JIPK 19 (1) (2025)
Jurnal Inovasi Pendidikan Kimia
https://journal.unnes.ac.id/journals/JIPK

Development Of A Five-Level Diagnostic Instrument On The Rate Of Reaction Material: To Identify The Causes Of Students' Misconceptions And Representation

Deni Ainur Rokhim✉, Hayuni Retno Widarti, Munzil

Universitas Negeri Malang, Faculty of Science and Mathematics, East Java, Indonesia

Info Artikel

Diterima: Nov 2024
Disetujui: Des 2024
Dipublikasikan: Jan 2025

Keywords:
Development, Diagnostic Instrument, Five-Tier, Misconceptions

Abstrak

Penelitian ini bertujuan untuk mengembangkan dan menguji validitas instrumen diagnostik lima tingkat pada materi laju reaksi. Instrumen ini mencakup lima level: pilihan ganda, alasan jawaban, tingkat kepercayaan, gambar representasi, dan sumber informasi. Sebanyak 100 siswa SMA Negeri di Sidoarjo dilibatkan. Hasil menunjukkan bahwa instrumen memiliki reliabilitas tinggi (α = 0,81–0,90) dan validasi isi mencapai ≥90%. Visualisasi siswa diklasifikasikan dalam enam kategori untuk mengidentifikasi pemahaman konseptual dan miskonsepsi. Instrumen ini memungkinkan analisis profil representasi ganda dan penggalian informasi kognitif melalui gambar, yang mendukung pemetaan miskonsepsi secara mendalam. Temuan menunjukkan bahwa penambahan elemen gambar memperkaya diagnostik konseptual siswa serta mendorong berpikir tingkat tinggi dan penguasaan sains berbasis representasi. Instrumen lima tingkat ini terbukti efektif dalam mengidentifikasi miskonsepsi, tingkat keyakinan, dan sumber pemahaman siswa.

Abstract

This study aims to develop and validate a five-tier diagnostic instrument on reaction rate concepts. The instrument consists of five levels: multiple choice, reasoning, confidence level, visual representation, and information source. A total of 100 high school students from Sidoarjo participated. Results indicate high reliability (α = 0.81–0.90) and content validity above 90%. Students' visualizations were classified into six categories to identify conceptual understanding and misconceptions. The instrument enables detailed analysis of multiple representations and cognitive information through student-generated images, supporting in-depth misconception mapping. Findings show that incorporating visual elements enhances conceptual diagnostics, encourages higher-order thinking, and strengthens science learning through representational approaches. The five-tier instrument proves effective in identifying misconceptions, confidence levels, and the sources of students' understanding.

✉ Alamat korespondensi:
E-mail: ainurrokhim@gmail.com

p-ISSN 1979-0503
e-ISSN 2503-1244

INTRODUCTION

This level of understanding can be achieved by taking specific tests or assessments. Teachers usually give different tests to see the development of student learning outcomes. The current trial is not only required as a level of understanding but requires detailed information about the scientific understanding of students using a diagnostic test (Gurel et al., 2015; Kaltakçi, 2012; Rokhim, Widarti, et al., 2023; Rokhim et al., 2024). Diagnostic tests are a tool for identifying student learning difficulties. Test results are used as material to improve learning that is more in line with students' abilities, one of which is learning difficulties (Fortuna et al., 2013; Lemma & Ethiopia, 2012). Diagnostic tests require the right tools to study. The diagnostic instrument is an instrument that can be used to analyze the strengths and weaknesses of students' conceptual understanding of the material (Putra et al., 2020; Seçken, 2010). Solve if students have conceptual weaknesses so that appropriate solutions can be given. Diagnostic tests have been developed using various means, such as multiple-choice, image, and multi-level tests, such as levels two, three, and four (Kaltakçi, 2012; Rokhim et al., 2022; Rokhim, Rahayu, et al., 2023; Sözen & Bolat, 2011).

Diagnostic instruments that are starting to develop are multi-tiered or multi-tiered tests. This type of instrument was developed from a multiple-choice test type. The advantage of multiple-choice tests is that they have strong validity, easy to assess, implement, and analyze, have broad topic coverage, and have the flexibility to measure various levels of learning ability (Kaltakçi, 2012), but besides these advantages, there are some disadvantages such as correct answers can be obtained by guessing, and the diagnostic results obtained are not can identify students' conceptual understanding in detail. This instrument makes it possible to identify whether students have misconceptions but cannot distinguish errors due to the lack of knowledge of these errors and cannot identify whether students' answers are correct based on understanding or guessing (Nofiana et al., 2016; Usta & Ayas, 2010).

