Teaching Professional Committee of China Education Society. Six other lessons were ordinary ones from daily classroom teaching. Zheng divided the chemical knowledge into chemical symbol, element and compound, theory, and experiment (Zheng 2006). Among the lessons, Gas Molar Volume, Ionic Reaction (I) and Ionization Equilibrium, Ionic Reaction (II), Electrolyzation & Electrolytic Application, Ionization Equilibrium, and Chemical Reaction and Chemical Bond belong to theoretical knowledge; Mg & Al, Acetylene, Chemical Property of Metals, and Ethanol belong to element and compound knowledge; and Neutralization Titration belongs to experimental knowledge.

We transcribed the videos of the above-selected lessons and divided lesson into PrSs using the CPUP model theory. The PrSs were used as subjects (i.e., observations) for the subsequent EFA (exploratory factor analysis) and CFA (confirmatory factor analysis). The distribution of observations among the lessons is shown in Table 1.

Instrument Development and Validation

The development of the instrument, ESEPrSCT (Evaluation Scale of Effectiveness of Primitive System of Classroom Teaching), followed the following procedures:

We adopted focus group interview method to develop the ESEPrSCT. Steps for developing and validating the instrument included small-scale interviews with ten expert teachers on their perceptions of the characteristics of high-efficiency classroom teaching. These expert teachers mainly came from northeast Chinese cities. All of them have taught secondary chemistry lessons over 20 years, and they were awarded special-class teachers by their local provincial governments; besides, they all took part in national teaching research projects, and some of them have published articles in Chinese journals and books. Thus, we believed that these experts' views could be the mainstream ideas on the effective teaching in mainland China.

For the development of the instrument, based on previously discussed interviews, indicators were hypothesized to be associated with each of the five factors (i.e., rationality of primitive content, rational use of time, match degree, quality of teaching behavior chain, and quality of using resource and technology).

A detailed description of the five categories with tentative factor labels and respective sample items are presented below:

1. *Rationality of primitive content:* Several teachers held the same view that they chose teaching contents

Table 1 Description of sample lessons

Design quality	Exploratory factor analysis		Confirmatory factor analysis	
	Topic	PrS	Topic	PrS
Well-designed	Gas Molar Volume	24	Ionic Reaction (II)	38
	Ionic Reaction (I)	36	Chemical Property of Metals	28
	Ionization Equilibrium	24	Electrolyzation & Electrolytic Application	44
Ordinary	Neutralization Titration	39	Ionization Equilibrium	27
	Mg &Al	43	Chemical Reaction and Chemical Bond	29
	Acetylene	44	Ethanol	37
Total		210		202

mainly according to the curriculum standards and textbooks. As an expert teacher, Dong said, "It is an effective classroom teaching if it will achieve the requirement of curriculum standards and textbooks. And teaching contents must embody tri-dimensional goals." So we coded this argument as an item "this content is appropriate to reflect the curriculum standards and textbooks."

2. *Rational use of time:* Using time properly is a key factor for teaching effectiveness. One of them, Su said, "What is the effective teaching? Considering teaching effectiveness, it cannot ignore the rational use of time. As a chemistry teacher, you cannot waste time on making mistakes or re-presentation in your class." This item we coded "no waste time on making mistake or re-presentation."
3. *Match Degree:* It is important that teaching behavior chain must adapt to teaching content. A typical example was cited by Xu, "When teaching and learning sodium reacts with water, we should arrange for students to do experiments, organize them to discuss, and finally encourage them to get their conclusions." So we coded this point as an item "the type of this teaching behavior chain is consistent with the characteristics of this content."
4. *Quality of teaching behavior chain:* How to handle teaching behavior chain well plays a significant role in improving teaching effectiveness. As an interviewee, Liu said, "In my school, the reason why a teacher fails to organize students' discussion is that the question he posed is not clear." This item should be "question designed lead to students' effective thinking."
5. *Quality of using resource and technology:* Using resource and technology is a key means to promote students' understanding. Han said, "When teaching and learning structure of substance, teacher should display some related models to assist students' understanding deeply." So we coded this point as an item "choosing proper material object (or model, writing on the blackboard, multimedia, etc.) to assist students' understanding."

In this study, each category included from three to five items and was presented in a six-point Likert mode. These 21 items for response categories, namely Strongly disagree, Disagree, Slightly disagree, Slightly agree, Agree, Strongly agree. Scoring was accomplished by assigning a score of 1 to items receiving a "Strongly disagree" response, a score of 2 to "Disagree," and so on through the six response categories.

To further substantiate content validity of the instrument, three specialists (one each in the fields of chemical education, statistical analysts, and chemistry expert teacher in secondary school) examined the items on this evaluation scale. Some items were adjusted for syntax, discourse, and lexical cohesion.

In order to study the reliability of coding, we established a team of three experts. Of the three, one is an expert on curriculum and teaching theory in chemical education, another is a senior teacher who has more than 20-year teaching experience, and the last one is a post-doctoral researcher. The three experts coded the 412 PrSs independently. We then calculated Kendall's coefficient of concordance (W) as a measure of the agreement among the three coders (Kendall 1938; Legendre 2005). W ranges from 0 to 1, with 1 representing perfect concordance (Legendre 2010). The Kendall's W was found to be 0.721, and $\chi^2 = 49.744$, $df = 23$, $p = 0.001$, indicating that the experts were showing a significant agreement among them in coding (Salkind 2007).

We conducted exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to validate the ESE PrSCT. We randomly divided the entire sample of observations ($n = 412$) into two subsamples: for the EFA ($n = 210$) and for the CFA ($n = 202$), respectively. We used Statistical Program for the Social Sciences (SPSS), version 11.5, to conduct descriptive analysis and EFA. We carried out the steps to analyze the items, assess the adequacy of the matrix of correlations among the items (Ferketich 1991), extract the factors, rotate them, examine the factor loadings, interpret the factors, and determine the reliability of the scales (Glynn et al. 2009; Gorsuch 2003; Henson and Roberts 2006; Mainous 1993; Reise et al.

2000; Thompson 2004). In addition, CFA was conducted to test the structure of the scales. We used AMOS 17.0, a commonly used software package for the analysis of latent variable structures to conduct CFA (Kaplan 2000; Nokelainen 2007; Schumacker and Lomax 2004).

Results

Item Analysis

The relevance and discriminating power are essential qualities of a good educational test item (Ebel 1954). The comparison groups, in this study, was defined as upper and lower 27 percents based on the total scores of ESEPrSCT (Lange et al. 1967). According to Ebel (1954), we calculated the difference in scores between the upper group and the lower group as a measure of an item's relative effectiveness in achieving desired discrimination (Ebel 1954; Findley 1956).

An independent samples t test was conducted to compare the upper group and lower group. There was a significant difference in the scores for all items, $t\ (210) > 3$, $p < 0.001$, except for item 3. For item 3, the score for upper group ($M = 3.06$, $SD = 1.77$) and the score for lower group ($M = 2.68$, $SD = 1.20$) resulted in a $t\ (210) = 1.40$ ($p = 0.164$), suggesting that item 3 had no discriminating power. Besides, the intercorrelations between the score of each item and the total score of all items were calculated. According to Ferketich, the correlations of $r < 0.30$ or $r > 0.70$ indicated that the item is not sufficiently related or redundant and probably unnecessary (Ferketich 1991). Thus, except item 3, item to total corrected correlations were all above 0.30 which was good (Ferketich 1991; Nunnally et al. 1967). And the correlation between item 3 and the total was 0.14, indicating that item 3 did not contribute to measurement of the main factor. As the result, item 3 was removed, and remaining 20 items of the ESEPrSCT were kept for subsequent exploratory factor analysis.

Exploratory Factor Analysis

To create a valid measure of an underlying construct, factor analysis can play a crucial role in ensuring the discriminant validity of scales. In this study, 210 observations were above the suggested number from Gorsuch and the number was ten times of the number of the items (Clark and Watson 1995; Gorsuch 1983; Guadagnoli and Velicer 1988).

Prior to conducting the EFA, the KMO measure of sampling adequacy index was found to be 0.829, and Bartlett's test of sphericity, $\chi^2 = 2514.691$, $df = 190$, $p < 0.001$, indicating that the sample was appropriate for such an analysis.

The goal of factor extraction is to identify the number of latent dimensions (factors) need to accurately account for the common variance among items. To extract factors, a principal component analysis with an oblique rotation was performed on the items of ESEPrSCT. Using the Kaiser–Guttman rule, we identified five factors that had an eigenvalue greater than 1. We also used a Scree Plot to examine potential factors by plotting them with their eigenvalues in descending order. These five factors accounted for 67.99 % of the total variance, which is considered good (Glynn et al. 2009). The five factors were rotated, turning their reference axes around their origins. We used a varimax rotation to produce what is called a simple structure to facilitate interpretation. Table 2 presents the rotation results.

The factor loadings of items should be greater than 0.4 on the relevant factor and less than 0.4 on all other factors (Lee et al. 2008; Steven 1996). Table 2 shows that all of the items met the criterion of loading at least 0.35 on their respective factor (Glynn et al. 2009; Tabachnick and Fidell 2000). The communalities of all the 20 items are at least 0.517.

In addition, the Cronbach's coefficient alphas for the five factors ($n = 210$) were 0.914, 0.816, 0.820, 0.693, and 0.622 (see details in Table 3), and the overall alpha was 0.867, indicating that these factors had highly sufficient reliability in assessing the effectiveness of primitive systems of classroom teaching (Nunnally et al. 1967).

Factor 1 contained five items and was about using time in the teaching system, so we labeled this factor rational use of time (RUT). This 5-item factor was the most important of the five factors because it explained 20.8 % of the total amount of variation in the instrument. Factor 2 contained five items also and all of them related to teaching behaviors; we labeled this factor quality of teaching behavior chain (QTBC). This factor was the second most important of the five, explaining 17.247 % of the total variation in the instrument. Factor 3 contained four items related to the matching adaptability between teaching behavior chain and other four aspects (teacher, students, content, and resource); thus, we labeled this factor match degree (MD). MD accounted for 11.584 % of the total variation in the instrument. Factor 4 contained three items that reflected how teachers use resources and technology; we named this factor as quality of using resource and technology (QUR&T). It explained 9.190 % of the total variation in the instrument. The last factor, factor 5, also contained three items, and they were about teaching content; we named rationality of primitive content (RPrC) and accounted for 9.130 % of the total variation in the instrument.

Table 2 Rotated component matrix for the ESEPrSCT, communalities, means, and SD ($n = 210$)

Item #	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	M	SD	h^2
Factor 1: rational use of time(RUT)								
9	**0.922**	0.017	0.103	0.151	0.034	5.410	0.950	0.885
8	**0.912**	0.007	0.099	0.121	−0.010	5.400	0.979	0.856
10	**0.893**	0.081	0.105	0.087	0.086	5.424	0.900	0.830
7	**0.705**	0.089	0.267	0.084	0.099	5.319	0.711	0.593
6	**0.679**	−0.048	0.276	0.180	0.032	5.410	0.673	0.573
Factor 2: quality of teaching behavior chain (QTBC)								
16	−0.117	**0.799**	0.192	0.090	0.148	3.224	1.395	0.719
20	0.036	**0.773**	0.212	0.038	0.311	2.852	1.335	0.742
19	0.107	**0.739**	0.201	0.094	−0.026	3.414	1.409	0.608
21	−0.033	**0.709**	0.113	0.146	0.007	1.695	1.179	0.538
18	0.398	**0.561**	0.074	−0.193	0.029	3.614	1.444	0.517
Factor 3: match degree (MD)								
13	0.136	0.378	**0.712**	0.232	0.146	4.357	1.120	0.743
12	0.098	0.495	**0.682**	0.169	0.157	4.362	1.027	0.772
14	0.339	0.182	**0.665**	0.019	0.093	4.695	0.790	0.599
11	0.360	0.147	**0.652**	−0.125	0.149	4.938	0.739	0.613
Factor 4: quality of using resource and technology (QUR&T)								
2	0.094	−0.088	0.220	**0.767**	0.039	4.867	0.664	0.655
4	0.235	0.431	−0.092	**0.710**	−0.012	5.024	0.566	0.754
5	0.438	0.284	−0.032	**0.619**	0.075	5.100	0.728	0.661
Factor 5: rationality of primitive content (RPrC)								
17	0.018	0.062	0.037	−0.169	**0.842**	2.648	1.454	0.743
1	0.093	0.330	0.095	0.093	**0.695**	3.500	1.064	0.619
15	0.093	−0.049	0.301	0.255	**0.640**	3.562	1.528	0.577

a. h^2 = communalities of the measured variables

b. Pattern coefficients with absolute values of 0.40 or greater are in bold

c. Teaching Behavior Chain: in a certain chemistry classroom primitive system, teaching behaviors link together to form a specific function and meaningful chain pattern. The common teaching behavior chains are as follow: questioning–answering–summarizing, questioning–group discussion–report communication–summarizing, etc

Table 3 Eigenvalue, percent of variance explained, and Cronbach's coefficient alpha for each factor

Factor	Eigenvalue	% of variance	Cumulative %	Cronbach's alpha
RUT	4.168	20.838	20.838	0.914
QTBC	3.449	17.247	38.085	0.816
MD	2.317	11.584	49.669	0.820
QUR&T	1.838	9.190	58.859	0.693
RPrC	1.826	9.130	67.989	0.622

In the diagram, cov_1 means add covariation between e17 and e20; cov_2 means add covariation between e10 and e13; cov_3 means add covariation between e18 and e20

① RUT is rational use of time, QTBC is quality of teaching behavior chain, MD is match degree, QUR&T is quality of using resource and technology, and RPrC is rationality of primitive content

Confirmatory Factor Analysis ($n = 210$)

According to the result of exploratory factor analysis, we tested a first-order factor measurement model of five factors, the measurement model as shown in Fig. 3.

