

INTE 2014

Development of a two-tier diagnostic test to assess undergraduates' understanding of some chemistry concepts

Ayfer Mutlu^{a*}, Burcin Acar Sesen^b

^aKırklareli University, Vocational School of Health Service, Kırklareli, 39100, Turkey

^bIstanbul University, Hasan Ali Yücel Education Faculty, Science Education Department, 34070, Istanbul, Turkey

Abstract

The purpose of this study was to develop a valid and reliable two-tier diagnostic test to assess students' undergraduates' understanding of some concepts in the context of undergraduate general chemistry course. For developing this test, firstly, content boundaries, learning objectives and alternative conceptions were identified. The first version of the test with an open ended part was applied on 68 pre-service science teachers, and qualitative analyses were made on open-ended part. According to results, both tiers were made in multiple-choice format. The test was validated by five chemistry educators and test was applied with 151 pre-service science teachers. The final version of the two-tier diagnostic test was consisted of 44 items, reliability coefficient was found to be 0.84.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the Sakarya University

Keywords: Acids and bases, chemical equilibrium, chemical kinetic, electrochemistry, thermochemistry, two tier diagnostic test

1. Introduction

Many students hold on to their own personal views and inadequate interpretations of particulate phenomena that develop from their individual experiences, culture and classroom instruction (Duit & Treagust, 1995). These interpretations called misconceptions influence how students learn new scientific knowledge play an essential role in subsequent learning (Özmen, 2004). Therefore alternative conceptions at all school levels constitute a major problem of concern to science educators and some methods have been developed to identify them.

Two-tier diagnostic test have been regarded as an effective assessment tool to determine students' conceptual understanding and alternative conceptions (Treagust, 1988; Odom & Barrow, 1995). In this test, there are two tiers and the first tier of each item consists of a multiple-choice question includes content knowledge. The second tier of

* Corresponding author. Ayfer Mutlu Tel. +90-0288-214-38-25 fax: +0-000-000-0000
E-mail address: ayferkaradas@gmail.com.

each item contains reason for students' choosing to the first tier. The second tier included two, three, four or five responses, one of which was the expected answer. Distractors derived alternative conceptions from literature, interviews, and students justifications. This alternative assessment method has been used widely to assess students' understanding (e.g. Chandrasegaran, Treagust & Mocerino, 2007; Chu, Treagust & Chandrasegarana, 2009; Tan, Goh, Chia & Treagust, 2002; Tsui & Treagust, 2010)

1.1. Purpose of the Research

In this study, it was aimed to develop a two-tier diagnostic test to assess undergraduates' understanding of the subjects of *thermochemistry, chemical kinetics, chemical equilibrium, acids and bases and electrochemistry* in the context of undergraduate general chemistry course.

1.2. Participants

This study was conducted with 219 pre-service science teachers (18 to 21 years of aged) in an education faculty in Turkey. 68 of them were answered multiple-choice test with an open ended part. Two-tier diagnostic test was applied to 151 of them. The pre-service science teachers' socio- economic statuses were similar and all the participants achieved General Chemistry II course.

1.3. Procedures

In this study, two-tier diagnostic test was developed based on following Treagust's methods (1988) by the researchers. Development process was conducted on three phases and ten steps. The flowchart of instrument development process was presented in Figure 1.

Phase 1: Define the content area

In this phase, firstly, content boundaries and learning objectives were determined according to research purpose. Following learning objectives were documented related to research area.

