THE VALIDITY AND RELIABILITY OF FIVE-TIER CONCEPTION DIAGNOSTIC TEST FOR KINETIC THEORY OF GASES

Noriyatus S. Fajriyyah, Frida U. Ermawati
Physics Department, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya
Email: frida.ermawati@unesa.ac.id

Abstract

Misconceptions was reported to occur in many concepts in Physics subject in senior high school. One of them is in Kinetic Theory of Gases concepts. For example, students assumed if two balloons with different volumes are connected using a plastic straw, the air will flow from the bigger balloon to the small balloon. This assumption is not in accordance with the Kinetic Theory of Gases concept namely Boyle's Law, so it is called misconception. The aims of this study is to develop a five-tier conception diagnostic test for Kinetic Theory of Gases and determine the validity and the reliability. The author using two different samples, first, 27 students to collect reasons, second, 35 students to gain validity and reliability score. The validity aspect consists of internal and external validities. The internal validity has done by two Physics lecturers. The external validity consists of content and construct aspect. The external validity of the content was determined based on false positives (FP) and false negatives (FN) values that < 10%, while the construct external validity was determined using the Pearson Product Moment correlation equation. The reliability was calculated using Alfa Cronbach (r_{11}) with 5% significance level and $r_{theoritic}$ of 0.334. The internal validity results is 94.33%. The FP and FN values are 8.9 and 7.8% respectively, which means that the developed instrument is valid. The (r_{xy}) is 0.64, while the reliability value (r_{11}) is 0.885. Those values are greater than the $r_{theoritic}$, which means that the developed instrument is valid and reliable.

Keywords: Kinetic theory of gases, five-tier conception diagnostic test, validity, reliability

Abstrak

Miskonsepsi dilaporkan telah terjadi pada banyak konsep pada pelajaran Fisika. Salah satunya pada materi Teori Kinetik Gas. Sebagai contoh, peserta didik (PD) menganggap bahwa apabila dua balon dengan volume berbeda dihubungkan menggunakan sebuah plastik, maka udara akan mengalir dari balon bervolume besar menuju balon bervolume kecil. Anggapan ini tidak sesuai dengan konsep Teori Kinetik Gas yaitu Hukum Boyle, maka anggapan tersebut disebut dengan miskonsepsi. Tujuan dari penelitian ini yaitu mengembangkan instrumen tes diagnostik konsepsi berformat *five-tier* untuk materi Teori Kinetik Gas dan untuk menentukan tingkat validitas dan reliabilitasnya. Penulis menggunakan dua sampel yang berbeda yaitu 27 PD untuk menjaring alasan dan 35 PD untuk mendapatkan nilai validitas dan reliabilitas instrumen. Aspek validitas terdiri dari validitas internal dan validitas eksternal. Validitas internal ditentukan oleh dua dosen Fisika. Validitas ekternal terdiri dari isi dan konstruk. Validitas eksternal isi ditentukan berdasarkan hasil *false positives* (FP) dan *false negatives* (FN) harus < 10%, sedangkan validitas eksternal konstruk menggunakan persamaan korelasi *Pearson Product Moment.* Reliabilitas dihitung menggunakan *Alfa Cronbach* dengan taraf signifikansi sebesar 5% dan r_{tabel} yaitu 0,334. Nilai validitas internal yang diperoleh adalah 94,33%. Nilai FP dan FN berturut-turut adalah 8,9% dan 7,8%, yang berarti bahwa instrumen tersebut valid. Korelasi *Pearson Product Moment* (r_{xy}) yang diperoleh adalah 0,64 sedangkan nilai reliabilitas (r_{11}) adalah 0,885. Kedua nilai tersebut lebih besar daripada r_{tabel}, yang berarti bahwa instrumen yang dikembangkan tersebut valid dan reliabel.