Using AMOS 17.0, we obtained the initial model and final model fitting indicators shown in Table 4.

Wu suggested that the RMSEA value should be less than 0.08; GFI and AGFI more than 0.09; NFI, RFI, IFI, TLI CFI more than 0.09; and χ^2/df less than 2 (Wu, 2011). According to Table 4, χ^2/df was 2.860 (more than 2); RMSEA value was 0.096 (more than 0.08); GFI and AGFI were smaller than 0.09; NFI and RFI were less than 0.09. All of the above indicated that the initial model and observed data did not agree with each other. For the final model, we can see from Table 4 that the RMSEA value was less than 0.08, GFI and AGFI value close to 0.09; NFI, RFI, IFI, TLI, and CFI were larger than 0.09; chi-square/*df* was 1.838 (less than 2), suggesting that the revised final model improved over the initial model, and produced values close to ideal indices (Table 5).

Confirmatory factor analysis of the final measurement model showed that the measurement variable standardized factor loading was greater than 0.7, in line with the factor load more than 0.5, suggesting that factors on measuring model had strong capacity to explain. The comprehensive reliability (CR) values were more than 0.9. According to Wu's suggestion, a value of 0.6 would indicate that the scales had very good internal consistency reliability (Wu 2011). Five AVE values were greater than 0.7, higher than what Wu suggested value of 0.5 minimum, suggesting that measurement model had good convergent validity.

Table 6 shows the correlation among the five factors. From Table 6, we see that each factor has a square root of AVE greater than 0.7; most correlations had a correlation coefficient less than 0.5. The above suggests that the measurement instrument had good discriminate validity.

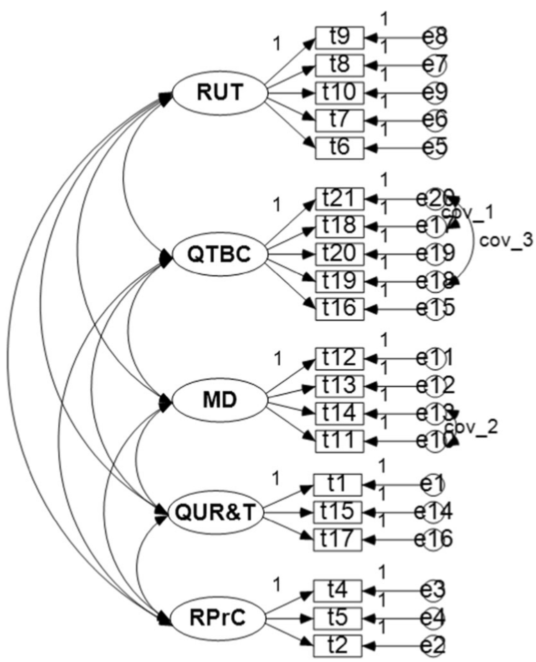


Fig. 3 Hypothesized measurement model

Table 5 Summative results of confirmatory factor analysis on the final model

Factors	Items	Loading	SE	CR	AVE
RUT	t8	0.919	0.155	0.956	0.812
	t7	0.903	0.185		
	t9	0.951	0.096		
	t6	0.859	0.262		
	t5	0.871	0.241		
QTBC	t20	0.795	0.368	0.931	0.731
	t17	0.788	0.379		
	t19	0.915	0.163		
	t18	0.876	0.233		
	t15	0.894	0.201		
MD	t11	0.955	0.088	0.941	0.800
	t12	0.962	0.075		
	t13	0.842	0.291		
	t10	0.809	0.346		
QUR&T	t1	0.956	0.086	0.956	0.879
	t14	0.966	0.067		
	t16	0.889	0.210		
RPrC	t3	0.896	0.197	0.940	0.834
	t4	0.980	0.040		
	t2	0.869	0.245		

① the CR value calculation formula: $(\Sigma \text{ standardized factor load})^2/[(\Sigma \text{ standardized factor load})^2 + \Sigma \text{ error variance}]$

② the AVE value calculation formula: $(\Sigma \text{ standardized factor load}^2)/[(\Sigma \text{ standardized factor load}^2) + \Sigma \text{ error variance}]$

Discussion and Conclusion

The measurement instrument (Evaluation Scale of Effectiveness of Primitive System of Classroom Teaching) consisting of five factors has been developed through an extensive literature review on effective teaching, critiques by experts in the field, and the classroom observation analyzed by EFA and CFA. Data analysis indicated that the instrument developed in this study has satisfactory validity and reliability measures. In this study, the main finding is the formulation of a five-factor model for assessing the effectiveness of primitive systems of chemistry classroom teaching. The five factors are the following: rational use of time (RUT), the quality of teaching behavior chain (QTBC), match degree (MD), the quality of using resource and technology (QUR&T), and the rationality of primitive content (RPrC). The five factors of the present study are similar to those of Cimer, who held the theoretical principles of constructivism

Table 4 Fitting index of initial model and final model

	χ^2	*df*	χ^2/df	RMSEA	GFI	AGFI	NFI	RFI	IFI	TLI	CFI
Initial model	457.616	160	2.860	0.096	0.816	0.759	0.899	0.880	0.932	0.918	0.931
Final model	288.515	157	1.838	0.065	0.883	0.844	0.936	0.923	0.970	0.963	0.970

Table 6 Correlations among the factor-based scales and AVE value of each factor

	RUT	QTBC	MD	QUR&T	RPrC
RUT	0.901(AVE)				
QTBC	0.006	0.855(AVE)			
MD	0.143	0.569	0.894(AVE)		
QUR&T	0.084	0.309	0.220	0.937(AVE)	
RPrC	0.369	0.181	0.097	0.108	0.913(AVE)

① The diagonal numerical values are square root of AVE value of each factor

② The non-diagonal numerical values are the correlation of factors

Table 7 Independent samples *t* test between well-designed lessons and ordinary lessons in factor-based scores and total scores

Factors	Well-designed lessons		Ordinary lessons		*t*
	M	SD	*M*	SD	
QUT	26.15	4.84	22.32	6.09	7.04***
QTBC	14.81	6.71	12.51	4.49	4.13***
MD	17.42	4.48	15.58	4.20	4.27***
QUR&T	10.54	4.02	8.41	3.26	5.93***
RPrC	14.86	2.38	13.71	2.95	4.39***
Total	83.77	13.65	72.53	14.47	8.12***

*** $p < 0.001$

and reviewed literatures of teaching models, summarized six main principles of effective teaching (Çimer 2007). And Gurney also found five key factors which could contribute to a good teaching (Gurney 2007). Regarding their findings, some factors such as "classroom activities," "assessment activities," "effective feedback," and "effective interaction" can be found in our instrument (see item 16, 21, 20 in "Appendix"). In contrast, the findings of the present study provide a more specific way to measure the effectiveness of classroom teaching. Furthermore, we constructed instrument of five dimensions not only based on system of classroom teaching theory (SCT), but also on the strength of data analysis through classroom observation. So we believe that the interaction of the five key factors provides a foundation for a good teaching in the creation of an effective learning environment. Chemistry teacher who focus on the areas will become an effective teacher in the near future (Table 7).

An independent samples *t* test was conducted to compare effectiveness of teaching between well-designed lessons and ordinary lessons in factor-based scores and total scores. The scores on the RUT scale were higher among the well-designed lessons ($M = 26.15$, $SD = 4.84$) than the ordinary lessons ($M = 22.32$, $SD = 6.09$), $t(412) = 7.04$, $p = 0.00$, suggesting that teachers in well-designed classes used time more properly than teachers in ordinary classes. The scores on the QTBC scale were higher among the well-designed lessons ($M = 14.81$, $SD = 6.71$) than the ordinary lessons ($M = 12.51$, $SD = 4.49$), $t(412) = 4.13$, $p = 0.00$, suggesting that teachers in well-designed classes controlled their teaching behavior chain more successfully than teachers in ordinary classes. The scores on the MD scale were higher among the well-designed lessons ($M = 17.42$, $SD = 4.48$) than the ordinary lessons ($M = 15.58$, $SD = 4.20$), $t(412) = 4.27$, $p = 0.00$, suggesting that teachers in well-designed classes are more skilled in selecting the proper teaching method on the basis of the content than teachers in ordinary classes. The scores on the QUR&T scale were higher among the well-designed lessons ($M = 10.54$, $SD = 4.02$) than the ordinary lessons ($M = 8.41$, $SD = 3.26$), $t(412) = 5.93$, $p = 0.00$, suggesting that teachers in well-designed classes are more skilled in using the resource and technology than teachers in ordinary classes. The scores on the RPrC scale were slightly higher among the well-designed lessons ($M = 14.86$, $SD = 2.38$) than the ordinary lessons ($M = 13.71$, $SD = 2.95$), $t(412) = 4.39$, $p = 0.00$, suggesting that teachers in well-designed classes handled the contents more expertly than teachers in ordinary classes. And thus, there was a significant difference in the scores between well-designed lessons ($M = 83.77$, $SD = 13.65$) and ordinary lessons ($M = 72.53$, $SD = 14.47$) with a $t(412) = 8.12$, $p = 0.00$. These results all suggest that lessons meticulously designed by teachers would be more effective than ordinary lessons.

As chemistry teachers, they can use the main factors of teaching effectiveness to improve lesson planning by themselves. This implies that teachers' knowledge of these five factors may assist themselves in enhancing effectiveness of their classroom teaching. For example, if chemistry teachers are more aware that the factor "rational use of time" is a key component of effectiveness of classroom teaching, then they will prepare lessons carefully for using time reasonably without mistakes or unreasonable generation. Thus, the results of this study provide a theoretical framework for efficient chemistry classroom teaching designs; furthermore, the instrument, ESEPrSCT, we have developed can be used as a standardized means to evaluate and improve chemistry classroom teaching by assessing the effectiveness of primitive systems. After finishing their lessons, according to the scores of ESEPrSCT, chemistry teachers will receive some micromesh advises by the expert coders that you should pay attention to "summarizing properly in an opportune moment" for 80 percent of PrSs got a low score in this item. Besides, you may be suggested exactly where the problem is, and how you can do better next time.

As mentioned above, this instrument is one of the first such attempts to explore the main factors of teaching effectiveness. In this way, the finding of this study may provide chemistry educators and even science researchers in the world invaluable insights regarding teaching effectiveness in science classroom teaching. Furthermore, we constructed system of classroom teaching (SCT) only in Chinese chemistry classrooms, and the potential fitness for science classrooms and for other countries should be further confirmed. This study presents a new approach to look into science classroom and to evaluate the effectiveness of science lessons for science educators in the world. However, the weakness of this study is that the outcomes are just based on a sample of new teaching lessons; the appropriateness of review lessons and exercise lessons remains to be researched in the future.

Further researches will expand the sampling extent to identify the adaptability of ESEPrSCT. Besides, the findings in present study can be further employed to do a series of researches about the five aspects of primitive system. And we have already finished some related researches on the important degree of primitive content, the difficulty degree of primitive content. Furthermore, as we have found how to assess the effectiveness of primitive system of classroom teaching (PrSCT), the next researches will focus on the evaluation standard for assessing the effectiveness of unit system (US), plate system (PS), and class system (CS).

Appendix: Initial Hypothesized Assessing Instrument

Evaluation scale of effectiveness of primitive system of classroom teaching (ESEPrSCT)

Items	StD	D	SiD	SiA	A	StA
1. Rich and innovative of this selective content material	1	2	3	4	5	6
2. This content is appropriate to reflect the curriculum standards and textbooks	1	2	3	4	5	6
3. This generated content is reasonable	1	2	3	4	5	6
4. The breadth and depth of this content is in students' zone of proximal development	1	2	3	4	5	6
5. No unreasonable deepen and widen to this content	1	2	3	4	5	6

Appendix continued

Items	StD	D	SiD	SiA	A	StA
6. Using time properly according to characteristics of this content	1	2	3	4	5	6
7. Compact and appropriate speed of the teaching process	1	2	3	4	5	6
8. No waste time on the lack of clarity	1	2	3	4	5	6
9. No waste time on the unreasonable generation	1	2	3	4	5	6
10. No waste time on making mistake or re-presentation	1	2	3	4	5	6
11. Teacher can well manage the type of this teaching behavior chain	1	2	3	4	5	6
12. The type of this teaching behavior chain is consistent with the learning characteristics of students	1	2	3	4	5	6
13. The type of this teaching behavior chain is consistent with the characteristics of this content	1	2	3	4	5	6
14. The type of this teaching behavior chain is consistent with the school resources	1	2	3	4	5	6
15. Choosing proper material object (or model, writing on the blackboard, multimedia, etc.) to assist students' understanding	1	2	3	4	5	6
16. Participating fully in teaching and learning activities (discussion, communication, question–answer, etc.)	1	2	3	4	5	6
17. Selecting experimental material properly to attract students' attention	1	2	3	4	5	6
18. Summarizing properly in an opportune moment	1	2	3	4	5	6
19. Question designed lead to students' effective thinking	1	2	3	4	5	6
20. Teacher–student and student–student communicate fully with each other	1	2	3	4	5	6
21. Encouraging students to self-evaluation	1	2	3	4	5	6

Key: *StD* strongly disagree, *D* disagree, *SiD* slightly disagree, *SiA* slightly agree, *A* agree, *StA* strongly agree

① Item 3 "The generated content is reasonable" was deleted from the final instrument

References

Alton-Lee A (2003) Quality teaching for diverse students in schooling: Best evidence synthesis. Retrieved from http://www.educationcounts.govt.nz/publications/series/2515/5959