The next diagnostic instrument development is a three-tier test containing the usual multiple-choice tests, multiple-choice reasons, choices of confidence levels or answers, and reasons. The test can confidently detect misconceptions from the presence or absence of knowledge (Damsi & Suyanto, 2023; Qodriyah et al., 2020; Soeharto & Csapó, 2022; Widarti et al., 2022). However, this instrument needs to identify the level of understanding of the concept because the level of confidence refers to both levels at once (Cengiz, 2009; Nofiana et al., 2016). This three-level diagnostic instrument was then developed into four levels: ordinary multiple-choice, reasons, and the confidence level for each answer and reason (Milenković et al., 2016).

The five-level diagnostic instrument is a further development of the four-level diagnostic instrument. Four-level diagnostic instrument Four-level diagnostic instrument consists of four levels, namely the first level, which contains multiple choice answers and distractors related to certain misconceptions; the second level includes the level of confidence from the solutions chosen at the first level; The third level consists of a choice of reasons for the selected answers at the first level (Yonata, 2021). The fourth level involves a level of trust in the third level reasons.

In the five-level diagnostic instrument, the test consists of five levels, namely: 1) Playing multiple choice questions; 2) Reasons or thoughts in the answers to the main questions; 3) The level of confidence from the previous solution; 4) Pictures/representation of responses; and 5) A questionnaire about sources of information was used as the basis for student answers. The five-level diagnostic instrument can analyze the multiple representation profile of students. Multiple representation profiles are obtained from the answers in the fifth tier in the form of an overview of multiple representations.

Drawing two interactions, namely representations and external mental models for students. When drawing, the brain uses spatial information to write down other types of information, such as verbal and visual information, to increase the brain's capacity to remember and learn verbally and visually(Achuthan & Murali, 2015; Bhaw et al., 2023). In processing verbal and visual information, the brain will select and organize verbal and visual information to be used and integrate elements into mental models (Alwan, 2011). Besides requiring mental processes when drawing, model-making also uses motor coordination to manipulate the image media into the desired image (Damsi & Suyanto, 2023; Quillin & Thomas, 2015). Pictures made by students in chemistry lessons can provide very useful insights into common understandings or misconceptions and alternative conceptions (Mawaddah & Haryani, 2021; Rusmini et al., 2021). Misconceptions are the views of someone who mismanages a concept, is not following what is justified according to experts, is difficult to change, and tends to persist. Students usually already have prior knowledge, which often relates their knowledge to current knowledge, which needs to be more accurate, thus creating wrong and harmonized concepts compared to experts' concepts (Mutlu & Sesen, 2015; Rusmini et al., 2021; Widarti et al., 2017, 2020).

This study develops a four-level diagnostic test instrument for response-level material by adding one-level questions called a five-level diagnostic test instrument. With this test tool, you can analyze several representative student profiles and data sources used to answer the questions and obtain students' double performance profiles from the results of student responses in the form of images related to multiple

appearances. Drawing tools should be added to diagnostic tools because drawing is an effective way of thinking and communicating in all disciplines. Additionally, drawing is a scientific processing skill for hypothesizing, designing experiments, visualizing and interpreting data, and presenting results.

METHOD

This study involved 100 students from three schools at SMA Negeri in Sidoarjo. This study used the survey test method to determine students' understanding of the reaction rate material. The instrument is a five-tier diagnostic instrument consisting of five levels of answers. The first tier is a question with a multiple-answer choice type; the second tier is in the form of answer choices for reasons in determining the answers to the first tier; the third tier is the level of confidence of students in determining answers in the first tier and second tier, the fourth tier contains instructions for students to describe the understanding of multiple representations and the fifth tier which contains questions that ask students to name sources that are used as references in answering questions.

This instrument has been modified for the topic of reaction rate in class XI high school students and added multiple representations and questions about sources of information which serve as the basis for students in answering questions on the test instrument. Based on the results of the reliability test of the items in which the reliability of the five-tier diagnostic instrument was tested using Cronbach's Alpha technique, the results were good. According to the reliability criteria, the items are included in the high category because they get results in the range of 0.81 - 0.90, which means the item reliability test is high. Because this instrument involved a drawing test, participants' drawings (visualization) were classified into six categories.