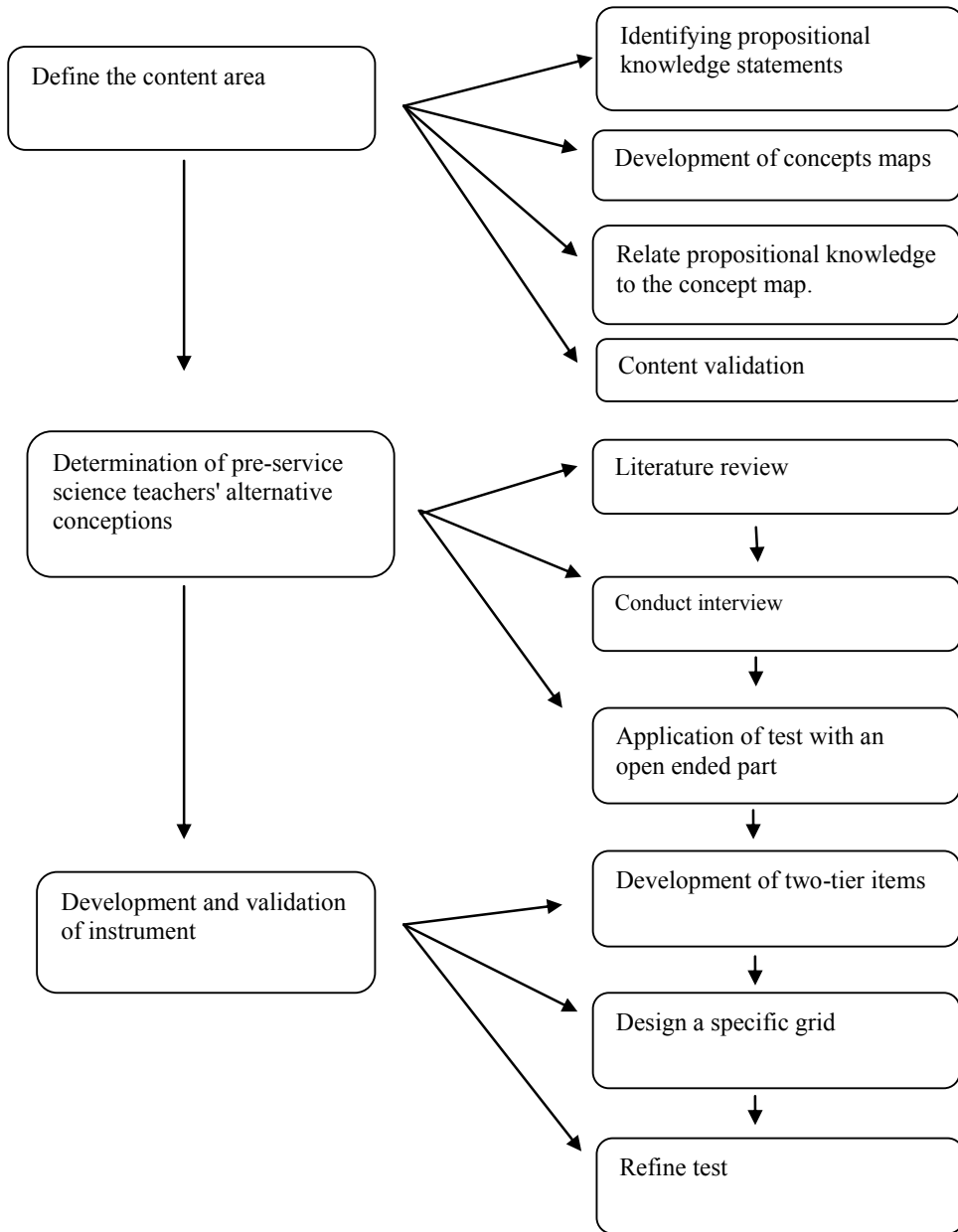
Thermochemistry;

- Students are able to explain reaction enthalpy.
- Students are able to calculate reaction enthalpy.
- Students are able to explain conversation of energy.
- Students are able to explain endothermic reactions.
- Students are able to compare different reactions' enthalpy by using calorimeter.
- Students are able to explain bond energy.

Chemical Kinetic

- Students are able to explain reaction rate.
- Students are able to compare value of reaction constant on multiple-step reactions.
- Students are able to interpret effective collisions.
- Students are able to interpret effect of concentration on reaction rate.
- Students are able to compare effect of temperature on endothermic and exothermic reactions' rate.
- Students are able to explain effect of temperature on reaction constant.
- Students are able to explain effect of surface on reaction rate.
- Students are able to interpret effect of catalyst on reaction rate.

Figure 1: Scheme of the development process of two-tier diagnostic test based on Treagust (1988) in this study



Chemical Equilibrium

- Students are able to explain equilibrium dynamics on a reaction.
- Students are able to explain Le Chatelier Principle.
- Students are able to explain effect of concentration on equilibrium.
- Students are able to explain effect of temperature on equilibrium

Acids and Bases

- Students are able to explain strength of acids and bases.
- Students are able to explain pH.
- Students are able to interpret neutralization.
- Students are able to interpret acid-base equilibrium.
- Students are able to explain indicators and principles of their using.
- Students are able to explain titration.
- Students are able to choose best indicator for titration.
- Students are able to explain properties of end point.
- Students are able to explain properties of equivalence point
- Students are able to compare end point and equivalence point.
- Students are able to describe buffers.
- Students are able to interpret buffers mechanism.

