Validity and Reliability of the Two-tier Diagnostic Test to Identify Students' Alternative Conceptions of Intermolecular Forces

Asriah N. Mardiyyaningsih[1], Erlina[2*], Maria Ulfah[2], Azwa Fadilla Wafiq[2]

[1] Department of Biology Education, FKIP, Universitas Tanjungpura, Pontianak, Indonesia.
[2] Department of Chemistry Education, FKIP, Universitas Tanjungpura, Pontianak, Indonesia.

Abstract: Alternative conceptions or misconceptions is believed as the main barrier for most students who learn chemistry. Students who hold alternative conceptions will struggle to understand the advanced concepts. Therefore, identifying those alternative conceptions is important. In this study a diagnostic test was designed and developed to identify students' alternative conceptions about intermolecular forces (IMFs). The diagnostic test developed in this study consist of 19 multiple choice questions with open reason. The test was administered to 88 university students who learn chemistry. Data collected were analyzed using SPSS. The validity and the reliability of the test were 0.526 and 0.878 respectively. The result indicates that the test is valid and reliable. The mean difficulty value and discriminatory index are 0.623 and 0.45 respectively. The findings of the item analysis showed that the test is in moderate difficulty and the discriminatory of the test is in good category. Meanwhile, the average distractor effectivity for the diagnostic test is 3 options (out of 4 options), which means the distractors are in good category and well function. Those findings indicate that the two-tier diagnostics test could be used to identify students' alternative conceptions and understanding about IMFs.

Keywords: Diagnostic test; Item analysis; Misconceptions; University student

Introduction

Alternative conceptions or well known as misconceptions is believed as the main issue of teaching and learning chemistry. Alternative conception is known by other terms such as misconceptions, alternate conceptions, alternative frameworks, and pre conceptions, but in this study the author prefer the term alternative conceptions to address students' understanding. According to Cho et al. (1985), alternative conception is "any conceptual idea whose meaning deviates from the one commonly accepted by scientific consensus". Due to the nature of the subject, chemistry relies on a number of models that rely on abstract concepts (e.g. models of atomic and molecular structure). These models are fundamental in understanding the more complex topics that students encounter in the later stages of their education. A means of identifying and resolving misconceptions in fundamental topics like these is an important consideration when designing a chemistry curriculum. In addition to the difficulty of interpreting representations, Keig & Rubba (1993) found that that majority of students were struggling to translate among formulae, electron configurations, and ball-and-stick models.

Nakhleh (1992) proposing a potential answer to that which is many students are not constructing decent comprehension of fundamental chemical concepts from the beginning of their studies, thus they cannot fully understand the more advanced concepts that build upon the basic concepts. This is also supported by the work of

---

How to Cite:
Mardiyyaningsih, A.N., Erlina, E., Ulfah, M., & Wafiq, A.F. (2023). Validity and Reliability of the Two-tier Diagnostic Test to Identify Students' Alternative Conceptions of Intermolecular Forces. *Jurnal Penelitian Pendidikan IPA*, 9(6), 4375–4381. https://doi.org/10.29303/jppipa.v9i6.2797

Taber (2009). A number of studies have shown that misconception occurs in the teaching of many fundamental topics in chemistry (Nicoll, 2001; Taber, 2003). In addition to that, Taber (2009) suggested students' prior educational experience is the primary contributory factor to misconceptions.

Understanding the physical basis, the principles required to describe intermolecular forces and consequences of Intermolecular Forces (IMFs) is an essential element in core chemistry education (Cooper et al., 2015; Kind, 2004; Tarhan et al., 2008). An understanding of intermolecular forces helps students predict a number of physical properties of substances, such as relative boiling points, changes in states of matter Schmidt et al. (2009) and the ability to predict whether a given solute will be soluble in a particular type of solvent. However, many students find it difficult to understand the concepts (Birk & Kurtz, 1999; Ma'rufah et al., 2022; Peterson & Treagust, 1989; Tan & Chan, 2003; Widarti et al., 2019). Therefore, students develop a wide range of alternative conceptions (Coll & Taylor, 2001). Tan & Chan (2003) found that some students at grade 5 and 6 (age 16-17 years and 17-18 years) had difficulty in understanding the nature of hydrogen bonding and dipole-dipole interactions. Tan & Chan (2003) found that some students find it hard to describe the intermolecular forces involved within the molecules. Students also found it hard to differentiate between intra and intermolecular forces (Vladušić et al., 2016).