Table 1. Student Image Criteria

Category	Description
Scientific Drawing (SD)	Respondents provide a comprehensive visualization by the scientific description.
Partial Drawing (PD)	Responses provide a visualization that closely matches the scientific picture with some drawbacks.
Misconception Drawing (MD)	Respondents provided visualizations that needed to be more accurate or different from scientific descriptions, but they drew their visualizations at the sub-microscopic level.
Undefined Drawing (UD)	Respondents provided incomprehensible visualizations, even though the visualizations provided were at the sub-microscopic level.
Non-Microscopic Drawing (NMD)	Respondents provide visualization, but not at the microscopic level.
No Drawing (ND)	Respondents did not provide visualization at all or only wrote answers.

Diagnostic tests equipped with pictures can help researchers or teachers to get more information regarding what students have learned through the learning process and find out the level of conceptual understanding of students (Dikmenli, 2010; Soeharto & Csapó, 2021). Based on research from Anderson (2014) said that by drawing, students' thoughts can move to think at a higher level and can develop at the conceptual level of science, the ability to observe, and the ability to think (Anderson et al., 2014).

DISCUSSION

This study's five-tier diagnostic test instrument consisted of multiple-choice questions along with reasons and confidence levels. Open questions in the rationale section are useful for identifying students' misconceptions about the reaction rate material. In general, the five-tier diagnostic test consists of the following:

The five-tier diagnostic instrument consists of five levels, namely the first level, which contains multiple choice answers with a distractor that refers to certain misconceptions; the second level contains a choice of reasons for the answers selected at the first level; the third level is the confidence level of the answers selected at the previous level, the fourth level contains pictures/representations of the answers, and the fifth level contains the sources of student answers used.

The test results using a five-tier diagnostic instrument can be used to analyze the multiple representation profiles of students. Multiple representation profiles are obtained from the answers in the fourth tier in the form of an overview related to multiple representations.

The validator previously validated this instrument. After validating the contents of the items obtained from the validator's assessment, 3 experts/validators consisted of 2 lecturers majoring in chemistry at Malang State University and 1 teacher from SMA Negeri 3 Sidoarjo, as can be seen in the following figure 1.

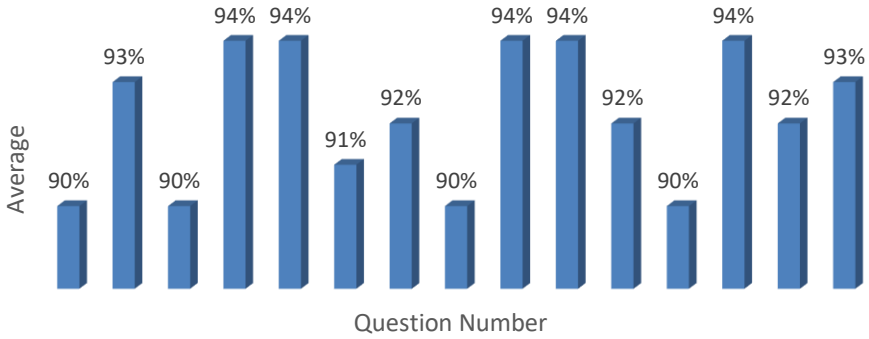
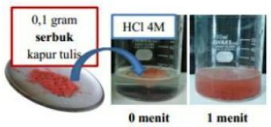
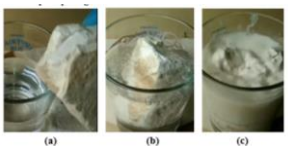


Figure 1. The average result of the validation percentage of the item content

Based on the picture above shows that the results of validating the contents of the item items carried out by the validator, as many as 15 questions can be said to be valid. Borich (1994) argues that R ≥ 75% can be classified as a good percentage of agreement by the validator. It is because the lowest average percentage is 90% and can still be said to be valid.

In addition, there are several suggestions and revisions from the validator. The validator suggests that the image in number 1 is better replaced with a clearer image. We received and corrected the suggestions as shown in Table 2.