Kata Kunci: Teori kinetik gas, *five-tier conception diagnostic test*, validitas, reliabilitas

INTRODUCTION

The physics subject contains many and various concepts. This causes many students to assume that Physics is a difficult subject and this leads to the misconception (Kurniawati & Ermawati, 2019). Many students in Science class in senior high school experient misconception on Thermodunamics concept (Jauhariyah et al., 2018; Suliyanah et al., 2018), one of them is in the Kinetic Theory of Gases. When the author conducted a teaching practice in 11-grade of Science class 3 of Senior High School 1 Gresik, the author demonstrated a simple experiment using two balloons each with different volume. Balloon A is the balloon with a large volume, while Balloon B is the one with a smaller volume. The two balloons were connected using a plastic straw. The middle of the plastic straw was clamped using a clip so the air on each balloon did not flow. The students were asked to determine how the final state of each balloon after the clamp was removed. Students assumed that the air in the Balloon A will flow into the Balloon B until the volume of the two balloons was equal. This presumption of students is called the initial understanding or preconception (Utari & Ermawati, 2018).

Meanwhile based on the Physics concept, the condition of the Balloon A and Balloon B is known as Boyle's Law. Boyle's Law states that at a constant temperature, the value of gas pressure is inverse to its volume (Tsokos, 2008, p. 176). Therefore, in the case of two balloons above, air will flow from Balloon B to Balloon A. This is because Balloon B has greater pressure than Balloon A. The difference between the initial understanding of student and the Physics concept is called misconception (Admoko et al., 2018; Kurniawati & Ermawati, 2019). Suprapto (2020) said that misconception not only occurs to students, but can also occur to anyone such as teacher, lecturers, and even professors.

Jauhariyah et al. (2018) also found student's misconceptions in the Kinetic Theory of Gases. The highest misconception that she reported occurred in the sub-concept of the Ideal Gas Law. Misconceptions experienced by students should not happen continuously, because it can inhibit students in understanding the next concepts (Ermawati et al., 2019). Diagnostic tests are the first step to identify misconceptions and even the level conception of students. Amin et al. (2016) and Anam et al. (2019) divided the level conceptions into scientific conception, partial understanding, misconception, and no understanding of concept. There are a variety of diagnostic tests that usually used to detect this level conceptions, including interviews (Griffiths & Preston, 1992), concept maps (Ingec, S.K., 2009), and multiple choice tests (Kirbulut & Geban, 2014). Multiple choice test is a more efficient way than interviews and concept map to identify potential misconceptions in class (Kirbulut & Geban, 2014). However, multiple choice test also has some disadvantages such as the answers chosen by students could be guessed (Anam et al., 2019; Bayuni et al., 2018). That is because students only choose the answers that are already available, not the answer that are truly understood and desired by students. To reduce these weaknesses, multiple choice tests developed to be multiple tier tests, ranging from one-tier, two-tier, three-tier, to four-tier (Kaltakci-Gurel et al., 2017). The multiple tier tests that currently used is four-tier test. The four-tier test consisting of questions and answers, the level of confidence in choosing an answer, the choice of reasons in choosing an answer, and the level of confidence in choosing a reason. The four-tier test can help teachers in identifying student's misconceptions deeper in terms of variations in choices given by students, both the choice of answers and the choice of reasons for answers and level confidence of each choice.

Although the four-tier test can provide more detailed identification results when compared with the identification results from the previous generation of multiple tier tests, the four-tier test was apparently considered by Anam et al. (2019 and Bayuni et al. (2018) to be unable to provide space for students to express their ideas about the concepts of given problems. Therefore, the possibility of an answer in the form of guesses written by the students still occurs. Thus, Bayuni et al. (2018) and Anam et al. (2019) added an extra tier in the form of an open-ended question. The answer to this question then used as a confirmation for the Assessor to justify whether the student's answer expresses student's understanding or no. Given the characteristics of each question from a different concept, the extra fifth tier question can also vary depending on needs. For example, according to Anam et al. (2019) for concepts that require student's understanding of confirmation in the form of a drawing, the extra fifth tier question should be a drawing question. Whereas if the confirmation required is the ability of students to express conclusions, the extra fifth tier question should be a conclusion question.