Misconceptions or alternative conceptions in chemistry can be identified using conceptual inventory for certain topics, such as the chemical concept inventory developed by Mulford & Robinson (2002), diagnostic test by (Halim et al., 2020; Milenković et al., 2016) , two-tiered multiple-choice diagnostic test by (Laliyo et al., 2019; Loh et al., 2014; Yamtinah et al., 2019). One of the most popular tools for determining conceptual learning level and misunderstandings has been diagnostic tests (Odom & Barrow, 1995; Tan, Goh, Chia, & Treagust, 2002; Treagust, 1988).

Identify students' alternative conception is crucial to improve the teaching and learning process (Sadler & Sonnert, 2016). When students have misconceptions, they may have difficulty understanding new information or applying what they have learned to solve problems. Addressing students' alternative conception will help teachers create more effective learning experiences in developing accurate and complete understanding of the concepts.

In this study, the diagnostics test developed is a two-tier multiple-choice test. The multiple-choice questions on the two-tier test are divided into items with two sub-questions (tiers) (Ivanjek et al., 2021). Moreover, (Ivanjek et al., 2021) mentioned that two-tier instruments

have the benefits of being simple to use and of giving insight into students' thought processes.

The first tier of the diagnostics test is the multiple choices test, while the second tier is the open-ended question to state their reason about the answer in the first tier. Thus, students were asked to choose the answer in the first tier and mention the reason in the second tier. By doing this, students' alternative conceptions will reveal (Treagust, 1988) as well as determined whether the misconceptions that students held relate to their previous alternative conceptions (Loh et al., 2014; Mann & Treagust, 1998; Uyulgan et al., 2014). The explanations provided by students are crucial for teaching scientific concepts. In two-tier assessments, they have the chance to choose an answer and its justification, and teachers also discover the causes of the students' misconceptions (Cengiz, 2009). In this article we present the development of the two-tier diagnostics test of Intermolecular Forces and assessing the quality of the test in term of item analysis, consist of validity, reliability, item difficulty, discriminatory index, and distractor information.

Method

Participants

There were 88 students of 3 different cohort of Chemistry Education Study Program of Department of Science and Mathematics Education, Universitas Tanjungpura who voluntarily participated in this study. The number of participants for each cohort can be seen on Table 1.

Table 1. Number of participants for each cohort

Cohort	Number of Participants
2019	22
2020	31
2021	34
Total	88

Methods and Procedures

This study employed a quantitative method in order to answer the research questions. All data collected were analyzed quantitively using IBM SPSS 27 to measure the quality of the diagnostic test developed. The procedures of the developing the diagnostics test were adapted to procedures proposed by (Peterson, Treagust, & Garnett, 1989) with some modifications. Peterson et al. (1989) proposing 3 steps, which are, initial testing, paper and pencil test, and test development validation. In this study, researcher proposing another two steps, therefore the procedures are consisting of: 1) Reviewing the concepts, in this step reviewing the chemistry curriculum and relevant studies on the topic of intermolecular forces to extract the important

concepts that will be asked. This process is the initial process in preparing the initial test. 2) Initial testing, this stage starts by administering a set of questions based on the first stage. The questions could be some open questions, or a multiple choice with open reason, to gather students' initial understanding. 3) Developing the prototype of the diagnostics test, based on students' answer in the previous stage the prototype of questions was design and developed to address all the possible alternative conceptions found in the second stage. 4) Validating the test, the prototype of the resulting diagnostics test will then be content validated by experts and empirically by involving participants to analyze the quality in terms of validity, reliability, difficulty index, discriminatory index and the effectiveness of the distractor. 5) The finalization of diagnostics test, in this step, the diagnostic test instrument will be revised based on the results of the empirical validity in the previous step, so that the final product is obtained, namely a two-tier test instrument that is feasible to use to identify students' alternative conceptions.