Table 1 Instrument Revision Result

No	Before	After
1	Deni experimented with dissolving chalk. As shown in the picture below, he put 0.1 gram of red chalk dissolved in 20 mL of 4 M HCl solution. 	Deni experimented with dissolving chalk. As shown in the picture below, he put 0.1 gram of red chalk dissolved in 20 mL of 4 M HCl solution. 
We have made suggestions on these numbers to correct or revise these questions. The validator intends that students who see the picture when working on the problem are not confused, and there are no misconceptions just because they see an unclear picture		
3	**First Tier** The formation of colorless ammonia gas can be done by reacting nitrogen gas and hydrogen gas on a laboratory scale. The reaction rate at any time in the construction of ammonia gas can be expressed as	**First Tier** The formation of colorless ammonia gas can be done by reacting nitrogen gas and hydrogen gas on a laboratory scale. The reaction rate at any time in the construction of amonia gas can be expressed as
The validator suggests that the correct way to write ammonia gas is ammonia gas in Indonesian language		
6	6. FIRST TIER Amir, in the afternoon, makes cold and hot green tea drinks while enjoying his family's	6. FIRST TIER Amir and Mira conducted an experiment to produce hydrogen gas. Amir reacted 1 gram of zinc powder with 1 M hydrochloric acid

	company. Amir put water and warm water into different glasses, then put the tea leaves and sugar into the two glasses. **If the sugar mass in the two green tea drinks is made the same, then the sugar's dissolution rate in the warm green tea drink will….**	solution in a test tube at room temperature. It turned out that the zinc powder had completely reacted after 2 minutes. After that, Mira did the same experiment with the same ingredients and amount. **However, Mira heated the HCl in her test tube when reacting the zinc powder with a 1 M hydrochloric acid solution. Based on this, analyze the experiment conducted by Mira; how is the reaction rate of zinc powder with a hydrochloric acid solution in Mira's experiment?**
Improve the diction of questions to make them clearer and more detailed with question number 6		
14	FIRST TIER **Data from the reaction experiment:** X + 2Y → Z at a certain temperature is as follows. Percobaan | [X] M | [Y] M | Laju Pembentukan Z (M/s) 1 | 0,1 | 0,1 | 0,16 2 | 0,1 | 0,2 | 0,32 3 | 0,2 | 0,2 | 0,64 4 | 0,3 | 0,3 | P The correct statement regarding this reaction is... a. The P value is 0.44 M/s b. The order for Y is 2 c. The total reaction order is 1 d. This reaction has a reaction rate constant of 16 M-1s-1 e. I do not know	FIRST TIER **Data from the reaction experiment:** X + 2Y → Z at a certain temperature is as follows. Percobaan | [X] M | [Y] M | Laju Pembentukan Z (M/s) 1 | 0,1 | 0,1 | 0,16 2 | 0,1 | 0,2 | 0,32 3 | 0,2 | 0,2 | 0,64 4 | 0,3 | 0,3 | P The correct statement regarding this reaction is... a. The value of the reaction rate in the 4th experiment was 0.44 M/s b. The order for Y is 2 c. The total reaction order is 1 d. This reaction has a reaction rate constant of 16 M-1s-1 e. I do not know
Improvements were made to the diction in the answer choice options. Previously we only used the parable P, but we have revised it to Reaction Rate Value.		
15	FIRST TIER The decomposition reaction $2KClO_3(s) \rightarrow 2KCl(s) + 3O_2(g)$ was followed by measuring the volume of O_2 produced. If the reaction is carried out at two different temperatures, T_1 and T_2, with $T_2 < T_1$, then the volume change of O_2 gas in the experiment with T_1 is more formed, Is this statement true? a. Right b. Wrong c. I do not know	FIRST TIER The decomposition reaction $2KClO_3(s) \rightarrow 2KCl(s) + 3O_2(g)$ was followed by measuring the volume of O_2 produced. If the reaction is carried out at two different temperatures, T_1 and T_2, with $T_2 < T_1$, then the volume change of O_2 gas in the experiment with T_1 is more formed, Is this statement true? a. Right b. Wrong
Improvement on answer option C, where there is an option "I do not know". Based on the validator's suggestion, it should be removed so that it does not trigger students to choose that answer		

Table 2 shows the results of the recapitulation of empirical validation obtained from a total of 100 students at SMA Negeri 3 Sidoarjo. These results recapitulate the 3 Tiers: Tier A, Tier R, and Multiple representatives.