Angelo T, Cross P (1993) Classroom assessment techniques: a handbook for college teachers. Josse Bass Higher and Adult Education: Jossey-Bass, San Francisco

Baker JA (1999) Teacher-student interaction in Urban at-risk classrooms: differential behavior, relationship quality, and student satisfaction with school. Elem Sch J 100(1):57–70

Birch SH, Ladd GW (1997) The teacher-child relationship and children's early school adjustment. J Sch Psychol 35(1):61–79. doi:10.1016/s0022-4405(96)00029-5

Byrom E, Bingham M (1998) Factors influencing the effective use of technology for teaching and learning: lessons learned from the SEIR-TEC intensive site schools. SERVE, Greensboro, NC

Cameron M (2002) Peer influences on learning. Set 3:36–40

Chickering AW, Ehrmann SC (1996) Implementing the seven principles: technology as lever. AAHE Bull 49(3):3–7

Çimer A (2007) Effective teaching in science: a review of literature. J Turk Sci Educ 4(1):20

Clark LA, Watson D (1995) Constructing validity: basic issues in objective scale development. Psychol Assess 7(3):309–319

Cui Y (2009) Effective teaching. East China Normal University Press, Shanghai (in Chinese)

Darling-Hammond L (2000) Teacher quality and student achievement: a review of state policy evidence. Educ Policy Anal Arch 8(1). Retrieved from http://epaa.asu.edu/ojs/article/view/392

Davis EA, Petish D, Smithey J (2006) Challenges new science teachers face. Rev Educ Res 76(4):607–651. doi:10.3102/00346543076004607

Delandshere G (2002) Assessment as inquiry. Teach Coll Rec 104(7):1461–1484

Dowing L, Harland D (2001) Critical factors in the effective use of technology. Walden University. Retrieved from http://www.dowlingcentral.com/gradschool/Edu6420/project1.html

Ebel RL (1954) Procedures for the analysis of classroom tests. Educ Psychol Measur 14(2):352–364. doi:10.1177/001316445401400215

Ferketich S (1991) Focus on psychometrics. Aspects of item analysis. Res Nurs Health 14(2):165–168. doi:10.1002/nur.4770140211

Findley WG (1956) A rationale for evaluation of item discrimination statistics. Educ Psychol Measur 16(2):175–180. doi:10.1177/001316445601600201

Floden RE (2001) Research on effects of teaching. In: Richardson V (ed) Handbook of research on teaching, 4th edn. American Educational Research Association, Washington, DC, pp 3–16

Glynn SM, Taasoobshirazi G, Brickman P (2009) Science Motivation Questionnaire: construct validation with nonscience majors. J Res Sci Teach 46(2):127–146. doi:10.1002/tea.20267

Goldhaber D, Anthony E (2007) Can teacher quality be effectively assessed? National Board Certification as a Signal of Effective Teaching. Rev Econ Stat 89(1):134–150. doi:10.1162/rest.89.1.134

Gorsuch RL (1983) Factor analysis, 2nd edn. Lawrence Erlbaum Associates, Hillsdale, NJ

Gorsuch RL (2003) Factor analysis. In: Schinka JA, Velicer WF (eds) Handbook of psychology: research methods in psychology. John Wiley, Hoboken, NJ

Guadagnoli E, Velicer WF (1988) Relation of sample size to the stability of component patterns. Psychol Bull 103(2):265–275

Gurney P (2007) Five factors for effective teaching. N Z J Teach Work 4(2):89–98

Haycock K (1998) Good teaching matters: how well-qualified teachers can close the gap. Thinking K 16(3):2

He X (2007) Reflect on the effectiveness of classroom teaching under the background of new curriculum. Teach Learn Ref Middle Sch Chem (in Chinese)

Henson RK, Roberts JK (2006) Use of exploratory factor analysis in published research. Educ Psychol Measur 66(3):393–416. doi:10.1177/0013164405282485

Howard ME (2008) Effective teachers' perceptions of the impact of the multiple dimensions of teaching on student success in the elementary classroom (Order No. 3310953, The University of North Carolina at Chapel Hill). ProQuest Dissertations and Theses, pp 167–n/a. Retrieved from http://search.proquest.com/docview/304530855?accountid=12840

Jones, B. F. (1995). New times demand new ways of learning. Retrieved from http://www.netc.org/cdrom/plug_in/html/newtimes.htm

Kaplan D (2000) Structural equation modeling: foundations and extensions. SAGE Adv Quant Tech Soc Sci Seri 10. ISBN 0-7619-1407-2

Kaplan LS, Owings WA (2001) Teacher quality and student achievement: recommendations for principals. NASSP Bull 85(628):64–73. doi:10.1177/019263650108562808

Kendall MG (1938) A new measure of rank correlation. Biometrika 30(1/2):81–93

Lahue-McCully M (2012) Traditional teaching strategies and effective teaching strategies in connection to student achievement: a comparison in one school district's high school English classrooms. Doctoral dissertation. Retrieved from http://search.proquest.com/docview/1019055700

Lange A, Lehmann IJ, Mehrens WA (1967) Using item analysis to improve tests. J Educ Meas 4(2):65–68. doi:10.1111/j.1745-3984.1967.tb00572.x

Lee M-H, Johanson RE, Tsai C–C (2008) Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis. Sci Educ 92(2):191–220. doi:10.1002/sce.20245

Legendre P (2005) Species associations: the Kendall coefficient of concordance revisited. J Agric Biol Environ Stat 10(2):226–245. doi:citeulike-article-id:201411

Legendre P (2010) Coefficient of concordance. In: N Salkind (ed) Encyclopedia of research design. SAGE Publications, Inc., Thousand Oaks, CA, pp 165–170. doi:10.4135/9781412961288.n55

Liu J (2008) Complexity: a philosophical view. Hunan Science and Technology Press, ChangSha, China (in Chinese)

Lv X, Wang Y (2007) Classroom instructional strategies of the training of students scientific quality in American science education. Studies in Foreign Education (in Chinese)

Mainous GA (1993) Factor analysis as a tool in primary care research. Fam Pract 10(3):330–336. doi:10.1093/fampra/10.3.330

Miao D (2010) Essentials of systems science, 3rd edn. China Renmin University Press, BeiJing (in Chinese)

Ministry of Education (2003) Chemistry curriculum standard of general senior secondary school. People's Education Press, Beijing (in Chinese)

Nokelainen P (2007) Introduction to structural equation modeling. Retrieved from http://www.uta.fi/arkisto/aktk/lectures/sem_en/pdf/1999_pn.pdf

Norman MM (2000) The human side of school technology. Educ Digest 65(7):45–52

Nunnally JC, Bernstein IH, Berge JMT (1967) Psychometric theory, vol 2. McGraw-Hill, New York

Oesterle M (2008) Student perceptions of effective teaching. Doctoral dissertation, Arizona State University. Retrieved from http://search.proquest.com/docview/304686040

People's Education Press, CMRI (2004) High school curriculum standard textbook: chemistry (Module 1). People's Education Press, Beijing (in Chinese)

Peterson PL, Fennema E (1985) Effective teaching, student engagement in classroom activities, and sex-related differences in learning mathematics. Am Educ Res J 22(3):309–335

Pianta RC (1999) Enhancing relationships between children and teachers. American Psychological Association, Washington, DC

Preece PF, Skinner NC (1999) The national assessment in science at Key Stage 3 in England and Wales and its impact on teaching and learning. Assess Educ Princ Pol Pract 6(1):11–25

Prosser M, Trigwell K (1997) Relations between perceptions of the teaching environment and approaches to teaching. Br J Educ Psychol 67(1):25–35. doi:10.1111/j.2044-8279.1997.tb01224.x

Reise SP, Wallar NG, Comrey AL (2000) Factor analysis and scale revision. Psychol Assess 12(3):287–297. doi:10.1037//1040-3590.12.3.287

Rui LJC, Yoau-chau J, Shi H (2010) The relationship between teacher quality and teaching effectiveness perceived by students from industrial vocational high schools. Arts Sci 1:167–187

Salkind NJ (ed) (2007) Encyclopedia of measurement and statistics, vol 1–3. SAGE Publications, Inc., Thousand Oaks, CA. doi:10.4135/9781412952644

Schumacker RE, Lomax RG (2004) A beginner's guide to structural equation modeling, 2nd edn. Lawrence Erlbaum Associates, Mahwah, NJ

Seidel T, Shavelson RJ (2007) Teaching effectiveness research in the past decade: the role of theory and research design in disentangling Meta-Analysis results. Rev Educ Res 77(4):454–499

Shelly RW (2000) From literacy to fluency in instructional technology: taking your staff to the next level. NASSP Bull 84(614):61–70. doi:10.1177/019263650008461408

Smith K (1995) Strategies for effective teaching a handbook for teaching assistants. Retrieved from http://www.engr.wisc.edu/services/elc/strategies.pdf

Song Q (2004) On teaching effectiveness. Curric Teach Mater Method 24(10):25–29 (in Chinese)

Steven J (1996) Applied multivariate statistics for the social science, 3rd edn. Erlbaum, Mahwah, NJ

Sun Y (2008) A study on the effective classroom teaching standards. Educational Science Publishing House, Beijing (in Chinese)

Tabachnick BG, Fidell L (2000) Using multivariate statistics, 4th edn. Allyn&Bacon, Boston

Tang Y (2003) Discussion of teaching communication effectiveness. J Chin Soc Educ 6:35–37 (in Chinese)

Thompson B (2004) Exploratory and confirmatory factor analysis: understanding concepts and applications. American Psychological Association, Washington, DC

Von Bertalanffy L (1950) An outline of general system theory. Br J Philos Sci 1:134–165. doi:10.1093/bjps/I.2.134

Wang X (2000) Classroom behavior difference research on effective teaching and inefficient teaching. Educ Theory Pract 20(9):50–53 (in Chinese)

Wang C (2005) The research of developing the scientific literacy of students in the chemistry teaching process of senior middle school. Master's thesis. Retrieved from http://epub.cnki.net/kns/default.htm

Wenglinsky H (2000) How teaching matters bringing the classroom back into discussions of teacher quality. Retrieved from http://www.ets.org/Media/Research/pdf/PICTEAMAT.pdf

Wolk S (2001) The benefits of exploratory time. Educ Lead 59(2):56–59

Wu M (2011) Structural equation modeling—operation and application of using AMOS. Chong Qing University Press, Chong Qing (in Chinese)

Yao L (2004) The elementary approach to the implications of effective teaching. Mod Univ Educ 5:10–13 (in Chinese)

Zhao Z (2002) Optimizing students mood—effective teaching strategy. People's Educ 5:45–46 (in Chinese)

Zheng C (2006) New chemistry curriculum and teaching. Graduated College of Northeast Normal University, Changchun (in Chinese)

J Sci Educ Technol (2014) 23:280–297
DOI 10.1007/s10956-013-9486-0

Enactment of Scientific Inquiry: Observation of Two Cases at Different Grade Levels in China Mainland

Lei Wang · Ronghui Zhang · David Clarke · Weizhen Wang

Published online: 17 January 2014

Abstract Enactment of scientific inquiry in classroom has attracted a great attention of science educators around the world. In this study, we examined two competent teachers' (one Grade 9 chemistry teacher and one Grade 4 science teacher) enactment of scientific inquiry in selected teaching units to reveal the characteristics of enacted inquiry at different grade levels by analyzing lesson sequence videos. The coding schemes for enacted inquiry consist of ontological properties and instructional practices. Pre-topic and post-topic teacher interviews and the two teachers' responses to a questionnaire were adopted to identify the factors influencing teacher's enactment. The results indicate that the two case teachers' enactment involved a range of inquiry activities. The enacted inquiry at fourth-grade level covered all the inquiry elements, tending to engage students in the whole procedure of inquiry. The ninth-grade chemistry class placed emphasis on the elements "making plans" to solve problems in authentic context. Important factors influencing the enactment include teacher's understanding about scientific inquiry, textbooks, assessment, students and resource. Implications for inquiry enactment and instruction improvement have been provided.

Keywords Scientific inquiry · Enactment · Factors · Observation · Ontological properties · Instructional practice

L. Wang (✉) · R. Zhang · W. Wang
College of Chemistry, Beijing Normal University, Beijing 100875, China
e-mail: wangleibnu@126.com

D. Clarke
Melbourne Graduate School of Education, University of Melbourne, Melbourne, VIC 3053, Australia

Backgrounds

Chinese Ministry of Education initiated a new round of general education reform nationwide at the beginning of 2000. Within 1 year, the new science curriculum standards for Grades 1 through 9 were released by Ministry of Education (Chinese Ministry of Education 2001a, b). The mission of this science education curriculum reform was to shift the emphasis from transfer of knowledge in the classroom to development of students' scientific literacy with inquiry-based teaching (Liu et al. 2012). As required by the reform document, integrated science curriculum was carried out at elementary level all over the country, while most provinces adopted separated science subjects at middle school and high school levels including chemistry, physics, biology and geography.

Similar to the situation all over the world, scientific inquiry has been a key aspect in the basic education reform in China. In the *Science Curriculum Standard for Grades 3–6* (Chinese Ministry of Education 2001a) and *Chemistry Curriculum Standard for Junior High School* (Chinese Ministry of Education 2001b), scientific inquiry is articulated as both a learning method and a learning goal, which indicates its important role in basic science education. As the reform was initiated at the national level, science teachers began to implement inquiry-based teaching in science classrooms (Wang 2010a, b). Consequently, researchers become interested in the question how scientific inquiry is implemented in classrooms, especially the characteristics of the situation at different grade levels.

