Scientific Research and Essay Vol. 4 (6) pp. 626-631, June, 2009
Available online at http://www.academicjournals.org/SRE
ISSN 1992-2248 © 2009 Academic Journals

Full Length Research Paper

Development of two-tier diagnostic instrument and assess students' understanding in chemistry

Cengiz Tüysüz

Department of Primary Education, Mustafa Kemal University, Hatay, Turkey. E-mail: ctuysuz@gmail.com. Tel: 90 326-2213077-112. Fax: 90 326 2455811.

Accepted 19 May, 2009

The objective of this study was to develop a two-tier test about the subject of "Separation of Matter" in 9th grade chemistry curriculum and to argue the efficiency of this test on students' achievement. First, we interviewed 21 students to develop multiple-choice items. Secondly, 24 students were asked to select the most appropriate answer to each question and then they gave explanations for their choices to develop a two-tier test. At the third phase of this study, a two-tier test including 15 questions was developed and applied to141 students for collecting data. The results of the study showed that the two-tier test was effective on determining the students' misconceptions and also it might be used as an alternative to the traditional multiple choice test for of assessment and evaluation of alternative students' achievement.

Key words: Two-tier test, chemistry, misconception, traditional multiple choice test, achievement.

INTRODUCTION

Many researchers agreed that students bring their pre-conceptions to class in science education (Ausubel, 2000; Driver and Oldham, 1986). Children develop ideas and beliefs about the natural world through their everyday life experiences. These include informal instruction like, sensual experiences, language experiences, cultural background, peer groups, as well as formal instruction. Studies have revealed that during science class students bring about certain ideas and explanations to natural phenomena that are inconsistent with the ideas accepted by the scientific community (Osborne et al., 1983). These existing ideas are often strongly held, resist to traditional teaching and form coherent though mistaken conceptual structures (Driver and Easley, 1978). Students may undergo instruction in a particular science topic, do reasonably well in a test on the topic, and yet, do not change their original ideas pertaining to the topic even if these ideas are in conflict with the scientific concepts they were taught (Fetherstonhaugh and Treagust, 1992). Duit and Treagust (1995) attributed this to students being satisfied with their own conceptions and therefore seeing little value in the new concepts. Osborne, Bell and Gilbert (1983) stated that students often misinterpret, modify or reject scientific viewpoints on the basis of the way they really think about how and why things behave, so it is not surprising that the research showed that students may persist almost totally with their existing views (Treagust et al., 1996).

Methods used to determine students' understanding of concepts included concept mapping (Novak, 1996), interviews (Carr, 1996) and multiple-choice diagnostic instruments (Treagust, 1995).

The researchers of science education use many methods to determine the alternative concepts of the students. Students' alternative conceptions have to be identified so that measures can be taken to help students improve themselves more scientifically acceptable concepts (Taber, 1998). Studies in which students' alternative conceptions were described cover a wide range of subject areas including science (Garnett et al., 1995; Barker and Millar, 2000; Pedrosa and Dias, 2000; Schmidt, 2000; Taber et al., 2000; Taber, 2001; De Jong and Treagust, 2002; Harrison and Treagust, 2002).

While primary multiple choice tests are mostly used, concept maps, guesses, observation and explanations, researches, concept phase diagrams, V diagrams, question forming, two-tier test and the other new information researching are also used in determining alternative concepts.

To date, several diagnostic tests have been developed

and were described in the literature for determining the alternative concepts (Tan and Treagust, 1999; Tan et al., 2002; Karatas et al., 2003; Chou and Chiu, 2004; Wang, 2004; Treagust, 2006; Chandrasegaran et al., 2007; Taber and Tan, 2007; Tan et al., 2008).

The rules of development of two-tier multiple-choice diagnostic instrument used in this study which described by Treagust (1988). In this instrument, the first tier of each item consist of a content question of five choices; the second part of each item contain a set of five justifications for the answers to the first part. Included in these justifications are the correct answer and two to five distracters. Distracters are derived from students' alternative explanations gathered from the literature, interviews and free responses. In advocating different assessment procedures to probe students' understanding of scientific concepts, Simpson and Arnold (1982) recommended that information relating to erroneous information held to be true by students should be included in the tests that have distracters. This line of research in assessment has included the development of multiple choice tests items that have distracters based on students' conceptions.