The five-tier test also requires interpretation to classify students based on the variations in the answers given by students. In the four-tier test, a student that provides an answer combination from the 1st to 4th tier questions as "*correct-sure-correct-sure*" is classified as scientific conception. In the five-tier question format, the level of student conception must still be determined by the students answer in the fifth tier. Thus, a five-tier diagnostic test can diagnose a deeper level of student

conception. Table 1 below summarizes the various answers given by student and student conception level based on the combination of student answers on a five-tier diagnostic test.

Table 1. Catagories of Combination Five-Tier Answers and Conception Level (Amin et al., 2016; Anam et al., 2019)

No	1st tier	2nd tier	3rd tier	4th tier	5th tier	Conception Level
1	Correct	Sure	Correct	Sure	SD/C PD/C UD/C MD/C ND/C	SC ASC LK UnC
2	Correct Correct Correct Correct Correct Correct Correct Wrong Wrong Wrong Wrong	Sure Not Sure Not Sure Sure Sure Not Sure Not Sure Sure Sure Not Sure Not Sure	Correct Correct Correct Wrong Wrong Wrong Wrong Correct Correct Correct Correct	Not Sure Sure Not Sure Not Sure Sure Sure Not Sure Sure Not Sure Sure Not Sure	PD/C or UD/C or MD/C	LK
3	Wrong Wrong Wrong	Sure Not sure Not sure	Wrong Wrong Wrong	Not sure Sure Not sure	PD/C or UD/C or MD/C	NU
4	Wrong	Sure	Wrong	Sure	PD/C or UD/C or MD/C	MSC
5	There is "tier" which are not answered or the answer is more than one available option					UnC

SC=scientific conception, ASC=almost scientific conception, LK=lack of knowledge, NU=no understanding on concept, MSC=misconception, UnC=uncode
SD/C=Scientific Drawing/Conclusion, PD/C=Partial Drawing/Conclusion, UD/C=Undifined Drawing/Conclusion, MD/C=Misconception Drawing/Conclusion, ND/C=No Drawing/Conclusion

The answers of each students in the fifth tier can vary, for example in the form of scientific, partial, undifined, misconception, or no drawing/conclusion as shown in Table 1 above. A description of the assessment to confirm the answer given by the student in the fifth tier is listed in Table 2.

Table 2. Category, Description and Score of Student's Drawing/Conclusion (Dikmenli, 2010; Köse, 2008)

No	Category	Description	Score (%)
1	**Scientific Drawing/ Conclusion**	Students provide correct answers with drawings/conclusions are in accordance with physics concept.	100
2	**Partial Drawing/ Conclusion**	Students provide drawings/conclusions are partly in accordance with physics concept.	99-70
3	**Undifined Drawing/ Conclusion**	Students provide answers that cannot be understood or the drawing/conclusion do not meet the physics concept.	69-40
4	**Misconception Drawing/ Conclusion**	Students provide wrong answers and the drawings/conclusions are diffenerent with physics concept.	39-1
5	**No Drawing/ Conclusion**	Students does not provide answers.	0

In the first combination (see Table 1 no. 1) where the student gives a correct-sure-correct-sure, and the answer in the fifth tier gets score 100 that belongs to the scientific drawing/conclusion, according to Anam et al. (2019) this student is called "understand the concept" or scientific conception. However, if the student drawing/conclusion gets score 99-70 (partial), then the conception level of student is almost scientific conception. Whereas if the student answer gets score 69-40 (undifined drawing/conclusion) or even gets score 39-1 (misconception drawing/conclusion), the student conception level is lack of knowledge. The last possibility is that if the student does not provide confirmation of an answer, the conception level is un-code.

Based on the description above, this article reports the work to develop instrument of five-tier conception diagnostic test on Kinetic Theory of Gases. The aim was to determine the validity (both internal and external validities) and reliability.

METHOD

This study was started by conducting a literature study and pre-study activities on January 8, 2020 at Senior High School 1 Pamekasan by providing three-tier diagnostic test questions with an open reason on the material Kinetic Theory of Gases to students. In addition, the author distributes student responses to the Physics questionnaire. The pre-study activity was carried out in order to explore data on potential misconceptions on the Kinetic Theory of Gases that might be experienced by sudents in the high school and explore information about student's interest in Physics, and find out what material students find difficult and easy.