Table 2. Item Validity

Question Number	Tier A	Tier R	Representative
1	0,000	0,000	0,769
2	0,003	0,011	0,234
3	0,001	0,002	0,000
4	0,000	0,000	0,000
5	0,007	0,000	0,000
6	0,000	0,033	0,000
7	0,008	0,014	0,000
8	0,013	0,641	0,000
9	0,001	0,000	0,000
10	0,001	0,000	0,000
11	0,000	0,641	0,000
12	0,000	0,002	0,000
13	0,044	0,000	0,000
14	0,001	0,079	0,000
15	0,024	0,899	0,000

This validity test uses Product Moment Pearson Correlations (Bivariate Pearson) analysis. According to Kimberlin & Winterstein (2008), validity refers to whether the information obtained from a test represents the proper understanding of the examinee. To test the validity of the items, you can use the value of the Pearson correlation index (r count)(Kimberlin & Winterstein, 2008). Based on the table above, most of the things are valid. The validity of the items is closely related to the index of discriminating power or discriminating power because if the items can trust to measure students' misconceptions, then the thing can distinguish high and low-achieving students (Jusniar et al., 2020).

Another possible reason why question number 15 is invalid is that even if it is wrong, the points are still positive, so it can still be used by checking the language of question number 15 because it is likely due to the language of the questions. Students need help understanding.

Relevant research on empirical validation tests written by Putri & Ernawati (2021) shows valid results because Rxy, compared to Rtable, produces Rxy > Rtable at a significance of 5%. By comparing the rxy and rtable values, it is found that the sixteen items in the Final Draft are valid, considering that the rxy > rtable values (Putri & Ermawati, 2021).

The reason for an invalid question is usually related to item differences. Item validity is closely related to the discriminating index or power because if an item can distinguish high and low-achieving students, the thing is reliable in measuring misconceptions (Jusniar et al., 2020).

Another possible cause of the invalidity of these questions is students who write down answers by guessing or not using the knowledge they get from school. Even though it is not valid, the item remains positive, so it can still be used by revising the language of the questions because it is likely caused by the language of the questions, which is difficult for students to understand.

Table 3. Reliability Test

N	Cronbach Alpha		
	Tier A	0,000	Tier A
15	0,707	15	0,707

The reliability test refers to evaluating items—calculations to determine the instrument's reliability using Cronbach's alpha with a significance value of 5%. Tier A has a Cronbach alpha coefficient of 0.707; Tier R is 0.664; in double representation is 0.957. The reliability results show that the five-level diagnostic device has high criteria, or it can say that the device is very reliable. Items are the test for reliability based on

the above information; According to Creswell (2012), the results of the test instrument have internal consistency or high regularity (Creswell, 2012). The statement is consistent with Utari et al. (2018) research on developing a four-level misconception diagnostic test instrument with a reliability test score of 1.067 is in the very reliable category (Utari & Ermawati, 2018).

This five-tier test on reaction rate material was developed to diagnose students' misconceptions in more detail by connecting multiple-choice questions to the main question and drawing sub-microscopically and symbolically. Through this test, researchers can also see the extent to which students understand the material being tested, which parts of the material they still feel unable to understand and become more aware of students' thoughts when answering questions. For this test to become an instrument for diagnosing misconceptions about reaction rates, it still needs to be tested further in future studies. Based on the results above, the ts diagnostic instrument identifies misconceptions about reaction rates in all educational units. This research has a Tier A reliable value with a Cronbach's alpha coefficient of 0.707; Tier R is 0.664, and Multiple representatives are 0.957.

CONCLUSION

This five-tier test on reaction rate material was developed to diagnose students' misconceptions in more detail by connecting multiple-choice questions to the main question and drawing sub-microscopically and symbolically. Through this test, researchers can also see the extent to which students understand the material being tested, which parts of the material they still feel unable to understand and become more aware of students' thoughts when answering questions. For this test to become an instrument for diagnosing misconceptions about reaction rates, it still needs to be tested further in future studies. Based on the results above, the ts diagnostic instrument identifies misconceptions about reaction rates in all educational units. This research has a Tier A reliable value with a Cronbach's alpha coefficient of 0.707; Tier R is 0.664, and Multiple representatives are 0.957.

ACKNOWLEDGEMENTS

The author would like to thank all parties who contributed to the process of making this article. The author hopes that this research can be useful for everyone. This work was supported by State University of Malang, East Java, Indonesia through the Office of Research and Technology.

