TÜRK FEN EĞİTİMİ DERGİSİ
Yıl 14, Sayı 1, Mart 2017

Journal of
TURKISH SCIENCE EDUCATION
Volume 14, Issue 1, March 2017

http://www.tused.org

The Development of a Three-tier Chemical Bonding Concept Test

Şenol ŞEN[1], Ayhan YILMAZ [2]

[1] Dr. Hacettepe University, Faculty of Education, 06800, Beytepe-Ankara, Turkey
[2] Prof. Dr. Hacettepe University, Faculty of Education, 06800, Beytepe-Ankara, Turkey

Received: 26.08.2016 **Revised:** 12.12.2016 **Accepted:** 15.02.2017

The original language of article is English (v.14, n.1, March 2017, pp.110-126, doi: 10.12973/tused.10193a)

ABSTRACT

The aim of this study was to develop a three-tier concept test to determine high school students' conceptual understanding in terms of chemical bonding, and at implementing a reliability and validity study for the test. A total of 175 high school students participated in the research. A three-tier test of 15 questions was employed in the study. Expert opinion was consulted for content and face validity. Furthermore, false positive and false negative values were also calculated for content validity, and on average these values were found to be less than 10%. Statistical analyses were performed for construct validity; Cronbach's Alpha coefficient was calculated as 0.74. To perform item analysis, the top 27% of the students and the bottom 27% of the students were determined. As a result, it was found that item difficulty indices ranged between .47 and .77, and that the discrimination indices were above .30. Following the analyses, both item difficulty indices and discrimination indices were found to be adequate. At the end of the study, Chemical Bonding Concept Test was considered to be reliable and valid.

Keywords: Chemical bonding, chemistry education, reliability, three tier tests, validity.

INTRODUCTION

Students bring into the classroom setting their ideas, concepts and interpretations of the actions happening around them to make their learning easy or difficult; and these ideas, concepts and interpretations differ from scientific thoughts, concepts, definitions or ideas (Chandran, Treagust, & Tobin, 1987; Hammer, 1996; Reynolds & Walberg, 1992; Taber, 2009). Champagne, Gunstone, and Klopfer (1985) and West, Fensham, and Garrard (1985) attribute students' having difficulty in science course to their preconceptions about the happenings around them which do not overlap with scientific concepts (As cited in Yürük, 2005). Such preconceptions or ideas are called misconceptions or alternative concepts in the relevant literature (Driver & Easley, 1978; Nakhleh, 1992). Students' misconceptions are among the main factors negatively influencing their learning process (Muller, Sharma, Eklund, & Reimann, 2007). Therefore, identifying students' misconceptions is important in the learning/teaching process.

Interviews (Boo, 1998; Thompson & Logue, 2006), open-ended questions (Çalık & Ayas, 2005; Şekercioğlu & Kocakülah, 2008; Tsaparlis & Papaphotis, 2002), concept maps (Aykutlu & Şen, 2012; Hazel & Prosser, 1994; Şen & Yılmaz, 2013), analogies (Aykutlu &

Corresponding author e-mail: schenolschen@gmail.com
© ISSN:1304-6020

Şen, 2012), and multiple-choice tests (Şen & Yılmaz, 2012; Uzuntiryaki & Geban, 2005) are used in the literature to determine misconceptions. Although interviews are useful for determining students' misconceptions, the process of collecting the required data and transcribing and analysing them, is too time consuming (Lin, 2004). The fact that concept maps are seldom used often makes it necessary to teach both teachers and students how to use them. Besides, much more time is needed to use concept maps (Kaya, 2008). For this reason, various studies have made use of multiple-choice tests in determining students' misconceptions (Arslan, Çiğdemoğlu, & Moseley, 2012; Çetin-Dindar & Geban, 2011; Gürcay & Gülbaş, 2015; Kanlı, 2015; Peşman & Eryılmaz, 2010; Sreenivasulu & Subramaniam, 2013; Wuttiprom, Sharma, Johnston, Chitaree, & Soankwan, 2009). Due to the fact that multiple-choice tests are easy to administer, mark objectively and are cost-effective, makes them suitable for statistical analyses and therefore more usually preferred by researchers (Wuttiprom *et al*., 2009).

Determining misconceptions through conventional multiple-choice tests is not a very popular or preferred method. The reason for this is that it is impossible to determine whether students' answers stem from mistakes, lack of knowledge or from another factor apart from those in conventional multiple-choice tests (Voska & Heikkinen, 2000). Yet, it is possible to work with larger sampling groups by using multiple-choice questions. For this reason, it is recommended that multiple-choice questions are prepared in two or three tiers and used accordingly (Eryılmaz & Sürmeli, 2002). In parallel to this recommendation, for the last ten years, two-tier (Artdej, Ratanaroutai, Coll, & Thongpanchang, 2010; Kanlı, 2015; Şen & Yılmaz, 2012) and three-tier (Çetin-Dindar & Geban, 2011; Gürcay & Gülbaş, 2015; Peşman & Eryılmaz, 2010; Şen, 2015; 2016) tests have been used in the literature rather than one-tier multiple-choice tests (Hırça, Çalık, & Akdeniz, 2008). As in conventional tests, content is tested in the first tier of the two and three-tier tests, and the reason for their answers given in the first tier is requested in the second tier (reason tier) (Tan, Goh, Chia, & Treagust, 2002). However, two-tier tests may not be adequate in finding out whether or not students' misconceptions are due to lack of knowledge. Thus, students can be asked to state how confident they are about their answer by adding a third tier (confidence tier). This third tier is used to distinguish between the students' incorrect answers, due to lack of knowledge and their misconceptions. If a student gives incorrect answers in the first and second tiers, despite being sure about his/her answer, it is asserted that the student has a misconception. But if a student is not sure that his/her incorrect answer is correct, the incorrect answer cannot be regarded as a misconception (Peşman, 2005). Misconception tests developed by researchers are also marked by assigning 1 point to correct answers and zero points to incorrect answers - as in achievement tests (Eryılmaz & Sürmeli, 2002).