Development Process of the Diagnostics Test Validity

Validating the diagnostics test developed in this study consist of two types, the first is content validity by the experts (the experienced lecturer who taught chemistry), the second is the empirical validity by administering the test to students who already learn the topic. Both validities are aimed to measure the quality of the diagnostics test. The result of the content validity then analyzed using experts judgement score suggested by (Gregory, 2015). The score ranges from 0 to 1 (see table 3). It is done by making contingency tables on two experts, with the first category that is not relevant and less relevant become the weak relevancy category, and the second category which is for quite relevant and very relevant that is created in a new strong relevant category. The experts' judgement score for content validity is a comparison of the number of items of the two experts with strong relevance category of overall items.

Table 2. Interrater agreement model for content validity

		EXPERT JUDGE #1	
		Weak Relevance (Item rated 1 or 2)	Strong Relevance (Item rated 3 or 4)
EXPERT JUDGE #2	Weak Relevance (Item rated 1 or 2)	A	B
	Strong Relevance (item rated 3 or 4)	C	D

The experts judgement for content validity is a comparison of the numbers of items from two experts as validators with strong relevance to the overall items category (Gregory, 2015). While the results of the relevancy tabulation (contingency tables) are presented in Table 2, the validity coefficient is presented in Formula 1.

$$\text{Content validity} = \left(\frac{D}{A+B+C+D} \right) \quad (1)$$

Description:
A: The number of items in cell A (weak-weak relevance)
B: The number of items in cell B (strong-weak relevance)
C: The number of items in cell C (weak-strong relevance)
D: The number of items in cell D (strong-strong relevance)

Table 3. Validity criteria of the item

Range Score	Category
0.80 – 1.00	Very Good
0.60 – 0.79	Good
0.40 – 0.59	Enough
0.20 – 0.39	Poor
0.00 – 0.19	Very Poor

Meanwhile, in the empirical validity, students' scores were analyzed using SPSS. The R-value of Pearson Correlation obtained from the SPSS analysis compared to R-table to conclude whether each item is valid or not.

Reliability

Internal consistency reliability method was used in the reliability study of the diagnostic test. In this method, each included in the test is analyzed after the diagnostic test is administered. After the analysis, Cronbach Alpha coefficient, which indicates to what extent the items consistent to each other is calculated. If the coefficient is higher than 0.60 the test is reliable (Ghozali, 2016). The item analysis is measured using IBM SPSS 27. In the item analysis difficulty index, discrimination index, and distractor effectiveness were analysis separately in addition to the reliability of the diagnostics test.

The category of difficulty index presented in table 4. Meanwhile, category of discrimination index and distractor effectiveness presented in table 5. The analysis was done using Microsoft Excel. The findings then compared to those categories in table 4 and 5.

Table 4. Category of difficulty index

Range of Difficulty Index	Category
0 – 0.30	Difficult
0.31 – 0.70	Moderate
0.71 – 1	Easy

(Arikunto, 2016)

Table 5. Category of discrimination index and distractor effectiveness

Discrimination Index		The Function of the Options (Distractor Effectiveness)	
Range	Category	Criteria	Category
< 0.20	Poor	Items with 4 functional distractors	Excellent
0.21 – 0.40	Satis-factory	Items with 3 functional distractors	Good
0.41 – 0.70	Good	Items with 2 functional distractors	Acceptable
0.71 – 1.00	Excellent	Items with 1 functional distractor	Remediable
Negative value	-	Nonfunctional distractors	Discarded

(Arifin, 2016; Sudijono, 2020)

Result and Discussion

The quality of the diagnostic test

Based on the review of chemistry curriculum and some relevant published papers, the concepts of IMFs for the diagnostic test consist of Intermolecular forces, London forces, Dipole-dipole forces, and Hydrogen bond. Nine indicators of the questions were included based on those concepts. Detail of concepts covered in the diagnostic test along with the indicators are presented in Table 6.

