IJECA
International Journal of
Education & Curriculum Application
http://journal.ummat.ac.id/index.php/IJECA
ISSN 2614-3380 | Vol. 5, No. 1, April 2022

Developing of Four-Tier Diganostic Test to Identify Test Profile on Acid and Base Materials

Isna Rezkia Lukman[1], Ratna Unaida[2], Fakhrah[3]
[1,2,3]Departement of Chemistry Education, Malikussaleh University, Indonesia
ratna.unaida@unimal.ac.id

Keywords:

Misconceptions,
4D Models,
Acid and Base

ABSTRACT

Misconceptions that occur continuously and protracted can interfere with the learning process. Therefore, educators must be able to identify precisely the misconceptions that occur. This research was conducted to find out the implementation of four-tier diagnostic tests in analyzing the misconceptions of students in high school in North Aceh regency and understand the results of analysis of the comprehension level of learners' concepts, whether they Scientific conception, Lack of knowledge, and misconception. Four-tier diagnostic test instrument has been developed by adapting 4D models (Define, Design, Development and Disseminate). Through these steps, the four-tier diagnostic test instrument is produced that is valid through content validation with content-validity coefficient of 0.89 and an average aspect of 4.5 and includes an excellent category. On average, the overall profile of learners got 44.47% who Scientific conception, 44.03% Lack of knowledge and 11.50% who experienced misconceptions in acids and bases materials.The implication of this study is that educators can analyze scientific conception, lack of knowledge and misconceptions on acids and bases materials in learners, and use suitable learning methods so that the misconception is not repeated.

Article History:
Received: 02-12-2021
Revised : 01-01-2022
Accepted: 04-01-2022
Online : 10-04-2022

https://doi.org/10.31764/ijeca.v5i1.6359

A. INTRODUCTION

Science develops based on theory and experimentation. There are five important aspects of science, namely: (1) empirical scientific knowledge, (2) reliable and tentative, (3) based on theory, (4) the inferential nature of science, which appreciates the fundamental difference between conclusion and observation results, (5) creativity that plays an important role in the development of science (Abd-El-Khalick & Akerson, 2009). Educational experts are of the view that science should be taught through experiments, conceptual changes, processes, and other approaches (Abd-El-Khalick & Akerson, 2009) (Magnusson et al., 1999). Therefore, concepts become very important in the teaching of science. Teaching science is the art of teaching, how to make students have problem solving and inquiry skills. Furthermore, in the teaching-learning process, students must be able to make learning meaningful (Leach &Scott, 2003). But it should be underlined, that if an educator fails to provide meaningful lessons, it will affect the skills of problem solving and inquiry students (Abd-El-Khalick & Akerson, 2009). A teacher must be thoroughly prepared and mastering the concept before teaching (McDermott, 2006). When

teachers have difficulty teaching in the classroom, it will hinder the meaningful learning process, causing misconceptions, the development of problem solving skills and student inquiry will be hampered (Mudau, 2013).

According to Eggen & Kauchak, (2004) concept is an idea, object, or event that can help a person in understanding a new problem faced. In chemistry learning, abstract concepts become one of the causes of learners having difficulty concluding concepts, or can infer concepts, but not in accordance with the actual concept. Concepts that are understood not in accordance with this scientific concept can last a long time and are difficult to improve throughout the time of formal education, because the concept can explain the problem faced, even though wrong Suparno (2013) This error should have been known from the beginning, so as not to interfere with the next learning process (Novick & Nussbaum, 1981) (Martin et al., 2002). On the other hand, misconception is interpreted as a misconception of a concept based on experience (Martin et al., 2002). Learners bring concepts that they understand themselves to understand the new knowledge they acquire. Unfortunately, the concepts he understood often did not conform to the concept of science in the classroom. (Bevir, 2003; Georghiades, 2000; Kang et al., 2004; Macbeth, 2000; Venville, 2004).