A review of the literature demonstrated that three-tier tests were developed and used in chemistry subjects, such as Acids and Bases (Çetin-Dindar & Geban, 2011), States of Matter (Kirbulut, Geban, & Beeth, 2010) and Electrochemistry (Şen, 2015; 2016), so as to determine students' misconceptions about chemistry. It was found that three different tests were developed in relation to chemical bonding (Othman, Treagust, & Chandrasegaran, 2008; Peterson & Treagust, 1989; Treagust, 1988); however, none of these was three-tier. Three-tier tests should be developed in order to be able to control external factors, such as lack of knowledge and chance in the process of determining students' misconceptions about chemical bonding; because research has shown that the subject of chemical bonding is difficult and problematic to students and, as a result, students have a large number of misconceptions (Coll, 2008; Coll & Taylor, 2002; Luxford, & Bretz, 2014; Nicoll, 2001; Peterson & Treagust, 1989; Şen & Yılmaz, 2013; Taber, 2011; Tan, & Treagust, 1999; Temel & Özcan, 2016; Yayon, Mamlok-Naaman, & Fortus, 2012). The fact that chemical bonding is an abstract concept and that it is not associated with learners' real life experiences makes it difficult to learn. As a consequence, that students have misconceptions about chemical bonding is an expected situation (Tan & Treagust, 1999). Using this as a starting point, the aim of this study is to develop a concept test on chemical bonding so as to determine the misconceptions of high school students in this respect, and to perform validity and reliability analyses. Studies available in the literature have shown that students have various misconceptions about the effects of physical changes on intramolecular bonds (Barker, 1995; Mirzalar Kabapınar & Adik, 2005), the formation and structure of ionic bonds (Luxford & Bretz, 2013; 2014), covalent bonds (Luxford & Bretz, 2013; 2014; Peterson & Treagust, 1989; Ünal, Coştu, & Ayas, 2010), metallic bonds (Acar & Tarhan, 2008), and intermolecular and intramolecular bonds (Othman *et al.*, 2008; Peterson & Treagust, 1989; Tan, & Treagust, 1999; Treagust, 1988). When considering these studies, performed in relation to misconceptions in the literature, the importance of a three-tier multiple-choice test to determine students' misconceptions about chemical bonding becomes manifest. For this reason, it is considered that this study will contribute to the literature concerning chemical bonding. The creation in this study of a valid and reliable test will facilitate its use by teachers. For this reason, the objective of this study is analysed using the following research problem:

- Are Chemical Bonding Concept Test results a valid and reliable measure of students' conceptual understanding of chemical bonding?

METHOD

a) Developing a Chemical Bonding Concept Test

The misconceptions concerning chemical bonding that are available in the literature were realized prior to preparing the chemical bonding concept test (Luxford, & Bretz, 2014; Nicoll, 2001; Peterson &Treagust, 1989; Şen & Yılmaz, 2013; Tan, & Treagust, 1999; Temel, & Özcan, 2016). The test prepared included three distractors and a correct answer. The misconceptions in the literature were used to create these distractors. The test contained three tiers and 15 questions. Expert opinion was obtained for the questions prepared for the test. For this purpose, the test was evaluated by two educators of chemistry education and by an expert chemistry teacher. Following the evaluations, the regulations were made, and the test was given its final shape. Subsequently, the test was administered to a group of 25 high school students, and efforts were made to determine whether or not there were any unclear or faulty questions. The questions were revised in accordance with the feedback from the students. Afterwards, the test was applied to 175 high school students for validity and reliability analyses to give it the final shape.

b) The Chemical Bonding Concept Test

Misconceptions included in the chemical bonding concept test (CBCT) are shown in Table 1.

Table 1. *The Distribution of Common Misconceptions included in the Items of CBCT*

Misconceptions	Items
Physical changes affect intramolecular bonds.	1, 2, 3, 11, 13
Compounds containing hydrogen bonds in the structure are in liquid phase.	2
Electrons are transferred from one atom to another during the formation of metal bonds, they can share electron.	4
An ionic bond is an intermolecular bond.	5
Polarity is determined by lone-pair electrons.	7
In all covalent bonds, bond-pair electrons are at equal distances from bonding atoms.	6
Whether a molecule is polar or not does not depend on the polarity of the bonds in that molecule and on the shape of the molecule.	8, 9
Intermolecular bonds are the forces available only in covalent compounds.	10
Energy is required for chemical bonding.	14
Heat and pressure change bond angles and the number of bonds.	11, 13
London forces and hydrogen bonds are available only between polar molecules.	10, 15
Non-bonding electron pairs do not have effects on geometric structure.	12
London forces are not real bonds.	10
London forces are both intermolecular and intramolecular.	10

A sample question included in the test is presented below. In the first tier of the question, the students were asked to explain the correlations between physical changes and heat. Option A is the correct answer, and the other three options are misconceptions. In the second tier, the students were asked to explain the reason for their answer in the first tier. Here, C is the correct option and the other three options are misconceptions. As is apparent from tier three, the students were asked to state whether or not they were sure about their answers.