Electrochemistry

- Students are able to determine anode and cathode according to several variables.
- Students are able to explain metal electrodes.
- Students are able to explain plating and galvanization.
- Students are able to explain effect of concentration on cell potential.

Then, concept maps for each subject were developed and relationship between propositional knowledge and concept maps was examined. All concept maps were validated by five chemistry educators. Development of the test depend the final version of the propositional knowledge statements and concept maps.

1.3.1. Phase 2: Determination of pre-service science teachers' alternative conceptions

In this phase, alternative conceptions were identified according to literature review (step 5) and pre-service science teachers' semi-constructive interviews (step 6). Later, 46 multiple-choice items with an open ended part, in where students were required to explain their reason for their answers to the first part, was constructed. Test was validated by five chemistry educators and it was applied 68 pre-service science teachers. The responses of the open-ended questions were analyzed. A list of pre-service science teachers' responses was constructed from their justifications that were provided to the 46 multiple-choice items, literature review and their interviews. Distractors included alternative conceptions were constructed according to this list. These alternative conceptions were subsequently used in the construction of the two-tier multiple-choice diagnostic instrument.

1.3.2. Phase 3: Development and validation of instrument

In the last phase, two-tier diagnostic test was developed in where the first tier is the conventional multiple choice step and the second tier is the possible reasons of the given answer for the first tier. The second tier included two, three, four or five responses, one of which was the expected answer. Distractors included high frequencies incorrect reasons and scientifically unacceptable conceptions held by pre-service science teachers, literature review and interviews. Distributions of items according to subjects and an example question were presented respectively in Table 1 and Table 2.

Table 1. Distributions of items according to subjects in the final version of the two tier diagnostic test

Subjects	Items
Thermochemistry	Q6, Q15, Q24, Q35, Q41, Q44
Chemical Kinetic	Q1, Q4, Q7, Q8, Q10, Q11, Q12, Q14, Q16, Q18, Q38
Chemical Equilibrium	Q27, Q30, Q32, Q34, Q36
Acids and Bases	Q2, Q5, Q9, Q13, Q17, Q19, Q20, Q21, Q23, Q25, Q28, Q29, Q31, Q37, Q42, Q43
Electrochemistry	Q3, Q22, Q26, Q33, Q39, Q40

For content validity, the test was examined by five chemistry educators and this test was applied on 151 pre-service science teachers. Then, item analysis and reliability analysis were made.

Table 2. An example question related to catalyst

$\text{H}_2\text{O}_2 + \Gamma \rightarrow \text{O}\Gamma + \text{H}_2\text{O}$ $\text{H}_2\text{O}_2 + \text{O}\Gamma \rightarrow \text{H}_2\text{O} + \Gamma + \text{O}_2(g)$ <hr style="border-top: 1px dashed black;"/> $2\text{H}_2\text{O}_2 \rightarrow 2 \text{H}_2\text{O} + \text{O}_2(g)$
<p>In the reaction above, which of the following is true about Γ ?</p> <p>a) It serves as an indicator. b) Products energy is increased by it. c) If it isn't in the reaction environment, reaction does not occur. d) It doesn't affect reaction equilibrium. e) Reaction enthalpy is decreased by it.</p> <p>BECAUSE;</p> <p>A) Reaction is gone to equivalence point quickly by it. B) Reaction yield is increased by it. C) Both forward and reverse reaction rate is increased same ratio by it. D) It provides energy to occur the reaction E) It causes to change potential energy of product and reactant.</p>

2. Results and Discussion

In this study, item analysis was conducted to identify items discrimination and difficulty. Results were presented in Table-3 for all tiers.

According to item analysis results, two items took off from the test because their discrimination indices were negative. Other items discrimination and difficulty indices were average level. The difficulty levels were 0.3 and 0.77 with an average of 0.50 based on the two-tier test scores item analysis. Moreover, discrimination levels were identified between 0.2 and 0.64 and average of them was calculated as 0.39. The final version of the two-tier diagnostic test was consisted of 44 items. Item discrimination indexes below 0.19 should be eliminated or completely revised (Crocker & Algina, 1986; Tan, 2006); and there were not any indexes below the criteria in our study as presented in the Table-3.

Moreover, the responses were analysed using a SPSS statistics software program. While analyzing these data, the answer was considered to be correct if both tiers were correctly answered (Tregust, 1988). Therefore one point was given for items only when both parts of the item were correctly answered, and zero point were given for items when either part was incorrectly answered. The reliability of the instrument was established by a Cronbach alpha coefficient of 0.84. If reliability coefficient is over 0.70-0.80, instrument is reliable (Özgülven, 1998) and our value was also consistent with other two-tier tests (e.g. Chandrasegaran, Tregust & Mocerino, 2007; Tsui & Tregust, 2010).