A few studies have examined how teachers enacted inquiry in the classrooms. The instrument DiISC (*Discourse in Inquiry Science Classrooms*) was developed to measure teachers' use of instructional strategies in their classrooms that support oral and written discourse, and academic language development embedded in inquiry according to learning principles (Arizona State University 2008). The DiISC has five scales in relation to five sets of instructional strategies. The scales are *Inquiry, Oral Discourse, Writing, Academic Language Development* and *Learning Principles*. Fu et al. (2007) investigated the frequencies and time length of inquiry using middle school and high school teacher questionnaires. Zhou et al. (2005) found that teacher's belief, teacher's subject knowledge and the ability to respond to class situation, assessment system, instruction time, school leader or coordinator had an impact on inquiry-based teaching by tracking and interviewing one ninth-grade chemistry teacher and another physics teacher.

While the above studies provided insight into the classroom practices in terms of scientific inquiry, few reported studies focused on the interaction between teacher and students in inquiry classroom. Moreover, few studies put sight into the features of inquiry classroom at different grade levels. Given the importance and emphasis of scientific inquiry in curriculum standards, this study explored the inquiry classroom at different grade levels to reveal their features and to identify the factors influencing teacher's enactment of scientific inquiry. As such, we intended to answer the following two research questions: (a) What are the characteristics of the enacted inquiry at different grade levels, especially at elementary grade and middle school? (b) Which factors influence teacher' enactment of scientific inquiry?

Theoretical Framework

Scientific Inquiry

A key aim of science teaching reform efforts has been to engage students in the epistemological aspects of science authentically. This aim is behind the considerable attention recently dedicated to inquiry and nature of science instruction (Ford and Wargo 2007).

The *National Science Education Standards* (NRC 1996) claimed that all students should develop abilities necessary to do scientific inquiry and understanding about science inquiry. Elements of inquiry in the standards were involved in the following text:

> Inquiry is a multi-faceted activity that involves making observations, posing questions, examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations. (NRC 1996, p. 23)

Inquiry is often framed as consisting of both process skills and understandings about the nature of science (e.g., NRC 1996). Process skills include designing investigations and collecting and analyzing data. Understandings about the nature of science consist of aspects of the philosophy and sociology of science, such as the tentative nature of theory or the role of creativity in experimentation. Together, the process skills and understandings are intended to provide an accessible, authentic image of how scientists engage in their practices of studying the natural world (Breslyn and McGinnis 2012).

The above views on scientific inquiry are rooted in early science education literature. According to Schwab (1962), "teaching science as inquiry" consists of two separate, identifiable parts: "teaching by inquiry" and "science as inquiry." These are best viewed as the process and the product of what might occur in a science classroom. Teaching by inquiry involves the means by which students gain knowledge. It includes the development of inquiry skills, such as the abilities (a) to identify and define a problem, (b) to formulate a hypothesis, (c) to design an experiment, (d) to collect and analyze data and (e) to interpret data and draw meaningful conclusions. Science as inquiry extends the image of science beyond that of a collection of facts, to include viewing science as a method by which facts are obtained (Eltinge and Roberts 1993). In this study, we focused on the process skills of scientific inquiry because our interest was on classroom inquiry activities to reveal the features of enacted inquiry.

Curriculum Enactment

This present study adopts an enactment framework instead of fidelity of implementation. Fidelity of implementation expects classroom teaching to follow step-by-step procedures and instructions from curriculum standards. On the other hand, the enactment assumes that textbook development and classroom instruction are creative and reflective processes. For enactment, it is not necessary or even impossible to demonstrate strict fidelity to the materials in order to be judged consistent with the intent of reform documents (i.e., curriculum standards). Instead, variations in enactments that meet student learning needs are considered reflective of reform and consistent with the intent of the materials (Schneider et al. 2005; Mcdonald and Songer 2008).

As for the factors influencing classroom inquiry enactment, a previous research has found that the most critical factor influencing a prospective teacher's intentions and abilities to teach science as inquiry is the prospective teacher's complex set of personal beliefs about teaching and views of science, wherein a prospective teacher's personal view of teaching science as inquiry is comprised of his or her knowledge of scientific inquiry and of inquiry-based pedagogy and his or her beliefs of teaching and learning (Crawford 2007). In a study of exemplary secondary science teachers (Breslyn and McGinnis 2012), the discipline (biology, chemistry, earth science and physics) in which a teacher taught was found to be a major factor on teachers' conceptions and enactment of inquiry. Two other contextual features of the classroom influencing enactment were curriculums and student abilities.

One of the reasons for the lack of inquiry in the science classroom is textbook portrayals of science as a collection of facts rather than as a process of inquiry (Eltinge and Roberts 1993). Germann et al. (1996) conducted a study to determine the degree to which the major high school biology laboratory manuals have portrayed inquiry. The results of the study indicated that the examined nine popular manuals seldom provided opportunities for students to pose a question to be investigated; formulate a hypothesis to be tested; predict experimental results; design observation, measurement and experimental procedures; work according to their own design; or formulate a new question or apply an experimental technique based on the investigation they performed.

Teachers' use of textbooks can also have an effect on student learning (Eltinge and Roberts 1993). For example, Forbes and Davis (2010) found that pre-service elementary teachers frequently added or substituted new elements into the curriculum materials they used and suggested that future research on pre-service teachers' use of curriculum materials should also characterize how these lessons with added elements actually play out in elementary classrooms.

Methods

Conceptualization of the Current Study

Scientific Inquiry in Chinese Curriculum Standards

The *Science Curriculum Standard for Grades 3–6* (Chinese Ministry of Education 2001a) states that "Scientific inquiry is the core of science learning. Inquiry is both a learning goal and a learning method." In the standard, scientific inquiry is described as one of the curriculum goals, with the other two goals being attitude & views, and scientific knowledge. Scientific inquiry activities are specified as the following elements: asking question, making hypothesis, making plan, conducting observation & experiment and making artifacts, collecting information, drawing conclusion and communicating.

Table 1 Scientific inquiry activities

Science standard (Grades 3–6)	Chemistry standard (Grade 9)	Inquiry activities included in this study
Asking question	Asking question	Asking question
Making hypothesis	Making hypothesis	Making hypothesis
Making plan	Making plan	Making plan
Conducting observations & experiment and making artifacts	Conducting experiment	Collecting evidence
Collecting information	Collecting evidence	
Drawing conclusion	Drawing conclusion	Drawing conclusion
Communicating	Reflecting	–
–	Communicating	–

The *Chemistry Curriculum Standard for Junior High School* (Chinese Ministry of Education 2001b) states that "scientific inquiry is not only an important method of learning but also major content." In the standard, scientific inquiry is presented as one of the five main content topics, with the other four topics being chemistry substances in daily life, the structure of substance, chemical reactions, and chemistry & society. The standard states that chemistry curriculum aims to promote student understanding of scientific inquiry processes and methods and to foster students' competence in scientific inquiry. It articulates scientific inquiry procedures to be composed of the following eight elements: asking question, making hypothesis, making plan, conducting experiment, collecting data, drawing conclusion, reflecting and communicating.

It is evident that inquiry activities identified in the two standards are quite similar. The two standards claim that only one or several activities of scientific inquiry (that is, elements of inquiry) may be involved in classroom teaching during a certain time. However, the 7 or 8 inquiry elements are too many to make detailed analysis about enactment, and some of them are overlapped with each other. To keep the clarity of the current study, considering the common properties of the elements, we combine "conducting observation & experiment and making artifacts" with "collecting information" to be "collecting evidence" for elementary grade level, and integrate "conducting experiments" with "collecting data" to be "collecting evidence" for middle school grade level. Although "communicating" (both elementary and middle school levels) is an important inquiry element, it does not show up as an independent inquiry activity in this study because

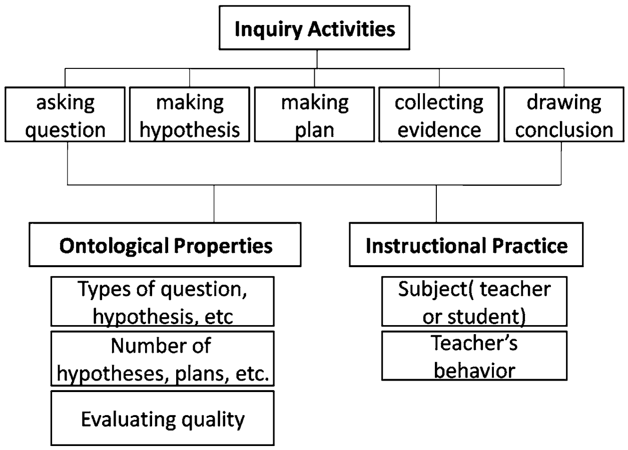


Fig. 1 Structure of coding scheme for enacted inquiry

communication always permeates in all other inquiry elements. A similar situation also applies to the element "reflecting." Table 1 shows the specific activities of scientific inquiry in the two standards and the inquiry activities included in this study.

Coding Scheme for Enacted Inquiry

To address the first research question, the coding scheme for each inquiry activity mainly involves two parts called *Ontological properties* and *Instructional practice*. Figure 1 shows the structure of the coding scheme. "Ontological properties" refers to the types of question, hypothesis, plan, evidence and conclusion, number of them in one inquiry activity and quality evaluation. For instance, the scientific questions proposed in class could be classified as the following categories: (1) relation question, which will lead to exploring the relationship between two objects; (2) explanation question, which will lead to seeking the explanation for specific phenomena; (3) description question, which directs to just describing objects or phenomena; (4) evaluation question, which will lead to evaluating something; and (5) designing question, which will direct to make design. Evaluating quality refers to evaluating the quality of questions, plans, etc. With regard to scientific question, quality evaluation indicates evaluating whether they are testable. Plan, hypothesis and so on sometimes get involved in the problem about number of them in one activity. "Instructional Practice" is related to the behaviors such as stimulation and scaffolding provided by teacher and the subject of the activity.

Examine the Factors that Influence the Enactment

The current study adopts a naturalistic approach to explore the factors influencing the enactment by placing the emphasis on teachers' delivery of the topic and the materials used to plan the topic instruction. The second author conducted a 30-min-long pre-topic teacher interview before start of the topic unit teaching and a 30-min-long post-topic teacher interview after each teacher finished their topic unit teaching. The protocols of interviews have been shown in the Appendix at the end of this article. Pre-topic interview focused on teacher's understanding of the topic, how the teacher decided what to teach and decisions about how to teach the topic. Post-topic interview included how well teacher achieved the instructional goals, the extent to which assessment influences the planning and implementation of teaching. The teacher questionnaire was administered to investigate the instructional objectives and the extent to which the student meets the learning goals. Teachers completed the questionnaire after each lesson of the topic. This part just provides qualitative evidence to claims about the factors.

Cases Introduction

To reveal the characteristics of enacted scientific inquiry at different grade levels and identify the factors influencing the enactment, the methods we employed combine a naturalistic approach (Wu and Krajcik 2006). Two competent teachers were purposefully selected.

Ms. CAI, Grade 4 science teacher, taught middle school biology and elementary school science for 1.5 years, in a newly established top-tier elementary middle school in Beijing. She received her M.A. degree in biological education, with a certification in middle school (biology).

Ms. LIU, Grade 9 chemistry teacher, who had more than 10 years of teaching experience in middle school and high school chemistry, has taught ninth-grade chemistry for 3 years in a top-tier middle and high school in Beijing. She held a certification in Grades 9–12 (chemistry).

Data Collection

The data collected for each case included 4 class periods of videos in specific topic units (Air of Grade 4 and Acids and Bases of Grade 9), teaching planning materials, students' worksheets and assessment materials. The second author made pre-topic and post-topic interviews with each teacher. As described in previous section, the two teachers completed the questionnaire which aimed to investigate their understanding of subject content, planning of teaching, understanding of students' learning and the assessment to be used. Data for Grade 9 were collected in March 2011, while data for Grade 4 were collected in December 2011.

The two exemplary teachers in this study were from Beijing city. Similar to most other areas in China that are under the same direction of new standards, science curricula of general education in Beijing consist of three parts: integrated science in Grades 3–6, traditional separated science subjects (physics, chemistry, biology and geography) in Grades 7–9 and traditional separated subjects in senior high schools (Grades 10–12). Junior high school students are required to take chemistry subject for 1 year in Grade 9.

Teaching Topics

The Air topic in Grade 4 and the Acids and Bases topic in Grade 9 are focuses of this study, since they both belong to physical science so that we can avoid the evident difference of disciplines. Table 2 is an overview of the Air unit (4S refers to Grade 4; L01 refers to the first lesson) by the lesson sequence as well as the lessons included in this study. Table 3 is an overview of the Acids and Bases unit by the lesson sequence as well as the lessons included in this study. Each lesson period lasts about 40 min in Beijing. Moreover, there are 36 and 40 students in Ms. CAI's class and Ms. LIU's class.

Data Analysis

To examine the classroom practice, several steps were taken. First, a detailed summary of each videotape was prepared, which included the teacher's and students' activities and conversations. Second, we coded the episodes on the videotapes that involved scientific inquiry activities using the coding scheme described previously with the software *Nvivo 8*. The durations of the episodes are not identical because they depend on the length of each inquiry activity. In relation to the second research question, the interview transcripts and teachers' responses to the questionnaire were repeatedly checked to find out the factors influencing teacher's enactment.