1a) In which of the following is a change observed during the transition of pure ice from solid phase into liquid phase?

***A)** The strength of hydrogen bonds between molecules

B) Sigma (σ) bond gravity force in the molecule

C) Attractive forces of the covalent bond in the molecule

D) The number of sigma (σ) bonds in the molecule.

1b) Which of the following best explains your reason for your answer?

A) Hydrogen bonds between molecules are not chemical bonds but are only forces/interactions.

B) Sigma bonds are not real chemical bonds but are only weak forces.

***C)** Molecules that have more kinetic energy while ice is melting begin to separate, and hydrogen bonds are preserved to some extent.

D) The attractive forces of covalent bonds need to decrease in order for molecules having more kinetic energy to separate, while ice melting.

1c) About the answer to the above question:

a. I am sure.

b. I am not sure.

* Correct answer.

c) The Study Group

A total of 175 12th grade high school students were included in the study. The average age of the students was 18.02 (SD=.23). Of the participants, 84 were female and 78 were male. Three of them did not state their age, and thirteen did not state their gender. No mark was given to the papers of the students, who answered the first tier of the test but did not answer the second and third tiers, or those who chose only one option for all three tiers and these were excluded from the study. In addition to this, the data for those students who answered only half of the test, leaving the other half untouched, were also not analysed.

d) Data Collection Process

The students were allowed one class hour to answer the 15 questions in the test. They were encouraged to participate on the basis of volunteering, and were assured that the scores they received from the test would not affect their average academic achievement score.

e) Types of Scores according to Students' Answers

The types of scores for the 15 test questions - which were obtained through analyses using the Microsoft Excel program - are shown in Figure 1. Figure 1 was prepared on the basis of the study by Arslan, Çiğdemoğlu and Moseley (2012).