Table 3. Item statistics of two tier diagnostic test

Items	First Tier		Second Tier		Items	First Tier		Second Tier	
	Disc. Index	Diff. Index	Disc. Index	Diff. Index		Disc. Index	Diff. Index	Disc. Index	Diff. Index
1	0.63	0.28	0.62	0.29	24	0.30	0.62	0.33	0.59
2	0.60	0.24	0.60	0.27	25	0.62	0.31	0.38	0.52
3	0.72	0.27	0.62	0.27	26	0.46	0.47	0.30	0.59
4	0.60	0.23	0.60	0.29	27	0.74	0.25	0.32	0.59
5	0.59	0.35	0.62	0.26	28	0.72	0.26	0.36	0.57
6*	0.38	0.56	0.15	-0.10	29	0.37	0.43	0.36	0.48
7	0.73	0.26	0.48	0.25	30	0.46	0.32	0.46	0.20
8	0.60	0.33	0.40	0.43	31	0.41	0.30	0.34	0.56
9	0.72	0.40	0.72	0.39	32	0.59	0.32	0.57	0.34
10	0.46	0.25	0.30	0.22	33	0.63	0.33	0.50	0.32
11	0.60	0.37	0.51	0.32	34	0.67	0.35	0.59	0.42
12	0.37	0.52	0.30	0.57	35*	0.44	0.47	0.20	-0.13
13	0.46	0.45	0.32	0.42	36	0.59	0.39	0.31	0.59
14	0.50	0.44	0.34	0.59	37	0.36	0.45	0.41	0.37
15	0.40	0.47	0.44	0.48	38	0.42	0.21	0.32	0.56
16	0.82	0.20	0.53	0.26	39	0.32	0.56	0.31	0.57
17	0.56	0.29	0.47	0.50	40	0.75	0.20	0.66	0.23
18	0.73	0.27	0.63	0.21	41	0.47	0.22	0.63	0.22
19	0.32	0.54	0.34	0.54	42	0.77	0.29	0.66	0.40
20	0.30	0.27	0.32	0.57	43	0.62	0.41	0.58	0.47
21	0.30	0.48	0.32	0.64	44	0.45	0.42	0.32	0.54
22	0.63	0.35	0.27	0.64	45	0.55	0.34	0.41	0.46
23	0.61	0.26	0.51	0.20	46	0.30	0.64	0.56	0.21

In the light of the results it can be concluded that the developed two-tier test provide to be a reliable and valid instrument to assess undergraduates' alternative conceptions and identify their conceptual understanding of the subjects of *thermochemistry*, *chemical kinetics*, *chemical equilibrium*, *acids and bases* and *electrochemistry* in the context of undergraduate general chemistry course.

References

- Chandrasegaran, A. L., Treagust, D. F., & 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. *Chemistry Education Research and Practice*, 8(3), 293-307.
- Chu, H. E., Treagust, D. F., & Chandrasegaran, A. L. (2009). A stratified study of students' understanding of basic optics concepts in different contexts using two-tier multiple-choice items. *Research in Science & Technological Education*, 27(3), 253-265.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Florida: Holt, Rinehart and Winston INC.
- Duit, R., & Treagust, D. F. (1995). Students' conceptions and constructivist teaching approaches in Fraser B. J. and Walberg H. J. (ed.), *Improving science education* (pp. 233-248). Chicago, IL: The National Society for the Study of Education.
- Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32(1), 45-61.
- Özgülven, İ. E. (1998). *Psikolojik testler*. Ankara: PDREM Yayınları.

- Özmen, H. (2004). Some student misconceptions in chemistry: A literature review of chemical bonding. *Journal of Science Education and Technology*, 13(2), 147-159.
- Tan, Ş. (2006). *Öğretimi planlama ve değerlendirme*. Ankara: Pegem A Yayıncılık.
- Tan, K. C. D., Goh, N. K., Chia, L. S., & Treagust, D. F. (2002). Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of research in Science Teaching*, 39(4), 283-301.
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159-169.
- Tsui, C. Y., & Treagust, D. (2010). Evaluating Secondary Students' Scientific Reasoning in Genetics Using a Two-Tier Diagnostic Instrument. *International Journal of Science Education*, 32(8), 1073-1098.