To measure the understanding of learners in accordance with scientific concepts, an evaluation is needed. But in fact, the use of the instrument given by the teachers only measures cognitive abilities without knowing whether learners already understand the concept, misconceptions or do not understand the concept. A misconception that occurs should not be allowed because it will cause adverse effects for students Lestari et al., (2019). In addition, if students are allowed to experience misconceptions, they will experience more complex misconceptions (Qurrota & Nuswowati, 2018). In identifying misconceptions, the easiest to use and apply to students is diagnostic tests (Suwarto, 2013). Diagnostic tests can make it easier for the teacher to classify students who experience misconceptions and learners who do not know the concept (Jubaedah et al., 2017). The use of diagnostic tests before and after learning can help teachers in diagnosing misconceptions experienced by students in the learning process (Diani et al., 2019). Information obtained on the realization of diagnostic tests can be used to support resolving difficulties experienced by learners (Suwarto, 2013). By knowing the difficulties and level of understanding of students, teachers can determine learning goals appropriately.

One form of diagnostic test is four-tier multiple choices. This test is a development of the three-tier diagnostic test, where the multiple choice consists of 3 distractors and 1 answer key, the level of confidence in choosing the answer, and the reason for choosing the answer. The level of confidence in choosing answers and reasons is divided over a scale of one to six. Scale one is selected if students guess, scale two if students are very unsure, scale three if students are unsure, scale four if students are sure, scale five if students are very confident and scale six if students are very confident (Wilantika et al., 2018).

B. METHODS

This study is an R&D (research and development) study with the aim to produce a 4-tier diagnostic test instrument to uncover the misconceptions of learners in acid and base materials. Research instruments that will be used in diagnosing the level of understanding and misconceptions of learners are developed in as many as 30 problems type 4-tierdiagnostic test, adaptation of the Tiagrajan 4D model namely 4 stages of Define, Design, Development and Dissemination.

Data collection techniques use questionnaires to find out the validity of products that have been developed. Validity used in the form of content validity consisting of visible validity and logical validity. The validity of the appearance is validated by 3 lecturers of chemical education as expert judgement and logical validity as many as 7 chemistry teachers, logical validation is used to find out the feasibility of problems that have been compiled based on aspects of substance, construction, language, validity, and practicality. The type of data obtained in this study in the form of qualitative data and quantitative data, qualitative data obtained based on the validation of the appearance where revisions and suggestions from expert judgment become the basis for improving the preparation of 4-tier instruments. Quantitative data is obtained from logical validation using non-test instruments. Analysis of validity data appears to be suggestions of improvement in the preparation of instruments, while for logical validity uses the aiken formula (Aiken 1980) to calculate content-validity coefficient coefficients of content validity based on the results of assessments from expert panels. After that stage, four-tier diagnostic tests are used to distinguish learners who Scientific conception, Lack of knowledge and misconceptions, by looking at the combination of learners' answers with their level of answer confidence. The answer combination can be seen in Table 1.

Table 1. Combination of answers *four-tier diagnostic test*

Tier				Decision
I	**II**	**III**	**IV**	
Correct	Sure	Correct	Sure	Scientific Conception
Correct	Sure	Correct	Not sure	Lack of Knowledge
Correct	Not sure	Correct	Sure	Lack of Knowledge
Correct	Not sure	Correct	Not sure	Lack of Knowledge
Correct	Sure	Incorrect	Sure	Misconceptions
Correct	Sure	Incorrect	Not sure	Lack of Knowledge
Correct	Not sure	Incorrect	Sure	Lack of Knowledge
Correct	Not sure	Incorrect	Not sure	Lack of Knowledge
Incorrect	Sure	Correct	Sure	Lack of Knowledge
Incorrect	Sure	Correct	Not sure	Lack of Knowledge
Incorrect	Not sure	Correct	Sure	Lack of Knowledge
Incorrect	Not sure	Correct	Not sure	Lack of Knowledge
Incorrect	Sure	Incorrect	Sure	Misconceptions
Incorrect	Sure	Incorrect	Not sure	Lack of Knowledge
Incorrect	Not sure	Incorrect	Sure	Lack of Knowledge
Incorrect	Not sure	Incorrect	Not sure	Lack of Knowledge

(Gurel et al., 2015)

The analysis was conducted on as many as 120 learners from 2 schools in North Aceh regency, using percentage techniques to determine learners who Scientific conception, Lack of knowledge and misconceptions. Furthermore, an analysis of the understanding of learners by summing up the percentage of learners who Scientific conception, Lack of knowledge, and misconceptions. To find out more about the cause of misconceptions experienced by learners, structured interviews are conducted to learners.