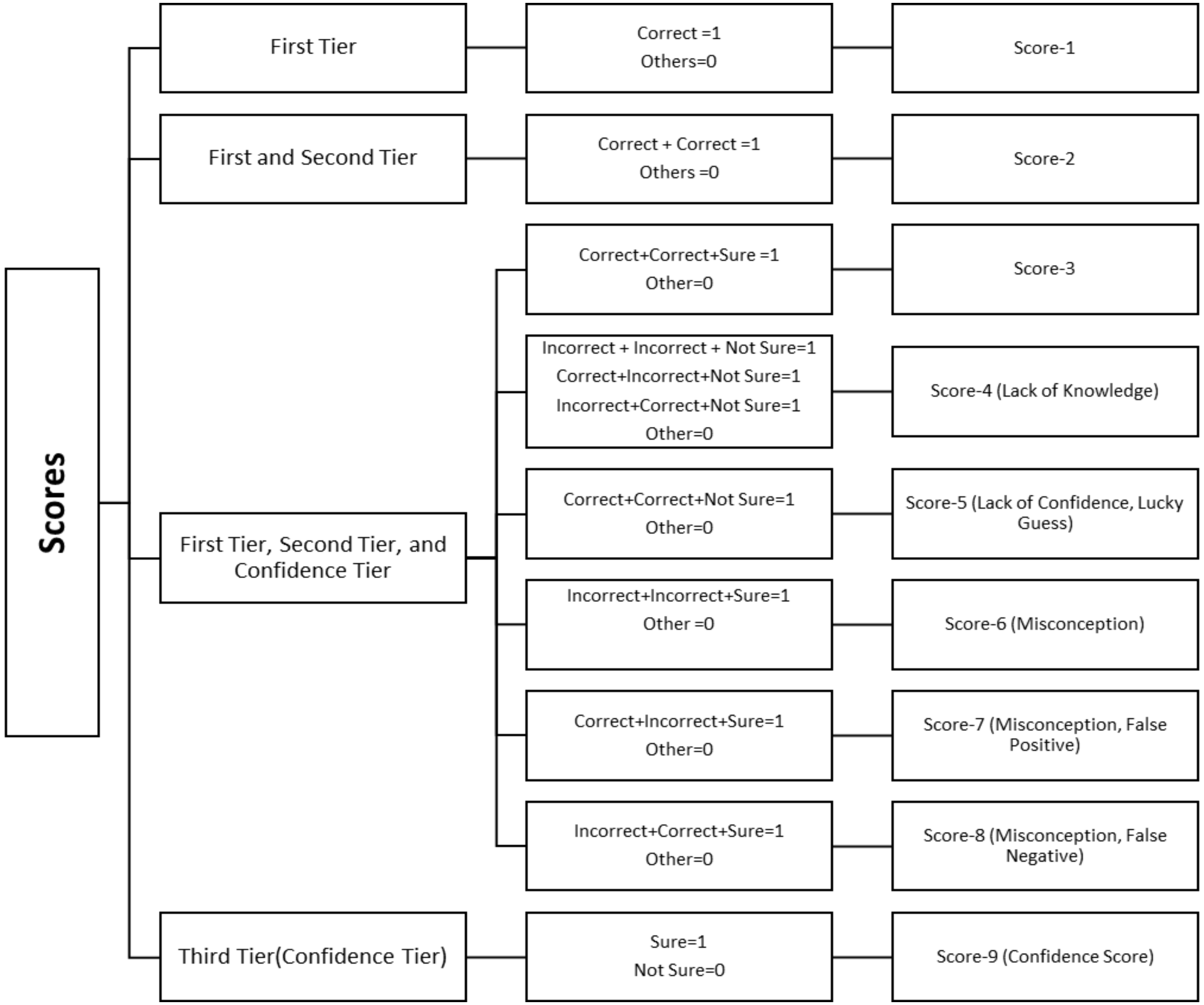


Figure 1. *The Coding Procedure for Scores.*

Score-1: This type of score was calculated from the students' answers to the first tier. Accordingly, if the students answered correctly, they received 1 point; if they answered incorrectly, they received zero points.

Score-2: This type of score was calculated from the students' answers to the first and second tiers. Accordingly, if they answered correctly in the first and second tiers, they were given 1 point; if they answered incorrectly in both tiers or in one of the tiers, they were given zero points.

Score-3: The students' answers to all three tiers were taken into consideration in this type of score. Accordingly, if the students answered correctly in the first and second tiers and said that they were sure about their answers, they received 1 point. In other cases, they received zero points.

Score-4 (Lack of Knowledge): This score was also calculated by taking the students' answers to all three tiers into consideration. Accordingly, if the students gave incorrect answers in the first and/or second tiers and stated that they were not sure about their answers, they received 1 point. In other cases, they received zero points. The scores the students received for each question demonstrated the students' lack of knowledge, in terms of the relevant concepts. In this way, the classification of each incorrect answer as a misconception was avoided.

Score-5 (Lack of Confidence or Lucky Guess): Here, the students also received 1 point, if they answered correctly in the first and second tiers and if they were not sure about their answers. They received zero points in other cases. In this case, the fact that they had been asked to state whether or not they were sure about their answers, even if they had answered correctly, demonstrated the lucky guess factor.

Score-6 (Misconception): If the students chose the incorrect option as the correct answer, chose the option to explain their reason for the answer and stated in the third tier that they were sure about their answer, then they were considered to have misconceptions. The students received 1 point for each question, as they answered incorrectly in the first and second tiers and they stated in the third tier that they were sure about their answer. In other cases, they received zero points.

Score-7 (Misconception, False Positive): This type of score was calculated when the students answered correctly in the first tier and incorrectly in the second tier, and in the third tier stated that they were sure about their answer. Here, the students received 1 point for each question they answered, and in other cases they received zero points, as the students had given correct answers with incorrect reasoning.

Score-8 (Misconception, False Negative): This type of score was calculated when the students answered incorrectly in the first tier and correctly in the second tier and stated in the third tier that they were sure about their answers. Here, they received 1 point for the question they answered, and they received zero points in other cases. In this case, students gave incorrect answers but with correct reasoning.

Score-9 (Confidence Score): In this type of score, only the students' answers to the third tier were essential. If they marked the option "I am sure," they received 1 point; and if they marked the option "I am not sure," they received zero points.

FINDINGS

The findings obtained in the study are presented under the following headings: results of content, face and construct validity (a), results of item analyses (b), and results of the reliability coefficient (c).

a) Results of Content, Face and Construct Validity

Firstly, the content, face and construct validity of the CBCT was analysed. Initially, expert opinion was obtained in relation to the content and face validity of the test. Two instructors of chemistry education and a chemistry teacher evaluated the test. The experts analysed the questions in the test, in terms of its level of suitability for high school students, its appropriateness to the school curriculum and the psychometric properties in the test, linguistic understandability and instructions for the questions. The items of the CBCT were revised, based on the suggestions and comments submitted by the experts. According to the experts' suggestions and comments, the distractors, linguistic understandability and instructions for the questions were revised. In addition, false positive and false negative values were also calculated for content validity - as is indicated by Hestenes and Halloun (1995). Researchers state that false positive and false negative proportions of less than 10% in a multiple-choice test increase content validity. In this study, the average percentage was found to be 6.51% for false positive and 6.02% for false negative. Therefore, having both false positive and false negative values of less than 10% is the evidence for content validity in this study. Çataloğlu (2002) recommends looking at the correlations between high scores received from the test and confidence levels to check for construct validity, and advocates that students with higher scores should have more self-confidence than those with lower scores. Çataloğlu (2002) and Peşman and Eryılmaz (2010) emphasise that there should be at least a moderate level of positive correlations between the two-tier scores and confidence tier scores. Therefore, the correlations between score-2 and score-9 were analysed in this study. The analyses revealed that there were positive and statistically significant correlations between the two types of scores ($r=.73$, $n=175$, $p<.0005$). Thus, the test was considered to have attained construct validity.