Two-tiered questions have two main benefits over conventional one-tiered questions. The first is a decrease in the measurement error. In a one-tiered multiple choice question with 5 possible choices, there is a 20% chance of correctly guessing the answer. These random, correct guesses must be accounted for in the measurement error. A two-tiered question is considered correct only if both tiers are answered correctly. As a result, a student responding to a question with 5 choices in the first tier and 5 in the second has only 4% chance of randomly correct guessing.

The second benefit to the two-tiered format is that it allows for the probing of two aspects of the same phenolmenon. In the first tier, students are asked to predict the outcome of a chemical change, and the second tier asks for an explanation. This allows the probing of the phenomenological domain with the first tier and the conceptual domain with the second.

In this study, it is aimed to develop a two-tier test about "Separation of Matter" in 9th class chemistry curriculum to argue the efficiency of this test on student' success. This unit is preferred because it is found difficult and not understood by students. Also, after the students have graduated from Anatolian Hotel and Tourism Occupational High School, they work in Hotels and need this information relate the separation of matter in their occupation. Başer (2006) reported that students respond questions correctly in science concepts but they hold misconceptions while applying to everyday situations.

Students' explanations are important for teaching of science concepts. They find opportunity to select an answer and its explanation and also teachers learn about reasons of the students' misconceptions in two tier tests.

For this reason, the most common misconceptions of the students about "Separation of Matter" are determined by using a two-tier test. Two-tier test and multiple choice tests were compared with respect to assessment and evaluation and furthermore, it is studied whether two-tier tests can be used as an alternative to multiple choice tests.

MATERIALS AND METHODS

This study is incorporating both qualitative and quantitative methods (Merriam, 1998). The two-tier multiple-choice diagnostic instrument was developed in three phases using procedures defined by Treagust (1995). The students were required to make drawings, to use Styrofoam, and to think aloud about their images of particles during the interviews. First, we interviewed the students about particles using open-ended questions to obtain original responses. Second, we developed multiple-choice items based on the data drawn from the students. Third, we conducted a paper-and-pencil test using the drawings from the students.

The first phase: interviewing

To understand 9th graders' preconceptions about the separation of matter, we interviewed 21 students individually using open-ended questions. The interview was conducted to collect data for the choices of the multiple-choice items.

The second phase: paper-and-pencil test

Responses to the interview questions were used to form multiple-choice items for a paper-and-pencil test. These items were considered to be student-oriented items that were different from traditional multiple choice test items. Once the items were constructed, an entire class of 24 students was tested. The students were asked to select the most appropriate answer to each question and then give explanations for their choices. The data obtained from the students were analyzed and developed into two-tier test items: the first tier for their image representations and the second tier for their explanations. The second tier was prepared by using student' misconceptions which were validated and determined by two chemistry professors by using data collected from 24 students' explanations. Students' responses related explanations for their choices on question 3 and 10 and researcher's inferences about them were given below.

Question 3
S-7: Sweet-water is a mixture between solid and liquid. For this reason, it dissociated with filter. While Sold is collected on filter paper, water is collected in container.
S-19: Sweet-water is homogeneous mixture. Distillation and crystallization are used separation of homogeneous mixtures. For these reasons answers II and III are correct.
S-11 Distillation method is used for separation homogeneous mixtures of liquid-liquid. But sweet-water is a homogeneous mixture of solid- liquid. For this reason answers III is correct.

Students were asked to determine whether they know or not differences between distillation and crystallization methods in this question and answer of first tier is B and second tier is D.

Interviews results showed that students have a misconceptions related to determine homogeneous and heterogeneous mixtures. This misconception is given A section of second tier in the question 3. Similarly, students believed that liquid evaporated and the solid is only obtained in the crystallization method. For this reason, students hold misconceptions in these concepts because they do

Table 1. Result of the TTT and TMT analysis.