C. RESULT AND DISCUSSION

1. Defining

The development of four-tier diagnostic tests aims to describe the profile of understanding, not understanding, and misconceptions experienced by learners in acid-base material, each item of the problem developed consists of four levels. The test consists of four levels, namely, (1). The first level is a multiple choice with five options, (2) The second level is the level of confidence of learners in choosing first level answers, (3) The third level is a concept question that is the reason in answering the question, and (4) The fourth level is the confidence level of the student's answer to the reason in the third level. The basic competency (KD) used is KD 3.10, explaining the concept of acids and bases as well as their strength and equilibrium in solution, as well as 4.10, analyzed the pH change trajectory of several indicators extracted from natural materials through experiments. There are 30 questions, consisting of 7 indicators of achievement, namely (1) Embed the theory of acid bases according to Arrhenius, Bronsted-Lowry, and Lewis, (2) Classifying compounds in everyday life that are acid or base properties, (3) Linking the strength of acids and bases with the degree of ionization and acidic setting or base setting (4) Determining the pH of acid and base solutions, (5) Distinguishing natural and artificial indicators (6) Determining the pH tray in a solution using an indicator (7) Proving acidic solution with various indicators

2. Designing

At this stage, the preparation of problem items based on the grid of questions that have been made at the stage of test development. The competency achievement indicator to be measured is spelled out in questions that correspond to the test grid. This step is done so that all learning indicators that have been set before can be covered in each point of the problem. The next step is the development of a tailing rubric. In this step, the preparation of assessment guidelines that contain the desired criteria in assessing the answers of test takers. The brushing rubric for each point of the problem is clearly arranged and used as a reference in diagnosing the level of understanding of learners. Understanding student concepts in the four-tier diagnostic test assessment instrument can be classified into four categories, namely understanding concepts, not understanding concepts, misconceptions and errors, category errors are intended if there are learners who do not answer one of the four levels. Here is an example of a four-tier diagnostic test that has been compiled used in table 2.

Table 2. Example of a question four-tier diagnostic test

Tier	Questions and answer options
Tier I	Mother is making cake. Mom looks very happy to see the results that expand perfectly, as mom adds baking soda to the bread dough. Baking soda contains a compound that can react with acidic compounds in other ingredients, such as milk, chocolate, yogurt, and buttermilk. The reaction of baking soda with the acid from these ingredients forms carbon dioxide gas which causes the bread to expand. The compounds in the baking soda are... a. Calcium carbide b. Sodium tetraborate c. Ascorbic acid d. Sodium bicarbonate e. Sodium bicarbonate and sodium tetraborate
Tier II	Answer Confidence Level a. Sure b. Not sure
Tier III	The reason you chose this answer is... a. Sodium bicarbonate and sodium tetraborate are both acidic so

	they quickly expand b. Sodium bicarbonate is contained in baking soda, so it can help inflate the cake. c. Sodium tetraborate may help inflate cakes d. Calcium carbide or carbine to accelerate maturation e. Ascorbic acid helps bread expand
Tier IV	Answer Confidence Level a. Sure b. Not sure

3. Developing

Validation is carried out by three expert lecturers in the field of chemical education as expert judgement. Validation is done to determine whether the instrument that has been prepared is worthy both in material concept, evaluation, and language. Data obtained from expert judgement in the form of qualitative data in the form of suggestions used as consideration in revising the problem. The diagnostic test instrument developed has been declared valid by the validator. This shows that the problems that have been developed have a material concept, evaluation and language. The concept of material in an instrument is assessed well by the validator, meaning that the material in the instrument is accurate and in accordance with KI and KD. As mentioned by (Firdaus et al., 2014) the feasibility of the material is seen from the conformity of the material with the core competencies (KI) and basic competencies (KD) of the current subjects, the accuracy of the material, and the supporting materials of learning.

Instrument validation is carried out to ensure that the test instruments developed are in accordance with learning indicators (Wardany & Anjarwati, 2020). Validator has assessed the evaluation aspect of this instrument, and this instrument is declared valid, so that the test instrument developed can measure the ability of learners and can distinguish learners who Scientific conception, Lack of knowledge, or who experience misconceptions. Kartikasari et al., (2015) states that language feasibility includes the use of language that is communicative, dialogical and interactive, straightforward, has a flow that is direct, coherent, in accordance with the rules of good and correct Indonesian, as well as using terms and symbols that are in accordance with the development of learners. Similarly, it is mentioned in BSNP the use of language to explain concepts or applications of concepts or discourses, texts, images, and illustrations up to abstract examples in accordance with the intellectual writing of learners. Because the language aspect in this instruction is considered valid by the validator, it means that the language aspect in this instrument has met the criteria mentioned above.