b) Results of Item Analyses

In this study, item analyses were made on the basis of the difference between the group averages of top 27% of the students and the bottom 27% of the students (Wiersma & Jurs, 1990). Item analyses for the items in the test were made on the basis of score-3. Table 2 shows the figures for the results concerning the analyses. Accordingly, item difficulty indices in the first column receive values between .47 and .77. The fact that the item difficulty indices are approximately .50 shows that those items are better and more useful. Crocker and Algina (2006) point out that the item difficulty index is not a property necessary for item quality, but that the value for the item difficulty index should be approximately .50 in order for the total variance and reliability to increase. However, it is stated in the literature that the item difficulty index can also range between .10 and .80 on the condition that it is approximately .50 for many items (Walsh & Betz, 1990). Another study emphasises that the value should in the .50- .80 interval (Wiersma & Yurs, 1990). Based on this, it was considered that the item difficulty indices for the items in the CBCT were at an adequate level. Another index shown in the table is the discrimination index. It shows to what extent the items discriminate individuals. Ebel (1965) points out that items with a discrimination index of .40 and above are appropriate but that the items with a discrimination index between .30 and .39 can be used in tests with little or no revision (As cited in, Crocker & Algina, 2006). That the discrimination index for the 15 items in the test was above .30 demonstrated that all the items distinguished between successful and unsuccessful students in terms of chemical bonding. An examination of item point biserial coefficient showed that the value was above .20, except for item 15. Item point biserial coefficients represent the correlations between students' correct answer to an item in the test and their correct answers to the other items in the test. A value above .20 is an acceptable value (Beichner, 1994; Peşman, & Eryılmaz, 2010; Wuttiprom *et al.*, 2009); nevertheless, the minimum acceptable point biserial value for the sample of 175 was found to be .15 according to the formula suggested by Crocker and Algina (2006), depending on the sample size. Thus, item 15 was not removed from the test, as it had the minimum acceptable point biserial value, which was adequate for this study, according to Crocker and Algina (2006). Besides, Crocker and Algina (2006) state that it would be more beneficial to use item discrimination indices in determining problematic items, and that the other values (item difficulty) could also be used in analysing the structure of questions. For this reason, the fact that the item discrimination index for question 15 was greater than .30 was interpreted as the question not being problematic.

Table 2. *Item Analyses*

Item	Item Difficulty Index (Pj)	Item Discrimination index (rj)	Item Variance	Item Standard Deviation	Item point biserial correlation coefficient
M1	.63	.62	.23	.48	.43
M2	.54	.70	.25	.50	.37
M3	.59	.62	.24	.49	.40
M4	.77	.38	.18	.42	.35
M5	.70	.55	.21	.46	.46
M6	.64	.30	.23	.48	.22
M7	.47	.51	.25	.50	.27
M8	.68	.51	.22	.47	.33
M9	.65	.62	.23	.48	.43
M10	.72	.43	.20	.45	.23
M11	.63	.70	.23	.48	.44
M12	.67	.62	.22	.47	.41
M13	.59	.57	.24	.49	.36
M14	.73	.45	.20	.44	.35
M15	.65	.36	.23	.48	.15

c) Results of Reliability Coefficient

Cronbach's Alpha coefficients were calculated separately for score-1, score-2 and score-3. The values calculated for all three types of scores are shown in Table 3. The fact that the Cronbach's alpha coefficients calculated from these three types of scores were above .70 showed that the test was reliable (Cronbach, 1951; George & Mallery, 2003; Pallant, 2005).

Table 3. *Reliability Statistics*

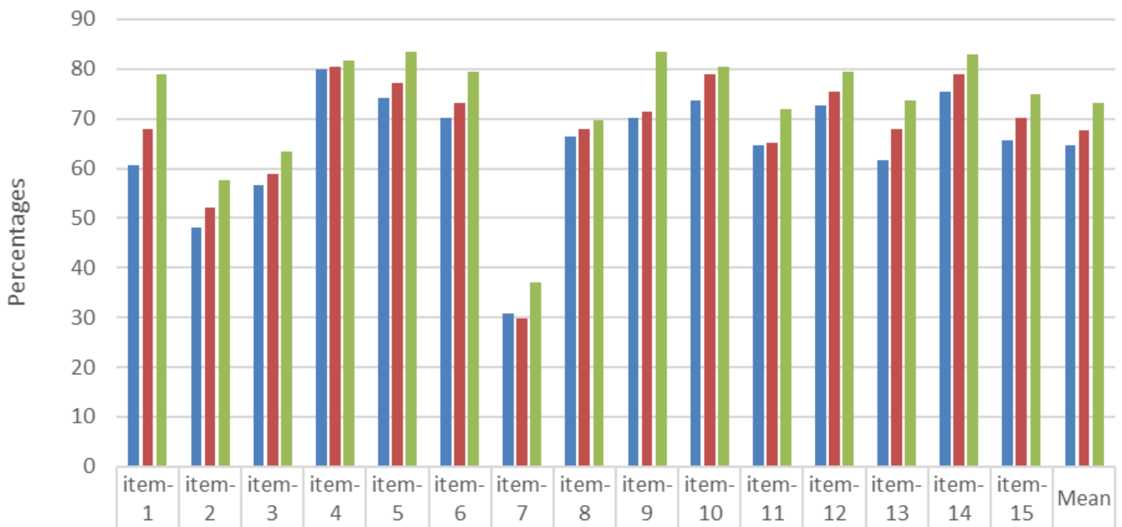
	Cronbach's Alpha	N of Items
Score-3	0,74	15
Score-2	0,77	15
Score-1	0,72	15