REFERENCES

Achuthan, K., & Murali, S. S. (2015). A comparative study of educational laboratories from cost & learning effectiveness perspective. *Advances in Intelligent Systems and Computing*, *349*(January 2015), 143–153. https://doi.org/10.1007/978-3-319-18473-9_15

Alwan, A. A. (2011). Misconception of heat and temperature among physics students. *Procedia - Social and Behavioral Sciences*, *12*, 600–614. https://doi.org/10.1016/j.sbspro.2011.02.074

Anderson, J. L., Ellis, J. P., & Jones, A. M. (2014). Understanding early elementary children's conceptual knowledge of plant structure and function through drawings. *CBE Life Sciences Education*, *13*(3), 375–386. https://doi.org/10.1187/cbe.13-12-0230

Bhaw, N., Kriek, J., & Lemmer, M. (2023). Insights from coherence in students' scientific reasoning skills. *Heliyon*, *9*(7). https://doi.org/10.1016/j.heliyon.2023.e17349

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.

Creswell, J. W. (2012). *Planning, Conducting, and Evaluating Quantitative and Qualitative Research*.

Damsi, M., & Suyanto, S. (2023). *Systematic Literature Review : Multiple-tier Diagnostic Instruments in Measuring Student Chemistry Misconceptions*. *9*(5), 8–21. https://doi.org/10.29303/jppipa.v9i5.2600

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

Fortuna, D., Chandra, E., & Gloria, R. Y. (2013). Pengembangan Tes Diagnostik Untuk Mengukur Miskonsepsi Siswa Pada Pokok Bahasan Sistem Regulasi Manusia Untuk Siswa SMA Kelas XI Semester II. *Jurnal Scientiae Eduactia*, *2*(2), 1–16.

Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education*, *11*(5), 989–1008. https://doi.org/10.12973/eurasia.2015.1369a

Jusniar, J., Effendy, E., Budiasih, E., & Sutrisno, S. (2020). Developing a three-tier diagnostic instrument on chemical equilibrium (TT-DICE). *Educación Química*, *31*(3), 84–102. https://doi.org/http://dx.doi.org/10.22201/fq.18708404e.2020.3.72133

Kaltakçi, D. (2012). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. Research in Science and Technological Education. *עלון המועצ*, *66*(September), 37–39.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, *65*(23), 2276–2284. https://doi.org/10.2146/ajhp070364

Lemma, A., & Ethiopia. (2012). Diagnosing the Diagnostics : Misconceptions of Twelfth Grade Students on Selected Chemistry Concepts in Two Preparatory Schools In Eastern Ethiopia. *ASEAN Journal of Community Engagement*, *2*(2), 16–31.

Mawaddah, I., & Haryani, S. (2021). Pengaruh Pembelajaran Berbasis MultirepresentasiTerhadap Keterampilan Metakognisi pada Materi LajuReaksi. *Chemistry in Education*, *10*(1), 22–29.

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

Mutlu, A., & Sesen, B. A. (2015). Development of a Two-tier Diagnostic Test to Assess Undergraduates' Understanding of Some Chemistry Concepts. *Procedia - Social and Behavioral Sciences*, *174*, 629–635. https://doi.org/10.1016/j.sbspro.2015.01.593

Nofiana, M., Studi, P., Sains, P., Pascasarjana, P., Sebelas, U., Surakarta, M., Studi, P., Sains, P., Pascasarjana, P., Sebelas, U., Surakarta, M., Pascasarjana, P., Sebelas, U., & Surakarta, M. (2016). *Pengembangan Instrumen Evaluasi Two – Tier Multiple Choice Question Untuk Mengukur Keterampilan Berpikir*. *2001*.

Putra, A. S. U., Hamidah, I., & Nahadi. (2020). The development of five-tier diagnostic test to identify misconceptions and causes of students' misconceptions in waves and optics materials. *Journal of Physics: Conference Series*, *1521*(2). https://doi.org/10.1088/1742-6596/1521/2/022020

Putri, W. K., & Ermawati, F. U. (2021). Pengembangan, Uji Validitas dan Reliabilitas Tes Diagnostik Five-Tier untuk Materi Getaran Harmonis Sederhana beserta Hasil Uji Coba. *PENDIPA Journal of Science Education*, *5*(1), 92–101. https://doi.org/10.33369/pendipa.5.1.92-101

Qodriyah, N. R. L., Rokhim, D. A., Widarti, H. R., & Habiddin, H. (2020). IDENTIFIKASI MISKONSEPSI SISWA KELAS XI SMA NEGERI 4 MALANG PADA MATERI HIDROKARBON MENGGUNAKAN INSTRUMEN DIAGNOSTIK THREE TIER. *Jurnal Inovasi Pendidikan Kimia*, *14*(2), 2642–2651.