This instrument involves aspects of substance, construction, language, validity and practicability, 7 chemistry teachers. In calculating content-validity coefficient based on assessments by 7 chemistry teachers, in proving the validity of contents through the linkert scale analyzed using aiken's index V. The results of the Aikens V index for each item are presented in table 3. Validation is carried out by three expert lecturers in the field of chemical education as expert judgement. Validation is done to determine whether the instrument that has been prepared is worthy both in material concept, evaluation, and language. Data obtained from expert judgement in the form of qualitative data in the form of suggestions used as consideration in revising the problem. The diagnostic test instrument developed has been declared valid by the validator. This shows that the problems that have been developed have a material concept, evaluation and language.

Table 3. Aikens V Value

Aspects	Grain	Aikens V index	Average of every aspect
Substance	1	0,82	0,81
	2	0,79	
	3	0,82	
Construction	4	0,89	0,88
	5	0,86	
	6	0,89	
Language	7	0,86	0,89
	8	0,86	
	9	0,93	
	10	0,93	
Validity	11	0,89	0,88
	12	0,86	
Practicality	13	0,89	0,95
	14	0,93	
	15	0,96	
	16	1,00	
Total			**0,89**

The results of validation from the teacher then analyzed using the aiken formula to determine aiken's formula V. Aiken (1985) states that to find out the sigmification of statistics from V, it can be determined by correlated the ratings category with the number of raters (experts). In this study uses seven raters with five category scales. So, in accordance with Aiken (1985) statement that with a signification level of 0.05 the limit for V for Aikens V per item of question is 0.76. The lowest problem item of 0.79 and the overall average of the category of 0.89 so that in the four-tier diagnostic test instrument on the base acid material proved valid in the validity of the contents.

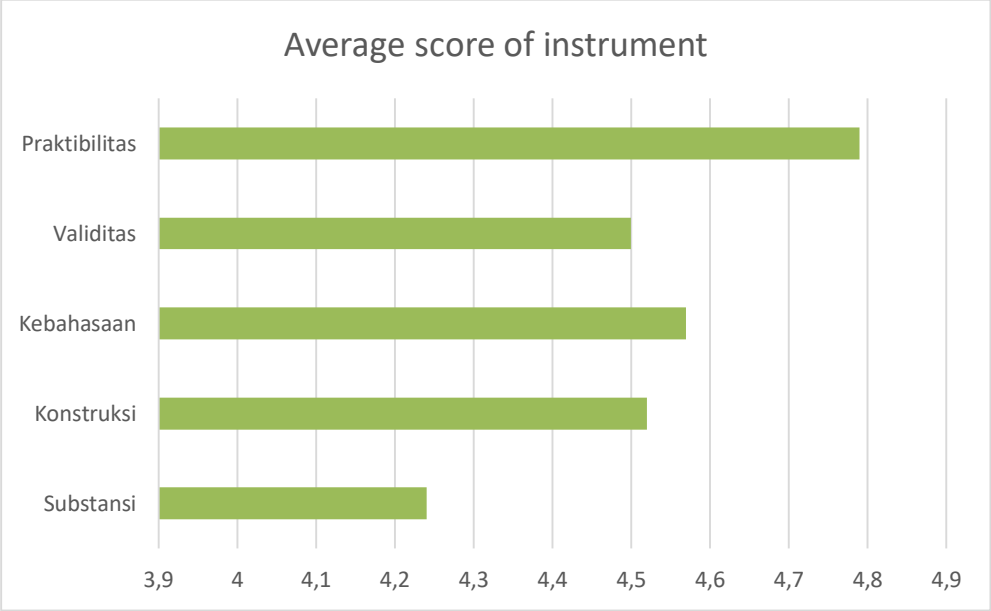

Figure 1. Average score of aspects on instruments with linkert scale of 1-5

Based on the results of the feasibility assessment of five aspects, namely aspects using the linkert scale, the substance obtained an average score of 4.24 in the range of scores $\bar{X}$> 4.2 categories very good. The construction aspect obtained an average score of 4.52 in the score range $\bar{X}$> 4.2 categories very well. The language aspect obtained an average score of 4.57 is in the range of scores $\bar{X}$> 4.2 categories very good. Valid aspects obtained an average score of 4.52 are in the range of scores $\bar{X}$ > 4.2 categories are very good. Aspects of practical ability obtained an average score of 4.79 in the score range of $\bar{X}$> 4.2 categories very well. Based on the overall feasibility analysis of aspects obtained an average score of 4.52 categories is very good, so that periodic diagnosis instruments are suitable for use. assessment by instrument experts aims to assess the feasibility of the instrument and the truth of the concept.