Table 2 Overview of the Air unit (Grade 4)

Lesson number	Content	Whether or not included in this study
4S-L01	Composition of air in the life	Yes
4S-L02	What is oxygen	Yes
4S-L03	What is carbon dioxide	Yes
4S-L04	Combustion	Yes

Table 3 Overview of Acids and Bases unit (Grade 9)

Lesson number	Content	Whether or not included in this study
9S-L01	Electrolyte (1)	No
9S-L02	Electrolyte (2)	No
9S-L03	Indicators for acids and bases	Yes
9S-L04	Properties of sulfuric acid	Yes
9S-L05	Properties of acids	Yes
9S-L06	Properties of bases	No
9S-L07	Neutralization reaction	Yes

Reliability

Two science education graduates (the second and fourth authors) observed and coded the science lessons and chemistry lessons. *Nvivo 8* was used to record the coding and make comparison between the two coder's coding, and it calculated kappa coefficient for each coding category to be the parameter of consistency. The final percentage of the inter-rater agreement ranged from 0.71 to 0.80.

Results

Scientific Inquiry Enacted in Science Classrooms

We identified the inquiry activity episodes involving inquiry activities in the filmed four lessons of each teacher and coded them according to the coding scheme to get the frequencies for each inquiry activity and the frequencies of the activities that include the corresponding subcategories of the inquiry activity. In what follows, we will present the frequencies along with the specific coding scheme of each inquiry activity.

Ms. CAI

At the beginning of the first lesson in the unit, Ms. CAI (Grade 4 teacher) posed several questions to direct students review what scientific inquiry is: How do scientists work? What steps constitute scientists' research process? Through a short story about Galileo's research on falling objects, she illustrated that the process of inquiry includes posing question through observation, formulating hypothesis, collecting evidence by conducting experiment or looking for information and drawing conclusion through evidence analysis.

Throughout the unit, the five inquiry activities were carried out, and the two activities "asking questions" and "drawing conclusions" occurred the most frequently (Table 4).

Asking Question

The *Ontological properties* of *Asking Question* consists of *type of question* and *quality of question*. The *Instructional*

Table 4 Frequencies of each inquiry activity in Grade 4 unit

Inquiry activity	Number of inquiry activity episodes			
	4S-L01	4S-L02	4S-L03	4S-L04
Asking question	7	6	9	5
Making hypothesis	3	2	1	1
Making plan	6	2	4	3
Collecting evidence	3	6	5	3
Drawing conclusion	5	8	7	4

practice includes *subject of posing question* (who proposes question) and *teaching practice*. Table 5 displays the specific coding categories of each aspect and the frequencies of all the categories in each lesson. In this unit, most of the research questions were formulated by Ms. CAI, such as what is air like according to your observation (4S-L01, description), what method can be used to collect air (4S-L01, design), how to prove that whether the bottle is fully filled with the CO_2 collected (4S-L03, design), why did the left candle extinguished, while the other right candle stayed burning (4S-L03, explanation) and whether is it good to have more and more oxygen (4S-L02, evaluation). The types of these sample questions are demonstrated in the parentheses.

Nevertheless, Ms. CAI also provided students opportunities to ask scientific question. Segment 1 shown in the following is taken from the first lesson (4S-L01) of the unit[In all the segments, the symbol "T" refers to the teacher, "S" refers to students of the whole class, and "S + Arabic numeral" refers to a certain student.].

Table 5 Frequency of asking question in Grade 4 classrooms

Aspects	Categories	4S-L01	4S-L02	4S-L03	4S-L04
Type of question	Relation question	0	0	0	0
	Explanation question	2	0	2	1
	Question directed to description or observation	2	4	3	0
	Question directed to evaluation	0	1	0	0
	Question directed to design	3	1	3	4
Quality of question	Teacher guided students to evaluate the questions	1	0	0	0
Who proposes question	Teacher	6	6	8	5
	Student	1	0	1	0
Teaching Practice	Teacher encouraged more than 3 students to propose question	1	0	1	0
	Teacher stimulated students to ask question	1	0	1	0

Teacher asked "what research questions can you formulate on the basis of your observation" so as to initiate students asking questions. Students tend to pose questions on "why" mostly, for instance, "why do we need air." Then, the teacher tried to stimulate students to evaluate these questions proposed by the students, but clearly they had no idea about the criteria for the investigable questions. In the third lesson (4S-L03), students had one more opportunity to raise questions as shown in Segment 2. Most questions students asked were assigned to the types of "explanation" and "prediction." Some questions were related to the composition and preparation of substance, while relation questions were involved (e.g., what is the difference between CO_2 and O_2?). It is evident that Ms. CAI stimulated students to formulate questions in a similar manner in segments 1 and 2 and encouraged more than 3 students to formulate question in Segment 1.

Segment 1

T On the basis of your observation, what questions can you propose?

S [students expressed their questions one by one] Why can't we see and touch the air? Why is the air odorless? How is the air formulated? How does the air occupy space? Is air invisible at any time? Why is air soluble in water? What constitute the air? Why can't we live without air? Can we touch the air?

T Now, can you identify which question is investigable and the most basic question?
(Students had no idea about this, keep silent.)

T "What is the composition of the air" is the most basic question. Long time ago, people viewed air as consisted of only one gas. Here we will come to the composition of air

Segment 2

T When we conduct science investigation, after observation we need to propose research questions. In regard to carbon dioxide, can you pose investigable questions?

S Why do people breathe out CO_2? What's the difference between CO_2 and O_2? Why CO_2 make people faint? How is the carbon dioxide formulated? What will happen for the concentrated carbon dioxide? What's the constituent of carbon dioxide?

Making Hypothesis

The ontological aspects of hypothesis include type, number and representation of hypothesis. *Type of hypothesis* involves two main categories, that is, *hypothesis on explanation* and *predicting phenomena. Number of hypothesis* refers to the

Table 6 Frequencies of making hypothesis in Grade 4 class

Aspects	Categories	4S-L01	4S-L02	4S-L03	4S-L04
Type of hypothesis	Hypothesis on explanation	1	0	0	1
	Predicting phenomena	2	2	1	0
Number of hypothesis in one activity	1 hypothesis	2	1	0	0
	2 hypotheses	0	0	1	0
	3 or more hypotheses	1	1	0	1
Representation of hypothesis	Written language on worksheet	3	1	0	1
	Communicate in verbal language	3	2	1	1
Who makes hypothesis	Teacher	1	0	0	0
	Student	3	2	1	1
Teaching practice	Teacher provides scaffold for making hypothesis	2	0	1	0
	Teacher stimulates students to formulate hypothesis	1	2	0	1

number of hypothesis in each inquiry activity which is identified as "making hypothesis." *Representation of hypothesis* includes written language and oral communication. The instructional aspects include the subject of the activity and teaching practice. Table 6 shows the coding framework of making hypothesis and the frequencies of the categories in each lesson.

Throughout the unit, Ms. CAI provided students opportunities to formulate hypothesis by raising questions such as "What will happen to the candle" "Why did the candle extinguish." Corresponding to these questions, the type of hypothesis involved explanation and prediction. Segment 3 from lesson 4S-L01 showed that Ms. CAI motivated students to make hypotheses about what would happen to the burning candle if it is covered with a glass bottle. The teacher made a talk with students to help them formulate the hypothesis whether "keep on burning" or "extinguished" and then asked students to write down the hypothesis on the worksheet (Fig. 2), which was developed to support student learning.

Segment 3

T (holding a glass bottle) There is air inside, right? Now, if the burning candle is covered, what will occur on the candle? Read the first table (Fig. 2), write down your guess or hypotheses in the first column of the table. Make a guess at the possible phenomena. We call the left candle as No. 1, and the right one as No. 2. What will happen to the burning No. 1 candle without treatment? Keep on burning. Then if we cover No. 2 with a bottle, what will happen?

S Extinguished

T Okay, write down your hypotheses

After the inquiry activity "making hypothesis" shown in Segment 3, Ms. CAI carried out the experiment to treat the two candles, directed students to observe that No. 2 candle extinguished in a while. Subsequently, the class came to another "making hypothesis" shown in Segment 4, which is assigned to be explanation. Ms. CAI stimulated students to make hypotheses on the reason why the candle burned out. Ms. CAI required students to fill their hypotheses in the worksheet (Fig. 2), and after that, she asked the students who raised their hands to communicate their opinion and organized a discussion. In this activity, more than one hypotheses were involved, and teacher encouraged more than one student to communicate their claims.

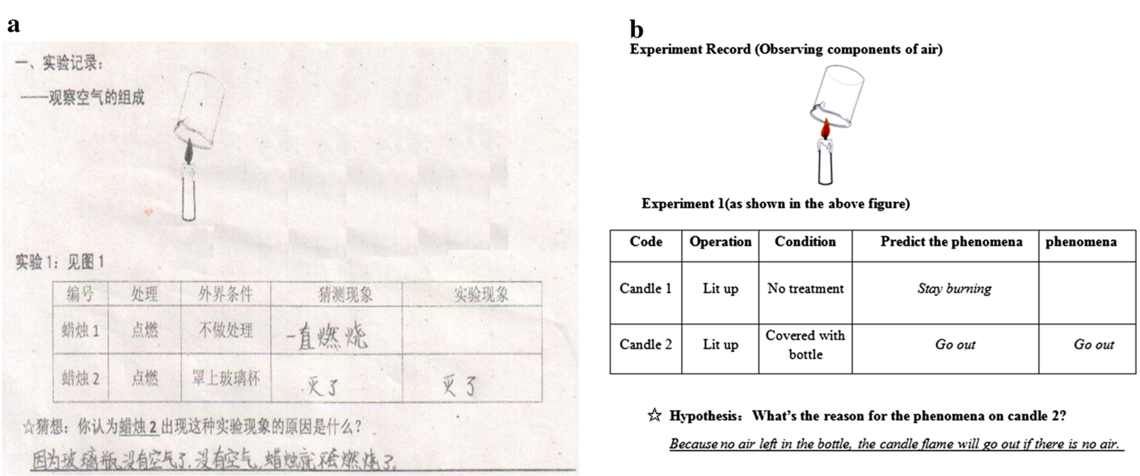


Fig. 2 A student's worksheet in lesson 4S-L01: **a** a scanning copy; and **b** a translated reproduction

Segment 4

T Why did the candle flame go out? Do you have any assumption? Write down your hypothesis on your worksheet. If you finish it, and want to share your opinion, please raise your hand
[After students finished the work, the discussion started]

T Finished? Okay, share your hypotheses with us

S1 Because fire needs air to keep on burning, while fire consumed air. When covered with the bottle, there is no air getting into the bottle, the fire extinguished because it's short of air. The candle outside can get air continuously, thus it keeps on burning

T Your hypothesis is long. Can you summarize it in a sentence?

S1 If there is no air, the fire can't stay lit

T Because there is no air. OK, what's your opinion?

S2 Maybe air is consisted of more than one kind of gas, while only part of the gas support combustion

T You mean, do you think there is air in the bottle?

S2 yes, there is. But there is no combustion-supporting air

T It's also a good hypothesis. Think about how to prove it

S3 I agree with S2. Ms. CAI collected a bottle of air, when it covered the candle, there is air. But it's possible that only oxygen in the air support combustion, there is air but no oxygen

T We know there is combustion-supporting gas. Just now, we said that air is combustion-supporting, but now is there air in this bottle?

S Yes, there is

T Who agree? [all students raised hands]

T Who hold the idea that there is no air left inside the bottle? [none of students raised hand]

Making Plan

The ontological aspects of making plan involve quality of plans, number of plans and evaluating plan. The instructional practices of making plan include the subject of the activity and teaching practice. Across this unit, Ms. CAI provided opportunities for student to making plans, and she also designed plans.

Segment 5 in what follows was taken from lesson 4S-L03. In this segment, Ms. CAI motivated students to design experiment to collect carbon dioxide and allowed several students to communicate their plans. Thus, more than one plan was involved to get carbon dioxide, and most of students' plans were feasible, but incomplete with a lack of apparatus elements. Ms. CAI advised to get carbon dioxide from air and guided students to evaluate this plan. Finally, teacher proposed two feasible and complete plans—one for laboratory and the other for family experiments. Both students and teacher made plans, and the process involved evaluating the plans (Table 7).

Segment 5

S1 Sometimes there is bubble in cola, and it tastes stimulating. That is carbon dioxide

T You mean, there is CO_2 in cola

S2 I don't agree with her. CO_2 was pressed into cola with high pressure. Because, CO_2 will dissolve in water when there is high pressure

T Her opinion is that we could get CO_2 from cola, is it right? So, it's okay

S3 We can blow up a balloon with mouth

T Excellent. The gas we breathe out also contain CO_2. We also know there is CO_2 in the air

S Yes

T Is the percentage of CO_2 in air high or low?

S low

T If we see air as a round plate, nitrogen occupies the largest area, oxygen takes the second large area, while the rest gases only have a percentage of 1 %. These gases include water vapor, rare gas, and CO_2. That means, whether the percentage of CO_2 is high or low?

S low

T Scientists found that the percentage of CO_2 is 0.03–0.04 %. Thus, it's hard to collect CO_2 from the air. [Then, teacher introduced the methods to prepare CO_2 including the materials and apparatus used.]