After the stage above, the author compiling a five-tier conception diagnostic test on the Kinetic Theory of Gases based on literature studies and pre-study activities that have been carried out. The first stage is to identify any potential misconceptions, and developed the instrument design (open ended test), the initial trial was carried out to 27 numbers of student in Physics Dept. In order to collect some possible reasons in answering the third tier question, revise the initial instrument, the developed instrument given in Table 3, conducted Trial 1 to 35 students in 11-grade of Science class A of Senior High School 1 Pamekasan to get external validity and reliability value, analysis of external validity and reliability.

Table 3. One of Questions in The Developed Five-Tier Conception Diagnostic Test

Tier	Question and the Multiple-Tier Test
1st	*Questions and Answers*

Tier	Question and the Multiple-Tier Test
tier	Pernyataan berikut ini yang menunjukkan sifat dari ukuran partikel gas ideal adalah.... a. Partikel gas cukup besar dan jarak antar partikel berjauhan b. Partikel gas cukup besar dan jarak antar partikel berdekatan c. Partikel gas memiliki ukuran/volume yang sama dengan volume ruang yang ditempati d. Partikel gas sangat kecil dan jarak antar partikel berdekatan e. Partikel gas sangat kecil dan jarak antar partikel berjauhan The following statements which show the nature of the ideal gas particle size is .... a. Gas particles are quite large and the distances between particles are far apart b. Gas particles are quite large and the distances between particles are close c. Gas particles have the same size/volume as the volume of space occupied d. Gas particles very small and the distances between particles are close e. Gas particles very small and the distances between particles are far apart
2nd tier	*The Level of Confidence in Choosing an Answer* Apakah kamu yakin terhadap jawabanmu? □ Yakin □ Tidak Yakin Are you sure about your answer? □ Sure □ Not sure
3rd tier	*The Choice of Reasons in Choosing an Answer* Alasan pilihan jawaban a. Gaya antar partikel gas ideal menyebabkan ukuran partikel gas besar b. Gas ideal akan stabil apabila ukurannya cukup besar dan jaraknya berdekatan c. Gas ideal tidak memiliki gaya antar partikel d. Dalam suatu wadah, semakin besar ukuran partikel maka jarak antar partikel gas ideal semakin kecil e. Gas adalah zat yang menempati ruangnya secara penuh, sehingga partikel gas berukuran besar dan berdekatan f. Gas memiliki volume yang sama dengan volume ruang yang ditempati a. The forces between ideal gas particles causing the size of pasrticles are large b. The ideal gas will be stable if its size is large enough and the distance is close together c. Ideal gas has no inter-particle force d. In a container, if the size of particles are large then the distance between ideal gas particles are small e. Gas is a substance that occupies its full space, so that the gas particles are large and close together f. Gas has the same volumes as the volume of space occupied
4th tier	*The Level of Confidence in Choosing a Reason* Apakah kamu yakin terhadap jawabanmu? □ Yakin □ Tidak Yakin Are you sure about your answer? □ Sure □ Not sure
5th tier	*Drawing/Conclusion* Gambarkan partikel-partikel gas ideal dan gas riil masing-masing dalam suatu wadah dengan dinding tertutup di bawah ini! Draw the ideal gas and real gas particles respectively in a container with a closed wall below! Ideal Gases Real Gases

As explained above, the analysis of validity was carried out twice, internal validity and external validity. The internal validity was carried out by two Physics lecturers who reviewed the instrument in terms of content, construct and language. Indicators for assessing content validation include: (a) the compatibility of the item test with the Kinetic Theory of Gases; (b) the suitability of the item test with the question indicators; (c) the suitability of the item test with the order of the content; (d) limits, answers as well as reasons for answers clearly given. While the indicators of construct validation include: (a) the clarity of test instructions; (b) the compatibility between the question criteria and Bloom's taxonomy and basic competencies; (c) each question can identify the students conception; (d) the choice of reasons can reveal the cause of misconceptions originating from within the students; (e) deception on the choice of reasons is rational and homogeneous with answers on the first tier; (f) tables, graphs, figures and the like according to the problem presented. The indicators of the language validation include: (a) the use of Indonesian language that is good and right; (b) the various sentences do not lead to multiple interpretations; (c) the problem is stated clearly and communicatively.