Quillin, K., & Thomas, S. (2015). Drawing-to-learn: A framework for using drawings to promote model-based reasoning in biology. *CBE Life Sciences Education*, *14*(1), 1–16. https://doi.org/10.1187/cbe.14-08-0128

Rokhim, D. A., Atikah, A., Vitarisma, I. Y., Rahayu, S., & Muntholib, M. (2022). Assessment of high school students' ability to solve structured problems with ideal model on acid-base. *Orbital: The Electronic Journal of Chemistry*, 276–284.

Rokhim, D. A., Rahayu, S., & Dasna, I. W. (2023). Analisis miskonsepsi kimia dan instrumen diagnosisnya: literatur review. *Jurnal Inovasi Pendidikan Kimia*, *17*(1), 17–28.

Rokhim, D. A., Widarti, H. R., & Sutrisno, S. (2023). Five-Tier Instrument to Identify Students' Misconceptions and Representation: A Systematic Literature Review. *Orbital: The Electronic Journal of Chemistry*, 202–207.

Rokhim, D. A., Widarti, H. R., & Sutrisno, S. (2024). Profile of Needs Analysis of Five-Tier Diagnostic Instrument Development for High School Chemistry Courses. *Pegem Journal of Education and Instruction*, *14*(2), 140–145.

Rusmini, Suyono, Jatmiko, B., & Yonata, B. (2021). The Diagnosis of Misconception on the Concept of Acid-Base Theory in Prospective Teacher Students Used a Three-Tier Test. *Journal of Physics: Conference Series*, *1899*(1). https://doi.org/10.1088/1742-6596/1899/1/012061

Seçken, N. (2010). Identifying student's misconceptions about SALT. *Procedia - Social and Behavioral Sciences*, *2*(2), 234–245. https://doi.org/10.1016/j.sbspro.2010.03.004

Soeharto, S., & Csapó, B. (2021). Evaluating item difficulty patterns for assessing student misconceptions in science across physics, chemistry, and biology concepts. *Heliyon*, *7*(11). https://doi.org/10.1016/j.heliyon.2021.e08352

Soeharto, S., & Csapó, B. (2022). Exploring Indonesian student misconceptions in science concepts. *Heliyon*, *8*(9). https://doi.org/10.1016/j.heliyon.2022.e10720

Sözen, M., & Bolat, M. (2011). Determining the misconceptions of primary school students related to sound transmission through drawing. *Procedia - Social and Behavioral Sciences*, *15*, 1060–1066. https://doi.org/10.1016/j.sbspro.2011.03.239

Usta, N. D., & Ayas, A. (2010). Common misconceptions in nuclear chemistry unit. *Procedia - Social and Behavioral Sciences*, *2*(2), 1432–1436. https://doi.org/10.1016/j.sbspro.2010.03.214

Utari, J. I., & Ermawati, F. U. (2018). *Inovasi Pendidikan Fisika ISSN : 2302-4496 PENGEMBANGAN INSTRUMEN TES DIAGNOSTIK MISKONSEPSI BERFORMAT Inovasi Pendidikan Fisika ISSN : 2302-4496*. *07*(03), 434–439.

Widarti, H. R., Permanasari, A., & Mulyani, S. (2017). Undergraduate Students' Misconception on Acid-Base and Argentometric Titrations: a Challenge To Implement Multiple Representation Learning Model With Cognitive Dissonance Strategy. *International Journal of Education*, *9*(2), 105. https://doi.org/10.17509/ije.v9i2.5464

Widarti, H. R., Rokhim, D. A., & Rahmaniyah, N. M. S. (2022). Multiple representation-based mobile apps with learning cycle 7e model on colligative properties of solutions. *Educación Química*, *33*(3), 115–126.

Widarti, H. R., Rokhim, D. A., & Syafruddin, A. B. (2020). The development of electrolysis cell teaching material based on stem-pjbl approach assisted by learning video: A need analysis. *Jurnal Pendidikan IPA Indonesia*, *9*(3). https://doi.org/10.15294/jpii.v9i3.25199

Yonata, B. (2021). Four-Tier Diagnostic Test on Chemical Kinetics Concepts for Undergraduate Students. *International Joint Conference on Science and …*, *209*(Ijcse), 457–463.