Test Type	N	$\overline{X}$	SD	t	P
TTT	141	5.48	1.14	-16.169	0.000*
TMT	141	9.40	2.02		

not know difference between the crystallization and distillation. These explanations are given in the B and C of the second tier in the question 3.

Question: 10
S-2 When alloys are heated; low melting point solid is first melted and accumulated in the bottom of the container. When F is heated, it dissociated C and T substances. For this reason, it is an alloy.
S-11 Substance F is given on the question that is composed of same particles. Elements are composed of same particles. For this reason F is an element.

S 14 Solid-liquid mixtures separated with heating the crystallization method. For example, when sweet- water is heated, water evaporated and square precipitated in the bottom of the container. First appearance of sweet water is homogeneous but when it is heated, dissociated water and square.

Students were asked to determine whether they know or not that compounds is divided into components by heating in this question and answer of first tier is E and second tier is C.

Interviews results showed that students have misconception whether states of matter or heat is used to differentiate dissociation of matters. When substance of F is heated, it is divided into components with chemically. Students mixed this situation with alloys are divided into components with heating. This misconception is given in A choice of second tier in this question.

Also students have difficulties which differentiate same particles, type of same atom and molecules. A sentence is given question 10 like ''a substance of F is consisted of same particles ''. Students mixes same type of atom with same type of particles and then they think that one atom is same particle and their answer is element in this question. This explanation is given in section B. Also students showed misconception when mixture is divided into components with heating. This explanation is given in section D.

The third phase: two-tiered test (TTT)

The two-tier test was developed to determine the misconceptions of the 9th grade students about the unit of "Separation of Matter" and to determine whether two-tier tests can be used as an alternative to multiple choice tests in the education area. With this aim, at the 3rd phase of this study, a draft of two-tier test including 25 questions was developed. The draft form of the test was applied to 156 students, and the factor loading were calculated. As a result, 15 questions were selected as TTT (Figure 1). After selecting the 15 questions, the TTT was applied to 141 students to collect the data in Anatolian Hotel and Tourism Occupational High School. Students have taken two hours chemistry in a week and application was done after students took unit of separation of matter.

The Cronbach α- reliability coefficient was found 0.86 for the two tier test. The minimum point was 0 and the maximum point was 15 for TTT. The points increase so did the achievement.

Participants

This study was performed with 141 (76 male, 65 female) 9th grade high school students from Anatolian Hotel and Tourism Ocupational Occupational High School, in the first term of the 2006-2007-academic year. After data was collected from this test, first tier was evaluated as a Multiple Choice Test (MCT) and both of tier was evaluated as Two-Tier Test (TTT).

Data analysis

In this study, both qualitative and quantitative data are analyzed. Analysis of the results of the quantitative data is done by using SPSS/PC 11.0 program. While analyzing the quantitative data, arithmetic means, standard deviations and standard errors are calculated. During the evaluation of MCT, students were given one point if only first tier was true; otherwise, zero point was given. During the evaluation of TTT, students were given one point if only both of tiers were true; otherwise, zero point was given for all of other choices. Then, two-tier test and traditional multiple choice test compared.

RESULTS

Table 1 shows the average scores of the students on items related to separation of matter concepts. After making pair wise comparisons (Least Significant Difference), we found that there were significant differences between the two tier test and the traditional multiple choice test (TMT) ($p = 0.000$).

Table 1 shows that there is a significant difference between the two tier test (TTT) and the traditional multiple choice test (TMT) ($t_{(141)} = -16,169$, $p < 0,05$). When we look at the arithmetic mean, it is seen that the average of the traditional multiple choice test ($X = 9.40$) is higher than the average of the two-tier test ($X = 5.48$). This result shows that students are more successful at the traditional multiple choice test. In other words, traditional multiple choice test shows the students' achievement more than they are. This may be because of many reasons;