The validation assessment was carried out in accordance with the assessment rubric by giving a score of one to four. Percentage of internal validity was calculated using the following equation.

$$P = \frac{S_R}{N.P_A.R}.100\% \quad (1)$$

Where P is the percentage of internal validity; S_R is the total of respondent's scores; N is maximum score in the questioner; P_A is the total items of the questioner; and R is the total of respondent. Table 4 shows the criteria of internal validity results.

Table 4. Percentage Score of The Internal Validity & The Critaria (Riduwan, 2013:18)

Percentage (%)	Criteria
0 – 20	Very low
21 – 40	Low
41 – 60	Valid enough
61 – 80	Valid
81 – 100	Very valid

Meanwhile, the external validity includes content and construct. Content validition of the developed instrument was determined by calculate the percentage of false positives (FP) and false negatives (FN) and have to less than 10% (Zahra & Suprapto, 2019). False positives happens when the combination answer of student in 1st tier until 5th tier is correct, confirm, incorrect, confirm, incorrect, respectively. Whereas it becomes false

negatives if the student's answers is incorrect, confirm, correct, confirm, and incorrect.

The construct validation of the developed instrument was calculated using the Pearson Product Moment Correlation Equation (r_{xy}).

$$r_{xy} = \frac{\sum r_{xy}}{\sqrt{(\sum x^2)(\sum y^2)}} \quad (2)$$

Where, r_{xy} is the correlation between x and y; x is the difference between the total score of the correct answer (1st and 3rd tier) in every question and the average of the correct answer score in all of questions; and y is the difference between the total of the sure answer score (2nd and 4th tier) in every question and the average of the sure answer score in all of questions. The instrument will be valid when the value of $r_{xy} > r_{theoritic}$ (Arikunto, 2016).

According to Sugiyono (2015), the reliability was calculated using the Cronbach Alpha Coefficient equation (r_{11}).

$$r_{11} = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum {\sigma^2}_i}{{\sigma^2}_t}\right) \quad (3)$$

Where,

$$\sigma_i^2 = \frac{\sum {X_i}^2 - \frac{(\sum X_i)^2}{N}}{N} \quad (4)$$

$$\sigma_t^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N} \quad (5)$$

Where r_{11} is reliability coefficient, k number of questions, $\sum {\sigma^2}_i$ is variant scores of each questions, ${\sigma^2}_t$ is total varian, N is number of students, X_i is students answer for each question, $\sum X$ is total of students answer for each question.

The choosen $r_{theoritic}$ value is 0.3338 with the samples are 35 students in 11-grade of Science class A of Senior High School 1 Pamekasan as samples and 5% significant level. Rohmanasari & Ermawati (2019) reported an instrument is reliable if the value of $r_{11} > r_{theoritic}$.

RESULTS AND DISCUSSION

The results of recapitulation of the validation assessment conducted by two Physics lecturers are shown in Table 5.

Table 5. Recapitulation of Percentage and Criteria of Internal Validity

Validity	Aspects	Validator		Percentage (%)	Criteria
		1	2		
Content	a	4	4	97.00	Very valid
	b	4	4		
	c	4	4		
	d	3	4		
Construct	a	3	4	94.00	Very valid
	b	3	3		
	c	4	4		
	d	4	4		
	e	4	4		
	f	4	4		
Language	a	3	3	92.00	Very valid
	b	4	4		
	c	4	4		
Average				94.33	Very valid

Based on the data in Table 5 and Table 4, the developed five-tier instrument test is very valid dan feasible to use. While the content validity results are given in Table 6.