Table 4 shows the percentages for false negatives, false positives, lack of knowledge, lack of confidence and misconceptions. These types of scores exhibit the importance of three-tier tests. As is clear from the table, on average the percentage for lack of knowledge is 20.20%. Hence, the value demonstrates that all of the answers in a three-tier test should not be perceived as misconceptions. If one or two-tier tests had been used, students' incorrect answers due to lack of knowledge would have been considered as misconceptions. It is also clear from the table that question two is the item in which students have the most lack of knowledge. The question is about whether or not the electrons in covalent and hydrogen bonds are shared. An examination of lack of confidence percentages shows that the percentage is 16.33%. According to this result, some of the students were not confident about their correct answers or their correct answers were lucky guesses. On examining the values for score-6 (misconceptions), it was found that students had the most frequent misconceptions in the first four questions. It was apparent that the largest number of misconceptions related to question four in particular. This question was related to metallic bonds. It was also found that for question 13 the percentage for false negatives exceeded 10%. Yet, due to the fact that the value was not very high and that the average percentage for false negatives was less than 10%, it was regarded that content validity was attained. Likewise, on examining the false positives for each item, it was found that they were less than 10% except for items 11 and 13. Since false positive values for items 11 and 13 were 10.9% and 10.35, respectively and the average for all items was 6.52%, it was deemed that content validity was attained.

Table 4. *Percentages of False Negatives, False Positives, Lack of Knowledge, Lack of Confidence and Misconception*

	Item																
Scores	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**M**	**SD**
Score-4 (Lack of Knowledge)	15	35	22	14	16	11	26	23	20	14	25	16	23	17	26	20,2	6,34
Score-5 (Lack Of Confidence)	42	21	25	9	13	18	47	13	13	7	10	9	9	5	4	16,3	12,8
Socre-6 (Misconception)	39	38	35	42	34	33	33	34	30	27	21	31	22	26	21	31,1	6,54
Score -7 (False Positive)	3,43	3,43	3,43	3,43	4,57	8,57	5,14	5,14	7,43	6,29	10,9	8,57	10,3	8,57	8,57	6,52	2,63
Score -8 (False Negative)	2,29	4	4	2,29	4,57	4,57	4	6,29	5,71	6,86	8	8	10,9	9,14	9,71	6,02	2,67

Figure 2 shows the percentages for the students' correct answers to the tiers (score-1, score-2, and score-3). Accordingly, it was found that the rate of correct answers to the three tiers of the questions decreased. The reason for this was that the probability of lucky guesses was lower in three-tier tests than in one or two-tier tests. What is important here is the reduction in the number of the students' lucky guesses. According to the graph, while the rate of giving correct answers for the 15 questions in the test was 73% in the first tier, the rate fell to 68% on evaluating the first and second tiers together. In score-3, however, in which the students were required to answer correctly in the first and second tiers and to state that they were sure about their answers, the average percentage fell to 65%.



	item-1	item-2	item-3	item-4	item-5	item-6	item-7	item-8	item-9	item-10	item-11	item-12	item-13	item-14	item-15	Mean
Score-3	61	48	57	80	74	70	31	66	70	74	65	73	62	75	66	65
Score-2	68	52	59	81	77	73	30	68	71	79	65	75	68	79	70	68
Score-1	79	58	63	82	83	79	37	70	83	81	72	79	74	83	75	73

Figure 2. *Percentages of Correct Responses in terms of Number of Tiers.*

DISCUSSION and CONCLUSION

This study has made an attempt at developing a three-tier concept test so as to determine students' misconceptions about chemical bonding. The test contained 15 questions and three tiers for each question. The first tier of the test was the content tier, and it was directed to finding out what the participating students knew. The second tier was for ascertaining the students' reasons for their answers. The distractors included in these two tiers contained the common misconceptions reported in the literature. In the third tier, on the other hand, the students were asked whether they were sure about their answers in the first two tiers, or not. Prior to the implementation of the test, expert opinion was received for content and face validity. Having received expert opinion, a pilot study was implemented with 25 high school students, and any necessary modifications were made to the test. The false positive and false negative values were calculated for the content validity of the test, and the value was found to be below 10%. Thus, the test was considered to attain content validity. For the construct validity of the test, the correlations between two-tier scores and confidence tier scores were analysed through correlation analysis, and the correlations between the two types of scores were found to be positive and significant. Thus, the test was regarded as having attained construct validity.

In order to carry out item analyses, the top 27% of students and the bottom 27% of students were determined. Consequently, it was observed that the item difficulty indices took on values in the .47-.77 interval. The values demonstrate the relatively difficult and relatively easy questions in the test. The questions in the test, however, were at a medium level of difficulty. That the discrimination indices for all the questions were more than .30 exhibited that all the items in the test discriminated between the successful and unsuccessful students. Cronbach's Alpha coefficients for the test were calculated for score-1, score-2 and score-3. As the values were above .70, this was interpreted as the test being reliable.

At the end of this study, the students' answers to the first three tiers combined (score-3), the answers they gave to the first two tiers combined (score-2), and finally their answers to the first tier only (score-1) were evaluated. In consequence, it was found that students' correct answers to the questions decreased gradually. According to these results, as well as other studies in the literature, three-tier tests are more useful than conventional multiple-choice tests in determining students' misconceptions, their incorrect answers stemming from lack of knowledge and identifying those students who lack confidence, as well as for pointing out their correct answers by means of lucky guesses (Çetin-Dindar & Geban, 2011; Kirbulut *et al.*, 2010; Peşman & Eryılmaz, 2010). The use of such tests in education would also contribute to the recognition of any misconceptions.