- In a traditional multiple choice question with 5 possible choices, the chance of guessing the correct answer is 20 percent. But, in a two-tiered, it is 4 percent. By lessening the chance of the achievement from 20 percent to 4 percent, the arithmetic means of the students might decrease.
- In our country, since the most important exams are done by using traditional multiple choice tests, students learn the problem solving techniques instead of learning the subject. So, the students can find the correct answers without knowing why this one is true.
- The students' learning levels take place at a lower step of the cognitive domain. When they come to a face to face two-tier test, which require higher cognitive skills, the students become confused.
- The students only memorize the knowledge.
- Misconceptions of the students may lead them to make errors. Because the explanations of two-tier test are chosen by considering the misconceptions of the students. Below are some common alternative conceptions that

3. To get sugar and water separately from the sugar-water mixture;
I. Filtering II. Distillation III. Crystallizing

Which one or ones can be used?
Only I
Only II
Only III
II and III
I, II and III

EXPLANATION

Solid-liquid heterogeneous mixtures can be differentiated by filtering method.
Solid-liquid homogeneous mixtures can be separated by distillation and crystallizing method.
By distillation method, liquid-liquid homogeneous mixtures can be separated.
For separation of the homogeneous mixture; crystallizing method is used to get solid and distillation method is used to get solid and liquid separately.
There is more than one truth in the explanations. (B-D)

10. F is made from the same particles and when F is heated it is separated into C and T.
According to this knowledge what is F?
Alloy
Element
Solution
Heterogeneous mixture
Compound

EXPLANATION

When alloys heated, the matter whose melting point is lower separated firstly.
Elements are made of the same particles.
Heating process is a chemical method and compounds are separated by chemical methods.
When solutions are heated, liquid is evaporated and solid is separated through precipitation.
Heterogeneous mixtures can be separated by filtering or sifting. (E-C)

Figure 1. A sample two tier question used in this study.

students have, which are taken from the explanation part of the two-tier test.

- 34% of the students can not become aware of the difference between heterogeneous and homogeneous mixtures. For example; they suppose that water-salt mixture is a suspension because of being solid-liquid mixture and therefore, they think that the filtration method is using for separating this mixture.
- 52% of the students cannot understand the differences between element, compound and mixture. Therefore, they can not understand the differentiating methods to separate them. They cannot distinguish which methods to use, is a chemical or physical.
- 69% of the students had problems about distinguishing properties of matter. For example; when it is asked which property has to be used while separating the mixture of water-ethyl alcohol? Students gave answers such as; density, solubility etc.
- 71% of the students had problems in separating the mixtures which contained three or more components.

DISCUSSION AND CONCLUSION

In the first step we developed two-tier test. In the second step the effects of the two-tier test were discussed. The results of the study showed that the two-tier test was efficient to determine the alternative conceptions of the students. It was found that students could not become aware of the difference between homogeneous and heterogeneous mixtures, could not understand the differences between element, compound and mixture, and could not choose the methods for separating matters. On the other hand, the students had difficulties in differentiated mixtures which contained three or more components. The results of the study were supported by other studies in literature (Chiu, Chiu and Ho, 2002; Wang, 2004; Chiu, 2005; Kwen and Cheng, 2005; Tan et al 2005).

Results showed that to determine the achievement of the students, a two-tier test might be used as an alternative to traditional multiple choice tests. Literature review supported the results (Chen et al., 2002). Tsai and

Chou (2002) pointed that two-tier test helped teachers teach and students learn better. Chen and Lin (2003) found that the two-tier multiple choice test provided a reliable and valid pencil-and-paper, easy-to-score instruments for science teachers and/or researcher to better evaluate students' idea.

When we compare a two-tier test with a multiple choice test, the two-tier test was more effective to determine the students' alternative conceptions on a subject and to reveal whether meaningful learning occurs or not (Treagust, 1995). Being as easy as the evaluation of conventional multiple choice test, and the same time, student knows why he is giving particular answer to the question could make the two-tier test much more effective than the other tests (Peterson and Treagust, 1989).