Table 6. Total of Positives and Negatives Falses for Each Questions

Item Number	False Positives (FP)	False Negatives (FN)
1	2	2
2	11	1
3	2	8
4	2	1
5	1	2
6	3	2
7	5	2
8	2	3
9	2	2
10	3	4
11	6	7
12	1	2
13	7	3
14	2	2
15	1	1
16	2	4
17	3	2
18	1	1
Total (Σ)	56	49
Σ students	35	
Σ question x Σ students	630	
Percentage	8.9 %	7.8 %

Based on the data in Table 6 above, it is known that false positives (FP) and false negatives (FN) obtained in this research is < 10%. Therefore, the external validity of content for the developed instrument is valid.

The results of the external validity of construct for each questions are given in Table 7.

Table 7. Calculation of The Value of r_{xy} for Each Questions

Item Number	r_{xy}	$r_{tehoritic}$	Category
1	0.7371	0.3338	Valid
2	0.3673		Valid
3	0.4179		Valid
4	0.2293		Invalid
5	0.1922		Invalid
6	0.6147		Valid
7	0.7625		Valid
8	0.4315		Valid
9	0.5763		Valid
10	0.5885		Valid
11	0.7398		Valid

Item Number	r_{xy}	$r_{tehoritic}$	Category
12	0.7354		Valid
13	0.6591		Valid
14	0.5594		Valid
15	0.5339		Valid
16	0.4097		Valid
17	0.6594		Valid
18	0.7728		Valid

Based on Table 7, it can be seen that out of eighteen diagnostic test instrument questions, there are two questions that are invalid because the $r_{xy} < r_{theoritic}$.

The external validity of construct of the whole instrument also used Eq (2). The obtained value of r_{xy} is 0.64 that the developed instrument is valid because $r_{xy} > r_{theoritic}$.

Table 8 gives the data of variant to calculate the reliability index of the developed instrument used Equation (3).

Table 8. Data of Total Variant and The Variant of Each Questions

Item Number	The variant scores of each questions (σ_i^2)	The total variant scores (σ_t^2)	The total of questions (k)
1	0.2482		
2	0.2041		
3	0.2155		
4	0.2253		
5	0.2335		
6	0.2498		
7	0.2482		
8	0.2155		
9	0.1224	20.7	18
10	0.191		
11	0.2335		
12	0.2449		
13	0.142		
14	0.2482		
15	0.2482		
16	0.142		
17	0.2041		
18	0.24		

Based on data in Table 8 above, we can calculate the reliability index (r_{11}) used Eq (3) and obtained r_{11}= 0.8849. It can be known that the developed instrument is reliable because $r_{11} > r_{theoritic}$.

Based on the analysis above, the results of this study in accordance with the criteria of a valid instrument by Arikunto (2016), Riduwan (2013), Zahra & Suprapto (2019), and it in accordance with the criteria of a reliable instrument by Sugiyono (2015). It means that the developed five-tier conception diagnostic test instrument was decalered valid and reliable.

CONCLUSION

Based on the results above, it can be concluded that the developed instrument is valid, both internal and external validation. The reliability value (r_{11}) is also greater than the $r_{theoritic}$ value, so that the developed instrument is reliable. Thus the developed five-tier conception diagnostic test instrument is feasible to be used in identifying student's level conceptions.

REFERENCES

Admoko, S., Sunarti, T., Jauhariyah, M. N. R., Suliyanah, S., & Suprapto, N. (2018). Analysis of College Students' Misconception on Geometrical Optics. *Atlantis Highlights in Engineering (AHE)*, 1(Icst), 896–903. https://doi.org/10.2991/icst-18.2018.181.

Amin, N., Wiendartun, W., & Samsudin, A. (2016). Analisis Intrumen Tes Diagnostik Dynamic-Fluid Conceptual Change Inventory (DFCCI) Bentuk Four-Tier Test pada Beberapa SMA di Bandung Raya. *Prosiding SNIPS 2016*, March 2017, 570–574.

Anam, R. S., Widodo, A., Sopandi, W., & Wu, H. K. (2019). Developing a five-tier diagnostic test to identify students' misconceptions in science: an example of the heat transfer concepts. *Elementary Education Online*, 18(3), 1014–1029. https://doi.org/10.17051/ilkonline.2019.609690.

Arikunto, S. (2016). *Dasar-Dasar Evaluasi Pendidikan* (Edisi 2). Jakarta: PT. Bumi Aksara.