Table 6. Concepts covered in the diagnostic test

Concepts	Indicator	Item
Intermolecular forces (IMFs)	Definition of IMFs	1
	The influence of the strength of Intermolecular forces to boiling point	17
London Force	Example dan definition of London force (Dispersion force).	4, 5, 11, 12, 13, 20
	The strength of London Force	16
	Comparison of London force and dipole-dipole force	10
Dipole-dipole force	Definition and example of dipole-dipole molecules	2,3
Hydrogen bond	Definition of Hydrogen bond	15
	The strength of the Hydrogen bond	6, 7, 14*
		8, 9, 18,
	The process of the forming of Hydrogen bond	19

Note: item no 14 is not valid

The validity
Content validity

Based on the analysis of expert judgements, the value of the content validity is 1 which is categorized as very good (see table 3). The result indicates that the content validity of the diagnostic test is high. Therefore, the test could be the continue for further analysis, which are empirical validity and reliability.

Empirical validity

The validity of the diagnostics test was calculated using IBM SPSS 27. The result of the average R-value is 0.526 while the R-table is 0.209, it means that the test categorized as valid. Detail of R-value of each item can be seen on table 7. Among those items, item no 14 has the lowest R value which is -0.080 which below the R-table, which mean the item is not valid. Due to item No. 14's lack of distinctiveness and usability, it was eliminated from the test. Thus, the item no 14 was removed from the diagnostic test. Detail of the empirical validity is tabulated in Table 7.

Table 7. Critical values of the Pearson correlation coefficient (R)

					Item
	1	2	3	4	5
	.578**	.534**	.340**	.551**	.714**
	6	7	8	9	10
Pearson Correlation	.673**	.351**	.731**	.603**	.574**
	11	12	13	14	15
	.671**	.648**	.593**	- .080	.375**
	16	17	18	19	20
	.441**	.215*	.692**	.666**	.668**

The reliability

Based on data analysis gained from SPSS, the coefficient of Cronbach Alpha is 0.878 (see table 8). According to (Ghozali, 2016) the coefficient of Cronbach Alpha > 0.60 is considered as very reliable. The result indicates that the diagnostic test is providing internal consistency (Eisingerich & Rubera, 2010). In addition to that, this result means that the diagnostic test could be used to measure students' understanding and alternative conceptions about IMFs.

Table 8. The reliability of diagnostic test

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.878	.879	19

Based on the item analysis of students' score of the diagnostic test, the result of the difficulty and discriminatory index is tabulated in table 9. As can be seen on table 9, the range of difficulty index is 0.39 - 0.82 between easy to moderate (see table 4). The percentage of the easy and moderate of the items are 21% and 79% (see Figure 1). Meanwhile, the range of the discriminatory index is 0.14 - 0.68, categorized as less, satisfactory, and good. Details of each category is presented in Figure 2. The average of difficulty and discriminatory index are 0.623 and 0.45 respectively. The findings of both indexes illustrated that the test is in

moderate difficulty and the discriminatory of the test is in good category.
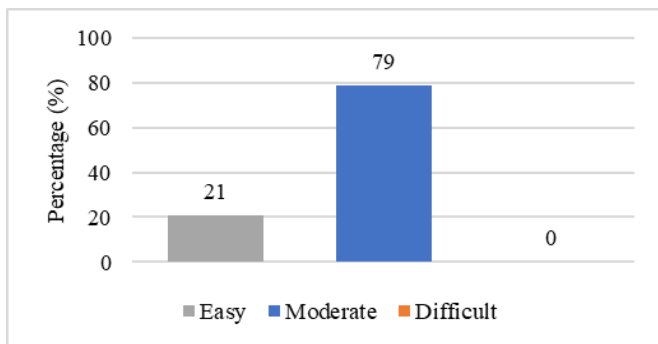

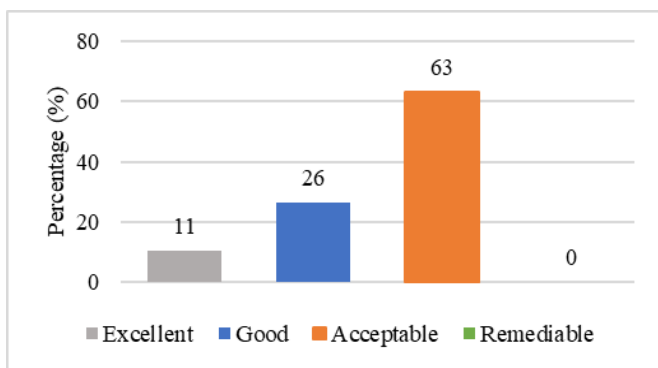**Figure 1**. Percentage of difficulty index for each category