Table 7 Frequency of making plan in Grade 4 classrooms

Aspects	Categories	4S-L01	4S-L02	4S-L03	4S-L04
Quality of the plans made by students	Infeasible	0	0	1	1
	Feasible but not intact	1	1	2	0
	Intact but not feasible	1	0	0	1
	Intact and feasible	3	1	2	2
Number of plans	Only one plan	4	1	2	1
	2 plans	2	0	2	1
	3 or more plans	0	1	1	1
Evaluating plans	Teacher guide students to evaluate plan	2	1	3	2
Who makes the plan	Teacher	3	1	3	1
	Students	3	1	3	2
Teaching practice	Teacher encourages 2 students to communicate	1	0	2	0
	Teacher encourages more than 2 students to communicate	2	1	1	2

Table 8 Frequency of collecting evidence in Grade 4 class

Aspects	Categories	4S-L01	4S-L02	4S-L03	4S-L04
Type of evidence	Describing phenomena	3	1	5	3
	Generalizing phenomena	0	3	0	0
	Extra information	0	2	0	0
Quality of evidence	Teacher guides students to evaluate whether the evidence matches the hypothesis	1	1	1	1
The subject and the approach to collecting evidence	Teacher guides student watching videos	0	4	1	0
	Teacher adds extra information	0	2	4	0
	Teacher does experiments demonstrating to students	3	0	0	2
	Student gets evidence from previous experiment	0	0	0	1
	Students' hands-on experiment in groups	0	0	0	0
Teaching practice	Teacher stimulates students to collect evidence	1	0	0	1
	Teacher directs students to observe and record phenomena	3	4	4	2

Collecting Evidence

Table 8 presents the categories and frequencies of *Collecting Evidence*. The ontological aspects of *Collecting Evidence* involve type of evidence and quality of evidence. The *type of evidence* includes the following categories: (a) Describing phenomena; (b) Generalizing phenomena; and (c) Extra information. The *instructional practice* involves the subject and the approach to collecting evidence and teaching practice. The *teaching practice* focuses on teacher's stimulation and guidance to students. In relation to the types of evidence, Ms. CAI' s class mainly involved "Describing phenomena" throughout the topic unit, while "Generalizing evidence" and "Extra information" just occurred in 4S-L02. Furthermore, Ms. CAI guided students to evaluate whether the evidence matches the hypothesis in each lesson of the unit. Teacher's demonstration, extra information and videos were predominantly the approach to collecting evidence. Besides, teacher stimulated students to collect evidence in 4S-L01 and 4S-L04, and frequently directed students to observe and record the phenomena.

For instance, when investigating combustion condition (4S-L04, Segment 6), Ms. CAI stimulated students to seek evidence from previous experiments in this unit to prove oxygen is a necessary criterion. With the guidance of teacher, students described the process and phenomena of those experiments, which is assigned to be "describing phenomena." After the first student (S1) provided his evidence to the question, teacher directed students to evaluate the fitness between hypotheses and evidence.

Segment 6

T How can we prove that oxygen is necessary for combustion? Think about the experiments we carried out previously

S1 When the burning stick was placed on the top of a bottle filled with carbon dioxide, it burned out

T We found the stick burned out in the bottle full of carbon dioxide, when we did experiment to test carbon dioxide. This just indicates that carbon dioxide can't support burning, but it doesn't suggest that oxygen is necessary

S2 We did an experiment, lit up the candle, and covered it with a bottle

T That experiment suggest, when we lit up two candles, covered one of them with a bottle, we found, which candle extinguished?

S The one covered

T Yes

S2 No oxygen left

T It extinguished because there was no oxygen left. That means we proved combustion needs oxygen

In lesson 4S-L02, Ms. CAI used video to demonstrate the experiment: put a lit candle into the bottle filled with oxygen and asked students to record the phenomena. Then, teacher summarized the different phenomena of burning in oxygen, nitrogen and air, which was assigned to be *generalizing phenomena*. In lesson 4S-L03, after the discussion about the method to confirm whether the bottle is full of carbon dioxide, teacher conducted the experiment and demonstrated the process to the students. She led students to pay attention to the phenomena and told students to write down "the burning sticks extinguished" in the worksheet.

Drawing Conclusion

The aspects and categories of *Drawing Conclusion* and frequency of each category are shown in Table 9. The ontological aspects include *type of conclusion*, *approach to drawing conclusion*, *fitness between evidence and*

Table 9 Frequency of drawing conclusion in Grade 4 class

Aspects	Categories	4S-L01	4S-L02	4S-L03	4S-L04
Type of conclusion	Explanation	1	0	1	0
	Relation	0	0	0	1
	Inference	4	8	6	3
Approach to drawing conclusion	Observation	2	5	3	0
	Generalization	0	2	0	1
	Phenomena-based reasoning	3	1	4	3
	Comparison	0	0	0	0
Fitness between evidence and conclusion	Teacher as subject	1	0	0	0
	Students as subject	1	0	0	0
Who draws conclusion	Teacher	3	7	5	3
	Students	2	1	2	1
Teaching practice	Teacher encourages more than 2 students to communicate the conclusion	2	0	2	0
	Stimulating	2	2	2	0
	Directing to record	0	5	3	0
	Summarizing	3	2	2	4

conclusion. *Type of conclusion* involves the following categories: (a) Explanation, which means that the conclusion is to answer "why" question; (b) Relation, the conclusion is to identify relationship between some objects; (c) Inference, which refers to further inference based on evidence. *Approach to drawing conclusion* includes the following: (a) observation, drawing conclusion by directly observing single phenomena; (b) generalization, generalizing series of phenomena or experiments; (c) phenomena-based reasoning, reasoning the conclusion by combination of knowledge and phenomena; and (d) comparison, comparing different phenomena to answer the question like "which one is better."

As shown in Segment 7 (in lesson 4S-L01) below, Ms. CAI motivated students to draw conclusions on properties of nitrogen by asking "What is nitrogen like through your observation of air? Why can you draw the conclusion, or why can't?" The conclusions on color and odor of nitrogen were assigned to be inference for the type of conclusions and observation for approach to drawing conclusions. Ms. CAI encouraged more than two students to communicate their conclusions.

Segment 7

S1 I see nothing (in the air), it's proved that this gas is colorless, and no reflection of light, because only when object reflects light we can see it

T good

S2 I don't agree with him. Although other gases in air only have a proportion of 1 %, but they might react with nitrogen and make nitrogen colorless

T You mean that if chemical reaction occurred, we might not see the matter. Okay, the nitrogen we are talking about does not react with other matter

S3 We can't identify whether it is soluble in water

T you mean, we can figure out what is nitrogen like through observing the air. That is because—what is the proportion of nitrogen in the air?

S a large amount

T yes, so we could draw conclusions on the color and odor of nitrogen by observing air

For the type of conclusion, the conclusions in lesson 4S-L03, such as "carbon dioxide is colorless and odorless," "carbon dioxide doesn't support burning, and also nonflammable," "carbon dioxide is heavier than air," were originated from the phenomena; thus, they were assigned to be inference in regard to the types of conclusion. In lesson 4S-L01, with teacher's guidance, students drew the conclusion that part of the air was consumed by the burning candle, and the other part of air left in the bottle, which was direct to answer teacher's question "why did the water get into the bottle." Such a conclusion was assigned to be explanation. Teacher summarized the three conditions of combustion, which revealed the relation among them; thus, this conclusion was assigned to be relation.

With regard to *approach to conclusions*, *observation, generalization and phenomena-based reasoning* were involved to draw conclusions in Ms. CAI's class. In lesson 4S-L02, Ms. CAI generalized the phenomena of different matters burning in oxygen to conclude that oxygen can support burning and make the flammable thing burn more vigorously; thus, this conclusion is assigned to be generalization. In lesson 4S-L03, Ms. CAI concluded that carbon dioxide is heavier than air through the reasoning on the experiments' phenomena, but not just observation or generalizing, so it is assigned to be phenomena-based reasoning.

Summary

Ms. CAI placed emphasis on all the five inquiry activities in an explicit way to show students how to do inquiry and make students experience the inquiry process. She often used the words "make a guess," "how to improve" and "what conclusions you can draw" to engage students in the inquiry activities. In addition, worksheets were developed to help student make inscriptions for inquiry.

We also found that Ms. CAI divided the whole inquiry task (e.g., composition of air) symbolized by research questions teacher proposed into several sequenced inquiry activities. This manner is helpful to involve students in the inquiry process and advocate primary students to understand inquiry. Thus, we could conclude that classroom performance reflected that Ms. CAI viewed scientific inquiry itself as learning content. However, students were guided in most of the inquiry activities, for example, teacher expressed the conclusion clearly and then reminded students to write down.

Ms. LIU

Throughout the unit, Ms. LIU's class involved four of the five inquiry activities including asking question, making plan, collecting evidence and drawing conclusion. Table 10 shows the frequencies of inquiry activities in each lesson. What follows below will discuss the features of each enacted inquiry activity by coding the episodes with the categories of each aspect for the four activities and summarize the overall features of Ms. LIU's enactment of scientific inquiry.

The coding framework of each inquiry activity for Ms. LIU's Grade 9 class is mainly the same as that of Ms. CAI's Grade 4 class, so as to uncover the characteristics of enacted inquiry at different grade levels. As such, the specific meaning of the categories will not be repeatedly introduced in this part.

Asking Question

In this unit, all the research questions were posed by Ms. LIU, and most of them were questions directed to design (Table 11). Ms. LIU raised questions such as "How can we identify the acidity of soil"(9S-L03), "We could buy concentrated sulfuric acid from agent store, but dilute sulfuric acid is needed in lab, how do we make the concentrated into dilute sulfuric acid" (9S-L04), "Thinking from the theoretical perspective, how do you prove whether the liquid brought from home contain acid" (9S-L05), "How can we prove that hydrochloric acid reacts with sodium hydroxide" (9S-L07). The activities initiated by these questions were to design investigations, and they are coded as *questions directed to design*. The only *Relation question* "Does all the acids have the same degree of acidity, all the alkaline have the same alkalinity" was involved in lesson 9S-L03. The question "we can observe, what is concentrated sulfuric acid like" in lesson 9S-L04 that belongs to the *questions directed to description or observation* were observed.

Table 10 Frequencies of inquiry activities in Grade 9 class

Inquiry activities	9S-L03	9S-L04	9S-L05	9S-L07
Asking question	2	2	1	1
Making hypothesis	0	0	0	0
Making plan	1	1	1	1
Collecting evidence	1	3	2	1
Drawing conclusion	2	4	3	1

Table 11 Frequency of asking question in Grade 9 class

Aspects	Categories	9S-L03	9S-L04	9S-L05	9S-L07
Types of proposed question	Relation question	1	0	0	0
	Explanation question	0	0	0	0
	Question directed to description or observation	0	1	0	0
	Question directed to evaluation	0	0	0	0
	Question directed to design	1	1	1	1
Quality of question	Teacher guides students to evaluate the questions	0	0	0	0
Who proposes question	Teacher	2	2	1	1
	Student	0	0	0	0
Teaching practice	Teacher stimulates students to ask question	0	0	0	0
	Teacher encourages more than 3 students to propose question	0	0	0	0

Besides, the proposed questions served as the driving question in lessons 9S-L03, 9S-L05 and 9S-L07. That is, the whole lesson was organized around each of the questions. This is different from the situation of Ms. CAI's class in which many questions were involved in a lesson.

Making Hypothesis

There was no explicit *making hypothesis* across the four lessons. That is, teacher did not stimulate students to make hypothesis. However, in the process of making plans, hypothesis might be implicitly permeated in student's thinking. For instance, in the lesson 9S-L05, Ms. Liu asked students to bring some liquids which may contain acid to classroom from home. When a student choosing the liquid, it might involve "making hypothesis" about which one is acidic. In addition, when the students designing plans to prove whether the liquid contains acid, hypotheses were also involved implicitly, for instance, they may thought about that "if there was acid, the liquid would turn red when adding litmus, or bubbles would arise in the solution when active metal was added to the liquid."

Making Plan

Table 12 displays the coding categories and frequencies of making plan in this unit. It indicates that "making plan" was involved in each lesson.

In lesson 9S-L03, Ms. LIU required students to bring some flowers, vegetables or fruits to the class which would be used to produce acid–base indicator. After introduction to what acid–base indicator is, Ms. LIU told students that the subsequent activity was to make acid–base indicator and to do experiments to observe the color of acid and base when litmus or phenolphthalein was added. Apparatus and operation steps were also introduced by Ms. Liu in the statement of the activities. That is, teacher constructed the plan for this experiment. However, students had the freedom to choose the materials, such as rose, apple, red cabbage.

According to Ms. LIU's response in the questionnaire, lesson 9S-L03 was aimed to provide experimental foundation for learning properties of acid and base, and to apply the concepts that were taught previously to daily life context. To some extent, this orientation could explain why not students but the teacher made plans in this lesson.

In lesson 9S-L04, Ms. LIU led students to think about how to make concentrated sulfuric acid into dilute sulfuric acid. First, the teacher and students discussed the two ways—adding water to sulfuric acid or adding sulfuric acid to water. Ms. LIU used a video to demonstrate the phenomenon that concentrated sulfuric acid was splashed out when water was added to it. Ms. LIU guided students to explain the cause of the phenomena with a common phenomenon in daily life. Finally, it came to the correct manner for diluting sulfuric acid (Segment 8). The whole process was progressed in the dialogue of teacher and students, but the plan was made by teacher.