There are also restrictions in three-tier tests. The most important of these restrictions is that these tests are not adequate for determining whether students are sure about their answers to the first and second tiers if they state in the third tier that they are not sure. Therefore, four-tier tests can be developed instead of three-tier tests (Caleon & Subramaniam, 2010). In a four-tier test, students can be asked if they are sure about their answers in both tiers.

High school teachers can easily use the CBCT developed in this study because these tests are more reliable and valid than the alternatives. For this reason, they are thought to be effective instruments in determining students' achievement and in uncovering their misconceptions. Furthermore, feedback received by high school teachers will also shed light on the development of such tests.

REFERENCES

Acar, B., & Tarhan, L. (2008). Effects of cooperative learning on students' understanding of metallic bonding. *Research in Science Education, 38*, 401–420.

Arslan, H. Ö., Çiğdemoğlu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess preservice teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain. *International Journal of Science Education, 34*(11), 1667-1686.

Artdej, R., Ratanaroutai, T., Coll, R. K., & Thongpanchang, T. (2010). Thai Grade 11 students' alternative conceptions for acid–base chemistry. *Research in Science & Technological Education*, *28*(2), 167-183.

Aykutlu, I., & Şen, A. İ. (2012). Determination of secondary school students' misconceptions about the electric current using a three-tier test, concept maps and analogies. *Education and Science*, *37*(166), 275-288.

Barker, V. (1995). *A longitudinal study of 16-18 year olds' understanding of basic chemical ideas*. Unpublished Doctorate Thesis, Department of Educational Studies, University of York.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American journal of Physics*, *62*(8), 750-762.

Boo, H. K. (1998). Students' understandings of chemical bonds and the energetics of chemical reactions. *Journal of Research in Science Teaching, 35*(5), 569-581.

Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education, 32*(7), 939–961.

Çalık, M., & Ayas, A. (2005). A comparison of level of understanding of eighth-grade students and science student teachers related to selected chemistry concepts. *Journal of Research in Science Teaching, 42*(6), 638-667.

Çataloğlu, E. (2002). *Development and validation of an achievement test in introductory quantum mechanics: The quantum mechanics visualization instrument (QMVI)* (Order No. 3060014). Available from ProQuest Dissertations & Theses Global. (305491729). Retrieved from http://search.proquest.com/docview/305491729?accountid=11248.

Çetin-Dindar, A., & Geban, Ö. (2011). Development of a three-tier test to assess high school students' understanding of acids and bases. *Procedia Social and Behavioral Sciences, 15*,600-604.

Chandran, S., Treagust, D. F., & Tobin, K. (1987). The role of cognitive factors in chemistry achievement. *Journal of Research in Science Teaching, 24*(2), 145-160.

Coll, R. K. (2008). Chemistry learners' preferred mental models for chemical bonding. *Journal of Turkish Science Education, 5*(1), 22-47.

Coll, R. K., & Taylor, N. (2002). Mental models in chemistry: senior chemistry students 'mental models of chemical bonding. *Chemistry Education: Research and Practice in Europe, 3*(2), 175-184.

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt College.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of the literature related to concept development in adolescent science students. *Studies in Science Education, 5(1),* 61–84.

Eryılmaz, A., & Sürmeli, E. (2002). *Üç-aşamalı sorularla öğrencilerin ısı ve sıcaklık konularındaki kavram yanılgılarının ölçülmesi [The assessment of students' misconceptions about heat and temperature concepts by means of three-tier questions]*.

Paper presented at the 5th National Conference on Science and Mathematics Education. Retrieved June 6, 2016, from http://users.metu.edu.tr/eryilmaz/TamUcBaglant.pdf

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference (11.0 update)* (4th ed.). Boston: Allyn & Bacon.

Gürcay, D., & Gülbaş, E. (2015). Development of three-tier heat, temperature and internal energy diagnostic test. *Research in Science & Technological Education*, *33*(2), 197-217.

Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research. *American Journal of Physics*, *64*(10), 1316-1325.

Hazel, E., & Prosser, M. (1994). First-year university students' understanding of photosynthesis, their study strategies and learning context. *The American Biology Teacher, 56*(5), 274-279.

Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory. *Physics Teacher, 33*,502–506.

Hırça, N., Çalık, M., & Akdeniz, F. (2008). Investigating grade 8 students' conceptions of energy and related concepts. *Journal of Turkish Science Education*, *5*(1), 75-87.

Kanlı, U. (2015). Using a two-tier test to analyse students and teachers' alternative concepts in astronomy. *Science Education International*, *26*(2), 148-165.

Kaya, O. N. (2008). A student-centred approach: Assessing the changes in prospective science teachers' conceptual understanding by concept mapping in a general chemistry laboratory. *Research in Science Education, 38*(1), 91-110.

Kirbulut, D., Geban, O., & Beeth, M. E. (2010, July). Development of a three-tier multiple-choice diagnostic instrument to evaluate students' understanding of states of matter. In *European Conference on Research in Chemical Education (ECRICE). Krakow, Poland*.

Lin, S. W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and Mathematics Education*, *2*(2), 175-199.

Luxford, C. J., & Bretz, S. L. (2013). Moving beyond definitions: What student-generated models reveal about their understanding of covalent bonding and ionic bonding. *Chemistry Education Research Practice, 14*, 214−222.

Luxford, C. J., & Bretz, S. L. (2014). Development of the Bonding Representations Inventory to identify student misconceptions about covalent and ionic bonding representations. *Journal of Chemical Education*, *91*(3), 312-320.