Figure 2. Percentage of discrimination index for each category

Table 9. Difficulty index and discrimination index of each item

Item	Difficulty Index		Discrimination Index	
	Value	Category	Value	Category
1	0.68	Moderate	0.41	Good
2	0.67	Moderate	0.43	Good
3	0.82	Easy	0.18	Poor
4	0.81	Easy	0.30	Satisfactory
5	0.5	Moderate	0.68	Good
6	0.78	Easy	0.43	Good
7	0.55	Moderate	0.32	Satisfactory
8	0.57	Moderate	0.68	Good
9	0.63	Moderate	0.52	Good
10	0.53	Moderate	0.57	Good
11	0.57	Moderate	0.64	Good
12	0.53	Moderate	0.57	Good
13	0.57	Moderate	0.45	Good
15	0.59	Moderate	0.27	Satisfactory
16	0.64	Moderate	0.36	Satisfactory
17	0.61	Moderate	0.14	Poor
18	0.76	Easy	0.39	Satisfactory
19	0.39	Moderate	0.59	Good
20	0.65	Moderate	0.57	Good
Average	0.62	Moderate	0.45	Good

The last analysis is the distractor effectiveness. The percentage of each option for every category are 42% (excellent), 16% (good), 37% (acceptable) and 5%

(remediable) (see figure 3). The "excellent" category means that 4 options (out of 4) in the item are well function, "good" means 3 options (out of 4) are well function, "acceptable" means only 2 options are well function, and "remediable" means that only 1 option is well function. Even though, there is 1 (out of 19) item in the acceptable category, the average of well function options for the diagnostics test is 3 options for the rest of the items (18 out of 19), which means the distractors are in good category. Number of the items for each category is presented in Table 10. These findings implies that the diagnostics test is feasible to use to identify students' understanding and alternative conceptions of IMFs.
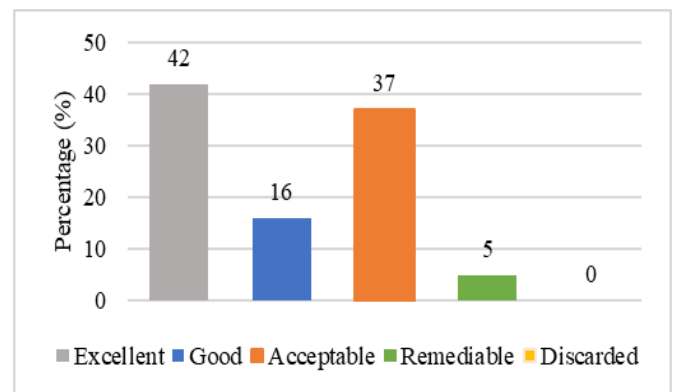

Figure 3. Percentage of distractor effectiveness for each category

Table 10. Item number for each category

Category	Total	Item
Excellent	8	5,9,10,11,12,13,15,19
Good	3	7,8,20
Acceptable	7	1,2,3,4,6,16,18
Remediable	1	17
Discarded	-	-

Conclusion

The feasibility of the diagnostic test is measured by item analysis, consist of validity, reliability, difficulty index, discriminatory index, and distractor effectiveness. The validity and the reliability of the test were 0.526 and 0.878 respectively. The result implies that the test is valid and reliable. Then, the results of the mean difficulty value and discriminatory index with the value of 0.623 and 0.45 respectively. The findings of the item analysis showed that the test is moderate difficulty and the discriminatory of the test is in good category. the average distractor effectivity for the diagnostic test is 3 out of 4 options which means the distractors are in good category and well function. Thus, it can be concluded that the two-tier diagnostic test can be used to identify students' alternative conceptions and understanding of Intermolecular Forces.

Conflicts of Interest
The authors declare no conflict of interest.

References

Arifin, Z. (2016). *Evaluasi Pembelajaran (Prinsip, Teknik, dan Prosedur), Cetakan Kedelapan*. Jakarta: Rosda Karya.