Numerous studies have shown that learners at different levels and ages have difficulties understanding science concepts. Stavy (1990) showed that even when 8th and 9th graders were taught about the concepts of particles in their science curricula, only about 15% of 8th and 9th graders who had studied under the same curricula were able to use the composition and arrangement of particles to explain phase changes.

Such difficulties were revealed by students' alternative conceptions which were generated during instruction. The reasons for such difficulties were some common features such as students' complex, abstract, and inconsistent intuition about same science concepts.

The results of this study showed that the two-tier test could help teachers teach and students learn better. Therefore, instead of multiple-choice tests, two-tier tests can be easily used by teachers to increase students' knowledge level and prevent students' alternative conceptions. Thus, two-tier tests help to improve teaching and learning.

REFERENCES

Ausubel DP (2000). The acquisition and retention of knowledge: a cognitive view, Kluwer Academic Publishers, Dordrecht.

Barker V, Millar R (2000). Students' reasoning about basic chemical thermodynamics and chemical bonding: what changes occur during a context-based post-16 chemistry course? Int. J. Sci. Edu. 21: 645-665.

Başer M (2006). Fostering Conceptual Change By Cognitive Conflict Based Instruction On Students' Understanding Of Heat And Temperature Concepts. Eurasia J. Math. Sci. Technol. Edu. 2: 96-114.

Carr M (1996). Interviews about instances and interviews about events. New York: Teachers College Press.

Chandrasegaran AL, Treagust DF, Mocerino M (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. Chem. Educ. Res Practice, 8 (3): 293-307.

Chen CC, Lin HS, Lin ML (2002). Developing a two-tier diagnostic instrument to assess high school students' understanding, The formation of image by plane mirror. Prac. Natl. Sci. Counc. ROC (D), 12(3): 106-121

Chen CC, Lin ML (2003). Developing a two-tier diagnostic instrument to assess high school students' understanding, The 4th International Conference of the European Science Education Research Association (ESERA), Netherlans.

Chiu MH (2005). A National Survey of Students' Conceptions in Chemistry in Taiwan. Chem. Edu. Int. 6 (1).

Chiu ML, Chiu MH, Ho CY (2002). Using Cognitive-Based Dynamic Representations to Diagnose Students' Conceptions of the Characteristics of Matter. Proc. Natl. Sci. Counc. ROC (D) 12 (3): 91-99.

Chou CC, Chiu MH (2004). A two-tier diagnostic instrument on the molecular representations of chemistry: comparison of performance between junior high school and senior high school students in Taiwan. Paper presented at the 18th International Conference on Chemical Education, Istanbul, Turkey

De Jong O, Treagust DF (2002). The teaching and learning of electrochemistry, in J.K. Gilbert, O. De Jong RJ Treagust DF, Van Driel KH (Eds.), Chemical education: towards research-based practice, Kluwer Academic Publishers, Dordrecht, pp. 317-337.

Driver R, Easley J (1978). Pupils and paradigms: a review of literature related to concept development in adolescent science students. Stud. Sci. Educ. 5: 61-84.

Driver R, Oldham V (1986). A constructivist approach to curriculum development in science. Stud. Sci. Educ. 13: 105-122.

Duit R, Treagust D F (1995). Students' conceptions and constructivist teaching approaches. Chicago: The National Society for the Study of Education.

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

Garnett PJ, Garnett PJ, Hackling MW (1995). Students' alternative conceptions in chemistry: a review of research and implications for teaching and learning. Stud. Sci. Educ. 25: 69-95.

Harrison AG, Treagust DF (2002). The particulate nature of matter: challenges in understanding the submicroscopic world, in J.K. Gilbert, O. De Jong, R. Justi., D.F. Treagust and K.H. Van Driel (Eds.), Chemical education: towards research-based practice, Kluwer Academic Publishers, Dordrecht, pp. 189-212.

Karatas FO, Kose S, Costu B (2003). Öğrenci yanılgılarını ve anlama düzeylerini Belirlemede kullanılan iki aşamalı testler. Pamukkale Üniversitesi Eğitim Fakültesi Dergisi, 13 (1): 54-69.