Table 12 Frequency of making plans in Grade 9 class

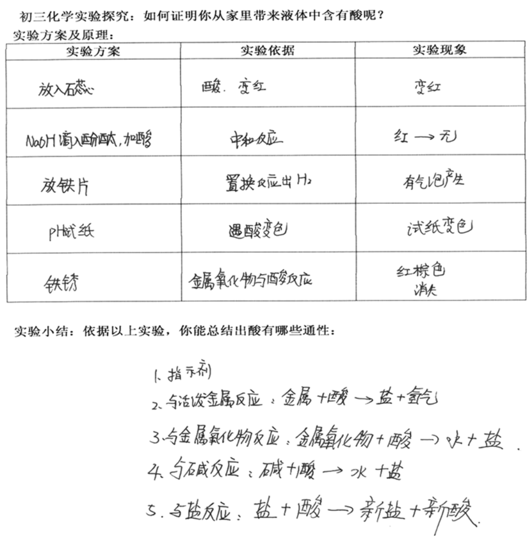
Aspects	Categories	9S-L03	9S-L04	9S-L05	9S-L07
Quality of the plans made by student	Infeasible	0	0	0	0
	Feasible but not intact	0	0	0	1
	Intact but not feasible	0	0	0	0
	Intact and feasible	1	1	1	0
Number of plans	Only 1 plan	1	1	0	0
	2 plans	0	0	0	0
	3 or more plans	0	0	1	1
Evaluating plans	Teacher guide students to evaluate plans	0	0	1	1
Who makes the plan	Teacher	1	1	0	0
	Students	0	0	1	1
Teaching practice	Teacher encourages 2 students to communicate	0	0	0	0
	Teacher encourages more than 2 students to communicate	0	0	1	1

Segment 8

T If you are a teacher who prepares agents for chemistry class, you will face such a problem: chemicals agent stores just sell concentrated sulfuric acid, but the laboratory needs dilute sulfuric acid. How can we make the concentrated sulfuric acid into dilute sulfuric acid?
T How many manners for diluting?
S Two
T Add water to acid, or add acid to water. In regard to concentrated sulfuric acid, which one is suitable?
S Add acid to water
T Let's watch a video (the video showed when adding water to concentrated sulfuric acid, and the acid was splashed out). What phenomenon is similar to this one?
S Water boiling
T When adding water to a pot filled with hot oil, water will spill out. The difference between oil and water includes density and boiling point. When the oil is hot, even if it is not boiling, the water is boiled and would spill out with oil. We can use this example to explain the phenomenon in the video. Heat is released when concentrated sulfuric acid is dissolved in water, and it is at the lower phase. Water at the upper phase, boiled by the heat, and splashed out with concentrated sulfuric acid. And we just mentioned, what property does concentrated sulfuric acid demonstrate?
S Corrosivity
T Hygroscopicity, strong oxidbillity, and corrosivity. If it is splashed out, it will be dangerous

In the lessons 9S-L05 and 9S-L07, students played as the subject to make plans, and more than one plan was designed to answer the research questions. Most of the plans designed by students were feasible, and some were not intact enough. Ms. LIU guided students to evaluate or revise plans. Segment 9 is taken from lesson 9S-L05, in which students made plans. Ms. LIU raised the question "The acids in laboratory usually include hydrochloric acid and sulfuric acid, but there are many acids in daily life. You have brought many liquids such as vinegar, lemon juice, and detergent. Now, please think about, how to prove that the liquid you brought contains acid(s)?" to engage students in making plans. Ms. LIU required the students to write down the key words of their plans on the worksheet

a



初三化学实验探究：如何证明你从家里带来液体中含有酸呢？

实验方案及原理：

实验方案	实验依据	实验现象
放入石蕊	酸，变红	变红
NaOH滴入酚酞，加酸	中和反应	红→无
放铁片	置换反应出H_2	有气泡产生
pH试纸	遇酸变色	试纸变色
铁锈	金属氧化物与酸反应	红棕色消失

实验小结：依据以上实验，你能总结出酸有哪些通性：

1. 指示剂
2. 与活泼金属反应：金属+酸→盐+氢气
3. 与金属氧化物反应：金属氧化物+酸→水+盐
4. 与碱反应：碱+酸→水+盐
5. 与盐反应：盐+酸→新盐+新酸

b

Inquiry with experiments: How do you prove that the liquid you brought contains acid(s)?

Plans and Principles:

Plans of Experiments	Rationales of the Plans	Phenomena Observed
Add litmus	*Acid turn red*	*Turn red*
Add Phenolphthalein to NaOH, then add Acid	*Neutralization reaction*	*Red → colourless*
Add iron sheet	*Displacement reaction occur, so H_2 is released*	*Bubbles arose*
pH test strips	*The colour change when the strip encounters acid*	*The test strip' colour changed*
Add iron rust	*Metallic oxide reacts with acid*	*The reddish brown colour faded*

To summarize: On the basis of the above experiments, the general properties of acids include:

1. Indicator

2. Reaction with active metal, Metal + Acid → Salt + Hydrogen

3. Reaction with metallic oxide, Metallic Oxide + Acid → Water + Salt

4. Reaction with base, Base + Acid → Water + Salt

5. Reaction with salt, Salt + Acid → new Salt + new Acid

Fig. 3 A student's worksheet for inquiry in lesson 9S-L05: **a** a scanning copy; and **b** a reproduction

(Fig. 3). What is important to know is that students have studied properties of sulfuric acid in classroom before and studied properties of hydrochloric acid on their own; thus, they developed plans independently at first before discussing with their group members.

In lesson 9S-L07, students discussed in groups to make plans on how to prove whether hydrochloric acid reacts with sodium hydroxide solution. After that, students reported their plans to the teacher and the class (see Segment 9). Some plans were feasible but not intact, and teacher led students to make evaluation and modification on the plans.

Segment 9

T Who wants to share your plans or your groups' plans?

S1 First, add hydrochloric acid to sodium hydroxide solution. Then add phenolphthalein after a moment

T Why, please illustrate your experiments' rationales

S1 Phenolphthalein is an acid–base indicator. It will become red in sodium hydroxide solution. If sodium hydroxide reacts with hydrochloric acid, because the amount of sodium hydroxide is small, so there will be none left, so phenolphthalein will be colorless

T He means that the reaction is at the first step, will the mixed solution's color change? It's ok, but is there any defect in this plan?

S1 Add excessive amount of sodium hydroxide

T For example, the amount of sodium hydroxide should be small or excessive. The sequence for adding agents. Who can help to modify it?

S2 First, add sodium hydroxide; Second, add phenolphthalein; Third, hydrochloric acid

Collecting Evidence

The coding categories and frequencies of each activity *Collecting Evidence* in this unit are given in Table 13. Collecting evidence was involved across all the four lessons, and all of them belong to the type *"Describing phenomena."* Students' group work of hands-on experiment was adopted in three of the four lessons. Videos and teachers' demonstration were also used as the approach to collecting evidence. Ms. LIU did not stimulate students themselves to collect evidence for hypothesis, but frequently directed students to observe and record the phenomena. The evaluation about whether the evidence matched the hypothesis did not occur explicitly.

Ms. LIU placed an emphasis on students' group work of hands-on experiments. The work was always organized after plans were adequately designed by teacher or students. When student groups were carrying out experiments, Ms. LIU reminded them to make experiment records, gave advice for their experiments and supervised the progress of their experiments.

When investigating properties of sulfuric acid in lesson 9S-L04, Ms. LIU used videos to demonstrate dehydration and hygroscopicity of concentrated sulfuric acid. Teacher led students to make observation and understand dehydration and hygroscopicity. Ms. LIU did an experiment to show the reaction of barium chloride and dilute sulfuric acid and then introduced the reaction to students through describing the phenomena. This activity is assigned to be *describing phenomena*.

Table 13 Frequencies of collecting evidence in Grade 9 class

Aspects	Categories	9S-L03	9S-L04	9S-L05	9S-L07
Type of evidence	Generalizing phenomena	0	0	0	0
	Describing phenomena	1	3	2	1
	Extra information	0	0	0	0
Quality of evidence	Teacher guides students to evaluate whether the evidence matches the hypothesis	0	0	0	0
The subject and approach to collecting evidence	Teacher guides student watching video	0	2	0	0
	Teacher adds other information	0	0	0	0
	Teacher does experiments demonstrating to student	0	1	1	0
	Student gets evidence from previous experiment	0	0	0	0
	Students' experiments in groups	1	0	1	1
Teaching practice	Teacher stimulates students to collect evidence	0	0	0	0
	Teacher directs students to observe and record phenomena	0	3	1	1

Throughout the four lessons, Ms. LIU always directed students to observe and record the experimental phenomena. When teacher's demonstration or video of experiments was adopted, Ms. LIU guided students to observe the phenomena. In the case of students' group work on hands-on experiments, Ms. LIU required students to make record on their worksheets. Students were always asked to report their experiments, phenomena and conclusions. Segment 10 shows an example of student groups' oral report in lesson 9S-L07, in which they illustrated their experiments to the class, which was used to prove whether hydrochloric acid reacted with sodium hydroxide in the solution.

Segment 10
(Students demonstrated the tubes used to do experiments, and illustrated their groups' experiments)

S1 (group 1)	This is the first plan. First, we added a lot of hydrochloric acid and phenolphthalein, the formed solution was colorless. Second, sodium hydroxide solution was added, maybe because it's a great quantity, the mixture looks deep red
S2 (group 1)	This is conducted according to the second plan. Sodium hydroxide solution was added first, then phenolphthalein, and hydrochloric acid last. The solution was red at the beginning and changed into colorless in the end
T	Is there somebody using pH papers?
S3 (group 2)	this pH paper was used to test the solution after reaction, that is the one in which no color change occurred. The solution is deep red
T	What's the pH value?
S3	The value is between 1 and 2
S3	This is the first one, the solution is red after reaction
S4 (group 3)	Our group adopted experimental comparison. This is the mixed solution of litmus and hydrochloric acid, it is red. This is the mixed solution of litmus and sodium hydrochloride solution, it is blue. When hydrochloric acid was mixed with sodium hydrochloride solution, it's purple
S5 (group 4)	We tested the temperature in the reaction process. 20 ml hydrochloric acid and sodium hydrochloride solution, it was 20 °C before reaction, 30 °C after the reaction

Drawing Conclusion

The categories and frequencies of each aspect are shown in the following Table 14. All the four lessons included the activity of drawing conclusion. All of the conclusions belong to the type "Inference." The approaches to drawing conclusion involved observation, generalization, phenomena-based reasoning and comparison. Both teacher and students played as the subject to draw conclusion. The teacher stimulated students to formulate conclusions, directed them to record conclusions and summarized the conclusions.

In lesson 9S-L01, students reported the color of litmus and phenolphthalein in hydrochloric acid, sodium chloride solution and sodium hydroxide solution according to their observation in group work. Ms. LIU generalized and reorganized the conclusion that students had reported as the color of litmus and phenolphthalein in acidic, basic and neutral solution. In this episode, the conclusion was drawn from the approach of observation and coded as inference for the aspect type of conclusion.

After students reported on the experiment of making acid–base indicator in the lesson 9S-L01, Ms. LIU motivated students to draw conclusions by asking "According to your reports, let's make a summary. Which material is suitable to be used to make indictors in family?" The

Table 14 Frequency of drawing conclusions in Grade 9 class

Aspects	Categories	9S-L03	9S-L04	9S-L05	9S-L07
Type of conclusion	Explanation	0	0	0	0
	Relation	0	0	0	0
	Inference	2	4	3	1
Approach to drawing conclusion	Observation	1	1	0	0
	Generalization	0	1	1	0
	Phenomena-based reasoning	0	3	1	0
	Comparison	1	0	1	1
Fitness between evidence and conclusion	Teacher as subject	0	0	0	0
	Students as subject	0	0	0	0
Who draws the conclusion	Teacher	1	3	3	1
	Students	1	1	0	0
Teaching practices	Teacher encourages more than 2 students to communicate the conclusion	0	0	0	0
	Stimulating	0	1	0	0
	Directing to record	1	1	1	0
	Summarizing	2	2	2	1

conclusions corresponding to this question were drawn from comparison and coded as inference.

Overall, Ms. LIU placed an emphasis on generalization of conclusions and always summarized the conclusions. Consequently, most of the conclusions were assigned to be "summarizing" for teaching practices and inference from the teacher (as shown in Table 14). Segment 11 was taken from the lesson 9S-L05 in which teacher summarized students' reports on experiments of testing the acidity of some liquids brought from home to formulate conclusions.

Segment 11

T According to your reports on the phenomena, considering the acidity, which one has the strongest acidity?
S The detergent, and lemon juice
T Considering the phenomena, which phenomena was the most evident? Which agent? For example, the bubbles in acid were very obviously visible
S pH test strips
T pH test strips, litmus. These methods resulted in obvious phenomena
T From laboratory to life in society, we can find that, although the acids are not the same, but the phenomena in our experiments were similar. Why?
T Different kinds of acids have some common properties because they can ionize to release H^+ in water

Summary

The enacted scientific inquiry in Ms. LIU's class was much more like problem-solving in authentic context. First, Ms. LIU usually set an authentic context to motivate the class and promote the progress of teaching activities. For example, at the beginning of lesson 9S-L07, to identify adulterated wine using the principle of neutralization reaction was demonstrated to students through TV show. Afterward, students were engaged in designing plans to prove whether hydrochloric acid reacts with sodium hydroxide solution. Then, Ms. LIU guided students to learn the types of neutralization reaction and to summarize the method for proving neutralization reaction. The class returned to the context of wine identification through the video of introducing rationales for the identification process.

Second, most of the research questions proposed by teachers were directed to design plans, and this is also evidence for the feature of problem-solving. Overall, the inquiry in grade 9 case placed more emphasis on inquiry-based teaching or learning rather than learning inquiry. The scientific inquiry was aimed to understand science ideas and apply science ideas in authentic contexts. In other words, Ms. LIU's class emphasized the inquiry skill "making plans" and dedicated to foster students' ability to plan investigations to solve the problems.

Factors Influencing Enactment of Scientific Inquiry

Teachers' Understanding About Inquiry

Through pre- and post-topic interviews with the two teachers and their responses to the questionnaire, we find that teachers' understanding of scientific inquiry is an important factor influencing the enactment of scientific inquiry, because their understanding can account for part of their enacted inquiry. In the questionnaire, there was a question "What was the main thing you wanted students to learn from today's lesson? Why do you think it is important for students to learn this?" Ms. CAI (Grade 4 Science teacher) answered this question after lesson 3S-L01 as follows:

> Students already had some common sense about air when reading science books in their rest time. Some students even knew the components of air. Thus, in addition to teaching them scientific knowledge "air consists of nitrogen, oxygen, carbon dioxide, water steam", I think teacher also needs to help students realize that scientific knowledge is not definite as what is written in the textbooks, but it is the patterns and rules which were generalized through constant investigations.