Arikunto, S. (2016). *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.

Birk, J. P., & Kurtz, M. J. (1999). Effect of Experience on Retention and Elimination of Misconceptions about Molecular Structure and Bonding. *Journal of Chemical Education*, *76*(1), 124. https://doi.org/10.1021/ed076p124

Cengiz, T. Y. Z. (2009). Development of Two-Tier Diagnostic Instrument and Assess Students' Understanding in Chemistry. *Scientific Research and Essays*, *4*(6), 626–631. Retrieved from https://academicjournals.org/journal/SRE/article-full-text-pdf/0B2C90A18467

Cho, H.-H., Kahle, J. B., & Nordland, F. H. (1985). An investigation of high school biology textbooks as sources of misconceptions and difficulties in genetics and some suggestions for teaching genetics. *Science Education*, *69*(5), 707–719. https://doi.org/10.1002/sce.3730690512

Coll, R. K., & Taylor, N. (2001). Alternative Conceptions of Chemical Bonding Held by Upper Secondary and Tertiary Students. *Research in Science & Technological Education*, *19*(2), 171–191. https://doi.org/10.1080/02635140120057713

Cooper, M. M., Williams, L. C., & Underwood, S. M. (2015). Student Understanding of Intermolecular Forces: A Multimodal Study. *Journal of Chemical Education*, *92*(8), 1288–1298. https://doi.org/10.1021/acs.jchemed.5b00169

Eisingerich, A. B., & Rubera, G. (2010). Drivers of brand commitment: A cross-national investigation. *Journal of International Marketing*, *18*(2), 64–79. https://doi.org/10.1509/jimk.18.2.64

Ghozali, I. (2016). *Aplikasi Analisis multivariete dengan program IBM SPSS 23 (Edisi 8). Cetakan ke VIII*. Semarang: Badan Penerbit Universitas Diponegoro.

Gregory, R. J. (2015). *Psychological testing: History, principles and applications seventh edition*. In Pearson Education.

Halim, A., Elmi, Elisa, Wahyuni, A., Ngadimin, Musdar, & Balqis, N. N. (2020). Development of concept maps diagnostic test for identification of students' misconceptions. *AIP Conference Proceedings*, *2215*, 050003. https://doi.org/10.1063/5.0000613

Ivanjek, L., Morris, L., Schubatzky, T., Hopf, M., Burde, J. P., Haagen-Schützenhöfer, C., … Wilhelm, T. (2021). Development of a two-tier instrument on simple electric circuits. *Physical Review Physics Education Research*, *17*(2), 20123. https://doi.org/10.1103/PhysRevPhysEducRes.17.020123

Keig, P. F., & Rubba, P. A. (1993). Translation of representations of the structure of matter and its relationship to reasoning, gender, spatial reasoning, and specific prior knowledge. *Journal of Research in Science Teaching*, *30*(8), 883–903. https://doi.org/10.1002/tea.3660300807

Kind, V. (2004). *Beyond appearances: Students' misconceptions about basic chemical ideas. School of Education, Durham University, UK.* London: Royal Society of Chemistry.

Laliyo, L. A. R., Botutihe, D. N., & Panigoro, C. (2019). The development of two-tier instrument based on distractor to assess conceptual understanding level and student misconceptions in explaining redox reactions. *International Journal of Learning, Teaching and Educational Research*, *18*(9), 216–237. https://doi.org/10.26803/ijlter.18.9.12

Loh, A. S. L., Subramaniam, R., & Tan, K. C. D. (2014). Exploring students' understanding of electrochemical cells using an enhanced two-tier diagnostic instrument. *Research in Science & Technological Education*, *32*(3), 229–250. https://doi.org/10.1080/02635143.2014.916669

Mann, M., & Treagust, D. (1998). A pencil and paper instrument to diagnose students' conceptions of breathing, gas exchange and respiration. *Australian Science Teachers Journal*, *44*(2), 55–59.