Kwen BH, Cheng AK (2005). Using two-tier reflective multiple choice questions to cater to creative thinking. International Education Research Conference, Singapore, AARE

Merriam SB (1998). Qualitative research and case study applications in education, Jossey-Bass Publishers, San Francisco, CA.

Novak JD (1996). Concept mapping: a tool for improving science teaching and learning, in Treagust DF, R. Duit and. Fraser BJ (Eds.), Improving teaching and learning in science and mathematics, Teachers College Press, New York, pp. 32-43.

Osborne RJ, Bell BF, Gilbert JK (1983). Science teaching and children's view of the world. Eur. J. Sci. Edu. 5: 1-14.

Pedrosa MA, Dias MH (2000). Chemistry textbook approaches to chemical equilibrium and student alternative conceptions. Chem. Edu. Res. Pract. 1: 227-236.

Peterson RF, Treagust DF (1989). Grade-12 students' misconceptions of covalent bonding and structure. J. Chem. Edu. 66: 459-460.

Schmidt HJ (2000). Should chemistry lessons be more intellectually challenging? Chem. Edu. Res. Pract. 1: 17-26.

Stavy R (1990). Children's conception of changes in the state of matter: From liquid (or solid) to gas. J. Res. Sci. Teach. 27 (3): 247-266.

Taber KS (1998). The sharing-out of nuclear attraction: or "I can't think about physics in chemistry", Int. J. Sci. Edu. 20: 1001-1014.

Taber KS (2000). Multiple frameworks? Evidence of manifold concept- tions in individual cognitive structure. Int. J. Sci. Edu. 22: 399-417.

Taber KS (2001). Building the structural conception of chemistry: some considerations from educational research. Chem. Edu. Res. Pract. 2: 123-158.

Taber KS, Tan KCD (2007). Exploring Learners' Conceptual Resources: Singapore A Level Students' Explanations in the Topic of Ionisation Energy. Int. J Sci Math. Educ. 5 (3): 375-392.

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 qualitative analysis, J. Res. Sci. Teach. 39: 283-301.

Tan KCD, Taber KS, Liu X, Coll RK, Lorenzo
M, Li J, Goh NK, Chia LS (2008). Students' Conceptions of Ionisation Energy: A Cross-cultural Study, Int. J. Sci. Edu. 30 (2): 263-283.
Tan KCD, Treagust DF (1999). Evaluating students' understanding of chemical bonding. Sch. Sci. Rev. 81 (294): 75-83.
Tan KD, Taber K, Goh NK, Chia LS (2005). The ionization energy diagnostic instrument: a two-tier multiple-choice instrument to determine high school students' understanding of ionisation energy. Chem. Educ. Res. Pract. 6 (4): 180-197.
Treagust DF (1988). Development and use of diagnostic tests to evaluate students misconceptions in science. Int. J. Sci. Edu 10 (2): 159-169.
Treagust DF (1995). Diagnostic assessment of students' science knowledge. In: Glynn SM, Duit R (Eds.), Learning science in the schools: Research reforming practice. Mahwah, New Jersey: Lawrence Erlbaum Associates. pp. 327-346.
Treagust DF (2006). Diagnostic assessment in science as a means to improving teaching, learning and retention. UniServe Science – Symposium Proceedings: Assessment in science teaching and learning, Uniserve Science, Sydney, Australia, p. 1-9.
Treagust DF, Duit R, Fraser BJ (1996). Overview: research on students' preinstructional conceptions – the driving force for improving teaching and learning in science and mathematics, in D.F. Treagust, R. Duit and B.J. Fraser (Eds.), Improving Teaching and Learning in Science and Mathematics, Teachers College Press, New York, pp. 1-14.
Tsai CC, Chou C (2002). "Diagnosing students' alternative conceptions in science through a networked two-tier test system". Int. J. Comp. Assisted Learn. 18 (2): 157-165.
Wang JR (2004). Development and Validation of a Two-Tier Instrument to Examine Understanding of Internal Transport in Plants and the Human Circulatory System. Int. J. Sci. Math. Edu. 2: 131-157.