Bayuni, T. C., Sopandi, W., & Sujana, A. (2018). Identification misconception of primary school teacher education students in changes of matters using a five-tier diagnostic test Identification misconception of primary school teacher education students in changes of matters using a five-tier diagnostic. *Conf. Series: Journal of Physics*, 1013.

Dikmenli, M. (2010). Misconceptions of cell division held by student teachers in biology: A drawing analysis. *Scientific Research and Essay*, 5(2), 235-247.

Ermawati, F. U., Anggrayni, S., & Isfara, L. (2019). Misconception profile of students in senior high school iv Sidoarjo East Java in work and energy concepts and the causes evaluated using Four-Tier Diagnostic Test. *Journal of Physics: Conference Series*, 1387(1). https://doi.org/10.1088/1742-6596/1387/1/012062.

Griffiths, A. K., & Preston, K. R. (1992). Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching*, 29(6), 611-628. doi: 10.1002/tea.3660290609.

Ingec, S. K. (2009). Analysing concept maps as an assessment tool in teaching physics and comparison

with the achievement tests. *International Journal of Science Education*, 31(14), 1897-1915. doi: 10.1080/09500690802275820.

Jauhariyah, M. N. R., Suprapto, N., Suliyanah, Admoko, S., Setyarsih, W., Harizah, Z., & Zulfa, I. (2018). The Students' misconceptions profile on chapter gas kinetic theory. *Journal of Physics: Conference Series*, 997(1). https://doi.org/10.1088/1742-6596/997/1/012031.

Kaltakci-Gurel, D., Erylmaz, A., & McDermott, L. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science & Technological Education*, 35(4), 1-23.

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. doi: 10.1007/s11165-007-9048-7

Kirbulut, Z. D., & Geban, O. (2014). Using Three-Tier Diagnostic Test to Asse ss Students ' Misconceptions of States of Matter. *Eurasia Journal of Mathematics, Science & Technology Education*, 10(5), 509–521. https://doi.org/10.12973/eurasia.2014.1128a.

Köse, S. (2008). Diagnosing Student Misconceptions: Using Drawings as a Research Method. *World Applied Science Journal*, 3(2), 283-293.

Kurniawati, D. M., & Ermawati, F. U. (2019). The Validity of Four-Tier's Misconception Diagnostic Test for Dynamic Fluid Concepts. *Inovasi Pendidikan Fisika*, 08(02), 668–671.

Riduwan, & Akdon. (2013). *Rumus dan Data dalam Analisis Statistika*. Bandung: Alfabeta.

Rohmanasari, F., & Ermawati, F. U. (2019). Validitas dan Reliabilitas Instrumen Tes Diagnostik Miskonsepsi Berformat Four-Tier Pada Materi Alat Optik. *Inovasi Pendidikan Fisika*, 08(02), 556–559.

Sugiyono. (2015). *Metode Penulisan Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

Suliyanah, Putri, H. N. P. A., & Rohmawati, L. (2018). Identification student's misconception of heat and temperature using three-tier diagnostic test. *Journal of Physics: Conference Series*, 997(1). https://doi.org/10.1088/1742-6596/997/1/012035.

Suprapto, N. (2020). Do We Experience Misconceptions ?: An Ontological Review of Misconceptions in Science. *Studies in Philosophy of Science and Education*, 1(2), 50–55.

Tsokos, K. A. (2008). *Physics for the IB Diploma*. United Kingdom: Cambridge University Pers.

Utari, J. I., & Ermawati, F. U. (2018). Pengembangan Instrumen Tes Diagnostik Miskonsepsi Berformat Four Tier untuk Materi Suhu, Kalor, dan Perpindahannya. *Inovasi Pendidikan Fisika*, 07(03), 434–439.

Zahra, Y., & Suprapto, N. (2019). Analisis Kualitas Instrumen Four-Tier Diagnostic Test untuk Mengidentifikasi Profil Konsepsi Siswa pada Materi Teori Kinetik Gas. *Inovasi Pendidikan Fisika*, 08(03), 830–834.