> The reason is that students should not just remember the conclusions when they study science, the most important thing for students is to understand how scientific conclusions are formed. On the basis of students' interest in science, I expect them to investigate scientific phenomena actively in future.

The interviews with Ms. LIU (chemistry teacher, Grade 9) indicate that she focused on the systematic structure of chemistry knowledge, students' ability to apply chemistry knowledge to solve authentic problems and their ability to design and conduct hands-on experiments. Her response to the same question in the questionnaire (9S-L07) was quite different from the answers of Ms. CAI:

1. the method used to prove neutralization reaction
2. the method used to confirm the reaction with no obviously visible change
3. Explain the application of neutralization reaction in life and society

> The reason is that these will provide foundation for neutralization titration, and promote students to apply chemistry knowledge and theories in community and society.

Other Factors Including Textbooks, Assessment, Students and Resources

The ninth-Grade teacher Ms. LIU said: "National curriculum standard, guideline of high school entrance exam and textbooks are the main materials which help me decide what to teach. I pay attention to the recommended inquiry tasks in the chemistry curriculum standard. Textbook is the most fundamental materials, because the province-responsible high school entrance exam is closely related to the textbook. We usually focus on the specific content in the textbook and integrate some exam item to them. I pay more attention to experiments in the textbook, that is, inquiry tasks. With regard to experiments, sometimes I try to transform the teacher's demonstrations into students' group work, or make some change and improvement on the experiments in textbook". Ms. LIU gave such a response to the question "how do you decide how to teach this topic": It depends on two aspects. The first one is the limitation of the content in the textbook, some are easy to design activities, some are difficult to be organized by activities. The second is about students diversity, students in a class may have several distinct features from students in another class. Although the main instruction activity is the same, I will make a little bit change in different classes.

The fourth-Grade teacher Ms. CAI described what to teach to the researcher (second author) according to the subtitles of the unit of the textbook. Other reference materials for her include the manual book for teachers and other teachers' planning on the topic, which can be found on the Internet. Furthermore, resource limitations such as shortage of apparatus also influence classroom enactment of scientific inquiry. For instance, in regard to "collecting evidence," all of the experiments in the case of Grade 4 were organized as teacher's demonstration rather than students group work.

Conclusions and Discussion

Characteristics of Enacted Inquiry at Different Grade Levels

In regard to the ontological properties of inquiry activities, the two cases covered kinds of questions, hypotheses, plans, evidence and conclusions. Both Grade 9 and Grade 4 cases involved "Questions directed to design"; however, it is devoid of "Explanation question" and "Questions directed to evaluation" in Grade 9 and there was less "Questions directed to description or observation" in Grade 9. According to the hierarchical categories of research questions used by Hasson and Yarden (2012), "Explanation question" and "Questions directed to evaluation" can be assigned to higher-order questions.

With respect to making hypothesis, the teacher in Grade 4 case provided opportunities for students to make hypotheses with different representation forms. For the activity "Making plan," the two case teachers allowed students to make more than one plan and directed them to evaluate the plans. With respect to collecting evidence, the Grade 9 teacher gave students opportunities to conduct experiments in groups. The two case teachers included the two types of evidence-describing phenomena and generalizing phenomena. In terms of drawing conclusion, both Grade 9 and Grade 4 cases covered different methods to draw conclusions, and most of the conclusions belonged to "inference."

The present study finds that the ninth-grade case lacked the opportunity for inquiry activity "Making hypothesis" and gave students less opportunity to posing research questions. However, the study also reveals that the ninth-grade chemistry class placed emphasis on engaging students in making design and conducting experiments to solve authentic problems, students are expected to use their knowledge in this process, and students' knowledge and conceptual understanding could be expanded. This can be a supplementary to Breslyn and McGinnis (2012)'s findings about chemistry teaching:

> the chemistry teachers in this study tended to enact inquiry with an emphasis on content knowledge, they

are less likely to enact items measured with the PII (The PII assesses the degree to which teachers engage their students in inquiry as defined by the National Research Council's Abilities Necessary to Do Inquiry). In this case, chemistry teachers were less likely than biology teachers to allow students a choice of questions to investigate, support students' use of questioning and discuss the use of hypotheses.

This study indicates that it is possible and important for teachers to involve different kinds of inquiry activities which belong to different categories of the ontological properties in classrooms in a topic unit or in a longer-term instruction. It shows a meaningful enactment of scientific inquiry by specifying it into such variant and abundant activities for students. What is important is that teachers need to lead students to evaluate quality of questions, fitness between hypothesis, evidence and conclusions, so that students would acquire critical thinking to do inquiry. Banchi and Bell (2008) suggested that there are four levels of inquiry-based learning in science education: confirmation inquiry, structured inquiry, guided inquiry and open inquiry. At the beginning of a new science subject curriculum, guided inquiry is necessary to show students how to do inquiry, just like the two cases in this study. As students progress on learning content knowledge and inquiry skills, open inquiry needs to be provided to them.

For the instructional practices in the enacted inquiry, the two case teachers set good examples to other teachers when enacting inquiry or conducting inquiry-based teaching. The behaviors such as stimulating students to actively pose research questions and engage in other inquiry activities, providing scaffolding for students especially for the early-year students and the difficult part of inquiry activities such as making plans for older students, asking students to record and report their questions, hypothesis, experiments phenomena and the conclusions, could be used to promote development of students' understanding about scientific inquiry and abilities to do inquiry. Requiring them to write down their opinions and encouraging more than 3 students to express their opinions are valuable to engage students in inquiry activities, especially for the situation that there are about 40 students in a class.

Furthermore, both the national science standard for elementary school and chemistry standard for ninth grade (Chinese Ministry of Education 2001a, b) call for scientific inquiry with similar identification of process skills. This study indicates that the scientific inquiry at elementary levels is to make students experience and understand the process of inquiry. As grade level increases to middle school, the focus will be to develop students' ability to do certain inquiry activities but not cover all the inquiry elements, just like the case of Ms. LIU.

Moreover, this study shed light on the research method for the enactment of scientific inquiry. The coding schemes for the enacted inquiry include important aspects of each inquiry element, such as types of the activities, approaches to the activities, subject who initiates the activity and teaching practices. These aspects provided multifaceted perspectives for each inquiry activity, which may contribute to describing teacher's interpretation and enactment in classroom.

Factors Influencing Enactment

This study evidences that teacher's understanding about scientific inquiry, textbooks, assessment materials, students and resources are the major factors that influence enactment of scientific inquiry. Findings from this study have implications for curriculum developers and teacher educators in science education. Curriculum developers should make the inquiry embedded in curriculum materials in a more explicit way, giving an emphasis on the inquiry activity elements. To promote instruction improvement, teacher educators need to pay more attention to teacher's understanding of scientific inquiry reflected by their instructional practices. Furthermore, assessment materials that contain items to assess students' ability to do inquiry will be helpful to promoting instruction.

Appendix: Interview Protocols and Questionnaire

(a) Pre-topic Interview Protocol

1. Tell me something about the topic.
2. How do you decide what to teach?
3. Which resources, such as documents or Web sites, did you refer to in planning this topic?
4. What are your main objectives in teaching this topic? What do you hope that your students will learn about this topic?
5. How do you decide how to teach the topic? What activities have you chosen for the teaching of this topic?
6. What do you think are the challenges for the students to learn about this topic?

(b) Post-topic Interview Protocol

1. To what extent have you achieved your instructional goals for this unit?
2. How did you know how well you have achieved your instructional goals for this unit?
3. To what extent did assessments of any type influence the planning and the teaching of this unit? And how does it affect any other aspects of teaching beside this topic?
4. What were the key decisions you made during the teaching of this unit?

(c) The Teacher Questionnaire for Each Lesson

1. Please describe the subject content of today's lesson.
2. What was the main thing you wanted students to learn from today's lesson? Why do you think it is important for students to learn this?
3. To what extent did the students meet your learning goals? How do you know?
4. Please describe what did not go according to your plan.

References

Arizona State University (2008) Discourse in Inquiry Science Classrooms (DiISC): Reference Manual. Retrieved from http://digitalcommons.unl.edu/teachlearnfacpub/121/

Banchi H, Bell R (2008) The many levels of inquiry. The Learning Centre of the NSTA. Retrieved from http://learningcenter.nsta.org/files/sc0810_26.pdf

Breslyn W, McGinnis JR (2012) A comparison of exemplary biology, chemistry, earth science, and physics teachers' conceptions and enactment of inquiry. Sci Educ 96:48–77

Chinese Ministry of Education (2001a) National curriculum standards series: science curriculum standards for grade 3–6. Beijing Normal University Press, Beijing (in Chinese)

Chinese Ministry of Education (2001b) National chemistry curriculum standard for junior high school. Beijing Normal University Press, Beijing (in Chinese)

Crawford BA (2007) Learning to teach science as inquiry in the rough and tumble of practice. J Res Sci Teach 44(4):613–642

Eltinge EM, Roberts CW (1993) Linguistic content analysis: a method to measure science as inquiry in textbooks. J Res Sci Teach 30(1):65–83

Forbes CT, Davis EA (2010) Curriculum design for inquiry: preservice elementary teachers' mobilization and adaptation of science curriculum materials. J Res Sci Teach 47(7):820–839

Ford MJ, Wargo BM (2007) Routines, roles, and responsibilities for aligning scientific and classroom practices. Sci Educ 91:133–157

Fu L, Zhou S, Zheng C (2007) Investigating the frequencies and time length of inquiry activities. Chem Educ 11:42–44 (in Chinese)

Gerrnann PJ, Haskins S, Auk S (1996) Analysis of nine high school biology laboratory manuals: promoting scientific inquiry. J Res Sci Teach 33(5):475–499

Hasson E, Yarden A (2012) Separating the research question from the laboratory techniques: advancing high-school biology teachers' ability to ask research questions. J Res Sci Teach 49(10):1296–1320

Liu X, Liang L, Liu E (2012) Science education research in china: challenges and promises. Int J Sci Educ 34(13):1961–1970

Mcdonald S, Songer NB (2008) Enacting classroom inquiry: theorizing teachers' conceptions of science teaching. Sci Educ 92:973–993

National Research Council (1996) National science education standards. National Academy Press, Washington, DC

Schneider RM, Krajcik J, Blumenfeld P (2005) Enacting reform-based science materials: the range of teacher enactments in reform classrooms. J Res Sci Teach 42(3):283–312

Schwab JJ (1962) The teaching of science as enquiry. In: Schwab JJ, Brandwein PF (eds) The teaching of science. Harvard University Press, Cambridge

Wang L (2010a) Progress and reflection on chemistry curriculum reform (part 1). Chem Educ 4:15–21 (in Chinese)

Wang L (2010b) Progress and reflection on chemistry curriculum reform (part 2). Chem Educ 5:20–24 (in Chinese)

Wu H, Krajcik JS (2006) Inscriptional practices in two inquiry-based classrooms: a case study of seventh graders' use of data tables and graphs. J Res Sci Teach 43(1):63–95

Zhou S, Li T, Zheng C (2005) Interpretation on the teaching activities of scientific inquiry in the context of new curriculum. Teach Educ Res 07:65–69 (in Chinese)

Journal of Science Education and Technology

Journal Website

www.springer.com/10956
Electronic edition: link.springer.com/journal/10956

Subscription Information

Journal of Science Education and Technology is published 6 times a year. Volume 23 (6 issues) will be published in 2014.

ISSN: 1059-0145 print
ISSN: 1573-1839 electronic

Orders, Inquiries and Subscription Rates

For information on subscription rates please contact Springer Customer Service Center:

The Americas (North, South, Central America and the Caribbean)
Springer Journal Fulfillment
233 Spring Street, New York,
NY 10013-1578, USA
Tel.: 800-SPRINGER (777-4643);
212-460-1500 (outside North America)
journals-ny@springer.com;
servicio-ny@springer.com
(Central and South America)

Outside the Americas
Springer Customer Service Center GmbH
Haberstr. 7, 69126 Heidelberg, Germany
Tel.: +49-6221-345-4303
subscriptions@springer.com

Office of Publication

Springer Science+Business Media LLC, 233 Spring Street, New York, NY 10013. Periodicals postage paid at New York, NY and additional mailing offices.

Postmaster: Send address changes to *Journal of Science Education and Technology*, Springer, 233 Spring Street, New York, NY 10013-1578, USA

Journal of Science Education and Technology

Volume 23 · Number 2 · April 2014

Special Issue: International Conference on Science Education (Nanjing, China, October 2012)

Guest Editors: Xiufeng Liu · BaoHui Zhang

Further articles can be found at link.springer.com

Indexed/abstracted in *Science Citation Index Expanded (SciSearch), Journal Citation Reports/Science Edition, Social Science Citation Index, Journal Citation Reports/Social Sciences Edition, Social SciSearch, SCOPUS, PsycINFO, Astrophysics Data System (ADS), Chemical Abstracts Service (CAS), Google Scholar, EBSCO, Academic OneFile, Academic Search, CSA Environmental Sciences, Current Contents/Social & Behavioral Sciences, Educational Research Abstracts Online (ERA), Educational Technology Abstracts, ERIC System Database, ERIH, Gale, JSTOR, MathEDUC, Multicultural Education Abstracts, OCLC, SCImago, Sociology of Education Abstracts, Studies on Women & Gender Abstracts, Summon by Serial Solutions, Vocational Education and Training Abstracts*

Instructions for Authors for *J Sci Educ Technol* are available at www.springer.com/10956