Ma'rufah, D. M., Effendy, & Wonorahardjo, S. (2022). Identification of intermolecular force misconceptions in students with different scientific thinking skills and the improvement efforts using conceptual change text. *Improving Assessment and Evaluation Strategies on Online Learning*, 123–128. https://doi.org/10.1201/9781003261346-19

Milenković, D. D., Hrin, T. N., Segedinac, M. D., & Horvat, S. (2016). Development of a Three-Tier Test as a Valid Diagnostic Tool for Identification of Misconceptions Related to Carbohydrates. *Journal of Chemical Education*, *93*(9), 1514–1520. https://doi.org/10.1021/acs.jchemed.6b00261

Mulford, D. R., & Robinson, W. R. (2002). An Inventory for Alternate Conceptions among First-Semester General Chemistry Students. *Journal of Chemical Education*, *79*(6), 739. https://doi.org/10.1021/ed079p739

Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education*, *69*(3), 191. https://doi.org/10.1021/ed069p191

Nicoll, G. (2001). A Report of Undergraduates' Bonding Misconceptions. *International Journal of Science Education*, *23*(7), 707–730. https://doi.org/10.1080/09500690010025012

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. https://doi.org/10.1002/tea.3660320106

Peterson, R. F., & Treagust, D. F. (1989). Grade-12 students' misconceptions of covalent bonding and structure. *Journal of Chemical Education*, *66*(6), 459. https://doi.org/10.1021/ed066p459

Peterson, R. F., Treagust, D. F., & 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. *Journal of Research in Science Teaching*, *26*(4), 301–314. https://doi.org/10.1002/tea.3660260404

Sadler, P. M., & Sonnert, G. (2016). Understanding Misconceptions Teaching and Learning in Middle School Physical Science. *American Educator, Spring*, 26–32. Retrieved from https://eric.ed.gov/?id=EJ1094278

Schmidt, H.-J., Kaufmann, B., & Treagust, D. F. (2009). Students' understanding of boiling points and intermolecular forces. *Chem. Educ. Res. Pract.*, *10*(4), 265–272. https://doi.org/10.1039/B920829C

Sudijono, A. (2020). *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Press.

Taber, K. S. (2003). Mediating mental models of metals: Acknowledging the priority of the learner's prior learning. *Science Education*, *87*(5), 732–758. https://doi.org/10.1002/sce.10079

Taber, K. S. (2009). Challenging Misconceptions in the Chemistry Classroom: Resources to Support Teachers. *Educació Química EduQ*, *4*(5), 13–20. https://doi.org/10.2436/20.2003.02.27

Tan, D. K. C., & Chan, K. S. (2003). Content framework for intermolecular forces. *Journal of Science and Mathematics Education in Southeast Asia*, *26*(2), 61–75. Retrieved from http://hdl.handle.net/10497/15466

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. https://doi.org/10.1002/tea.10023

Tarhan, L., Ayar-Kayali, H., Urek, R. O., & Acar, B. (2008). Problem-Based Learning in 9th Grade Chemistry Class: 'Intermolecular Forces.' *Research in Science Education*, *38*(3), 285–300. https://doi.org/10.1007/s11165-007-9050-0

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. https://doi.org/10.1080/0950069880100204

Uyulgan, M. A., Akkuzu, N., & Alpat, Ş. (2014). Assessing the students' understanding related to molecular geometry using a two-tier diagnostic test. *Journal of Baltic Science Education*, *13*(6), 839–855. https://doi.org/10.33225/jbse/14.13.839

Vladušić, R., Bucat, R. B., & Ožić, M. (2016). Understanding ionic bonding – a scan across the Croatian education system. *Chemistry Education Research and Practice*, *17*(4), 685–699. https://doi.org/10.1039/C6RP00040A

Widarti, H. R., Marfuaf, S., & Retnosari, R. (2019). Identifying Students' Misconception About Intermolecular Forces Topic in Organic Chmeistry I Course. *Unnes Science Education Journal*, *8*(1), 46–56. https://doi.org/10.15294/usej.v8i1.23527

Yamtinah, S., Indriyanti, N. Y., Saputro, S., Mulyani, S., Ulfa, M., Mahardiani, L., … Shidiq, A. S. (2019). The identification and analysis of students' misconception in chemical equilibrium using computerized two-tier multiple-choice instrument. *Journal of Physics: Conference Series*, *1157*(4). https://doi.org/10.1088/1742-6596/1157/4/042015