Mirzalar Kabapınar, F. (2007). A literature review of student misconceptions on chemical bonding I: Intramolecular bonding. *National Education, 176*, 18-35.

Muller, D.A., Sharma, M.D., Eklund, J., & Reimann, P. (2007). Conceptual change through vicarious learning in an authentic physics setting. *Instructional Science, 35*(6), 519–533.

Nakhleh, M. B. (1992). Why some students don't learn chemistry. *Journal of Chemical Education, 69*(3), 191-196.

Nicoll, G. (2001). A report of undergraduates' bonding misconceptions. *International Journal of Science Education*, *23*(7), 707-730.

Othman, J., Treagust, D. F., & Chandrasegaran, A. L. (2008). An investigation into the relationship between students' conceptions of the particulate nature of matter and their understanding of chemical bonding. *International Journal of Science Education*, *30*(11), 1531-1550.

Pallant, J. (2005). *SPSS Survival Manual: A Step-by-Step Guide to Data Analysis Using SPSS*. Allen and Unwin: Sydney.

Peşman, H. (2005). *Development of a three-tier test to assess ninth grade students' misconceptions about simple electric circuits*. Unpublished master's thesis, Middle East Technical University, Ankara, Turkey.

Peşman, H., & Eryılmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *The Journal of Educational Research*, *103*(3), 208-222.

Peterson, R. F., & Treagust, D. F. (1989). Grade-12 students' alternative conceptions of covalent bonding and structure. *Journal of Chemical Education*, *66*(6), 459–460.

Reynolds, A. J., & Walberg, H. J. (1992). A structural model of science achievement and attitude: An extension to high school. *Journal of Educational Psychology, 84*(3), 371-382.

Şekercioğlu, A. G., & Kocakülah, M. S. (2008). Grade 10 Students' Misconceptions about Impulse and Momentum. *Journal of Turkish Science Education*, *5*(2), 47-59.

Şen, Ş. (2015). *Investigation of students' conceptual understanding of electrochemistry and self-regulated learning skills in process oriented guided inquiry learning environment* (Unpublished Dissertation). Hacettepe University, Ankara, Turkey.

Şen, Ş. (2016). The relationship between secondary school students' self-regulated learning skills and chemistry achievement. *Journal of Baltic Science Education, 15*(3), 312-324.

Şen, Ş., & Yılmaz, A. (2012). Effect of conceptual change texts assisted dual situated learning model on achievement. *Journal of Hacettepe University Education Faculty, 42*,367-379.

Şen, Ş., & Yılmaz, A. (2013). A phenomenographic study on chemical bonding. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 7*(2), 144-177.

Sreenivasulu, B. & Subramaniam, R. (2013). University students' understanding of chemical thermodynamics. *International Journal of Science Education, 35*(4), 601-635.

Taber, K. S. (2009). Challenging Misconceptions in the Chemistry Classroom: Resources to Support Teachers. *Educació química*, *4*, 13-20.

Taber, K. S. (2011). Models, molecules and misconceptions: A commentary on" secondary school students' misconceptions of covalent bonding". *Journal of Turkish Science Education*, *8*(1), 3-18.

Tan, D. K. C., & Treagust, D. F. (1999). Evaluating students' understanding of chemical bonding. *School Science Review*, *81*(294), 75-84.

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*, 283–301.

Temel, S., & Özcan, Ö. (2016). The analysis of prospective chemistry teachers' cognitive structure: the subject of covalent and ıonic bonding. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(8), 1953-1969.

Thompson, F., & Logue, S. (2006). An exploration of common student misconceptions in science. *International Education Journal, 7*(4), 553-559.

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.

Tsaparlis, G., & Papaphotis, G. (2002). Quantum-chemical concepts: Are they suitable for secondary students? *Chemistry Education Research and Practice*, *3*(2), 129-144.

Ünal, S., Coştu, B., & Ayas, A. (2010). Secondary school students' misconceptions of covalent bonding. *Journal of Turkish Science Education*, *7*(2), 3-29.

Uzuntiryaki, E., & Geban, Ö. (2005). Effect of conceptual change approach accompanied with concept mapping on understanding of solution concepts. *Instructional Science, 33*(4), 311-339.

Voska, K. W., & Heikkinen, H. W. (2000). Identification and analysis of student conceptions used to solve chemical equilibrium problems. *Journal of Research in Science Teaching*, *37*(2), 160-176.

Walsh, W. B., & Betz, N. E. (2000). *Tests and assessments* (4th Ed.). Upper Saddle River, NJ: Prentice Hall.

Wiersma, W., & Jurs, S. G. (1990). *Educational measurement and testing* (Second Ed.). Needham Heights, MA: Allyn and Bacon.

Wuttiprom, S., Sharma, M. D., Johnston, I. D., Chitaree, R., & Soankwan, C. (2009). Development and use of a conceptual survey in introductory quantum physics. *International Journal of Science Education*, *31*(5), 631-654.

Yayon, M., Mamlok-Naaman, R., & Fortus, D. (2012). Characterizing and representing student's conceptual knowledge of chemical bonding. *Chemistry Education Research and Practice*, *13*, 248–267.

Yürük, N. (2005). *An analysis of the nature of students' metaconceptual processes and the effectiveness of metaconceptual teaching practices on students' conceptual understanding of force and motion*. Unpublished Dissertation, Ohio State University, Ohio, USA.