Table 4. Percentage of Student Conception Profile

Competency Achievement Indicators	Question Items	Scientific Conception (%)	Lack of Knowledge (%)	Misconceptions (%)
Distinguishing acid-base theory according to Arrhenius, Bronsted-Lowry, and Lewis	1	55,83	39,17	5,00
	2	35,83	57,50	6,67
	3	35,00	60,00	5,00
	4	28,33	65,00	6,67
	5	35,83	57,50	6,67
Average	1-5	38,17	55,83	6,00
Classify compounds in everyday life that have acid or base properties	6	37,50	59,17	3,33
	7	46,67	45,83	7,50
	8	35,83	45,00	19,17
	9	65,00	25,00	10,00
Average	6-9	46,25	43,75	10,00
Connects the strength of acids and bases with the degree of ionizing and acid remains or base keepings	10	72,50	16,67	10,83
	11	26,67	70,83	2,50
	12	37,50	56,67	5,83
	13	41,67	39,17	19,17
	14	35,83	54,17	10,00
Average	10-14	42,83	47,50	9,67
Determines the pH of acid and base solutions	15	54,17	33,33	12,50
	16	35,83	45,00	19,17
	17	17,50	65,00	17,50
Average	15-17	35,83	47,78	16,39
Distinguishing natural and artificial indicators	18	28,33	53,33	18,33
	19	46,67	43,33	10,00
	20	47,50	34,17	18,33
	21	56,67	23,33	20,00
Average	18-21	44,79	38,54	16,67
Determines the pH tray in a solution using indicators	22	71,67	18,33	10,00
	23	56,67	25,00	18,33
	24	56,67	33,33	10,00
	25	71,67	24,17	4,17
	26	45,00	44,17	10,83
Average	22-26	60,33	29,00	10,67
Prove acidic solution with various indicators	27	28,33	68,33	3,33
	28	26,67	55,83	17,50
	29	45,00	35,83	19,17

	30	55,83	26,67	17,50
Average	27-30	38,96	46,67	14,38
Overall average		**44,47**	**44,03**	**11,50**

4. Disseminating

After the validity of the contents by way of visible validity and logical validity, diagnostic tests were conducted on 120 learners from 2 high schools in North Aceh Regency. This diagnostic test aims to find out the level of understanding of learners towards acid-base matter. Based on Table 4 the most learners experience misconceptions in the point of problem 21, which is as much as 20%. While the material that has the least misconception is in the point of question 6 and 27, which is3.33% 3.33% each. In question point number 10, as many as 72.5% of learners already understand the material, and is the most widely understood material. The most widely understood material is the material in the 11th point of the problem, the acid-base theory, which is as much as 70.83% of learners. On the achievement indicator distinguishes the theory of acid bases according to Arrhenius, Bronsted-Lowry, and Lewis the lowest level of misconception of learners and those who lack of knowledge the most. While on the achievement indicator determines the pH trajectory in a solution by using the indicator the majority of students already Scientific conception with a percentage of 60.33%. On average, learners get 44.47% who Scientific conception, 44.03% lack of knowledge and 11.50% who gets misconceptions.

This study shows results that are in line with previously conducted research by Juliani et al (2021) showing that diagnostic tests are able to analyze a number of misconceptions and classify the level of understanding into misconceptions, lack of concepts, and errors. Research analysis of student misconceptions on acids and bases materials produces information on how the student's mindset when answering a question related to the concepts contained in the material. Based on the information obtained, it is expected that teachers can determine the appropriate and suitable learning methods used in pressure learning and its application so that misconceptions on the material are not experienced by other students.

D. CONCLUSION AND SUGGESTIONS

The four-tier diagnostic test developed in this study has been assessed by a validator, and assessed valid, with an average score of 4.52, or including an excellent category, so that it can be used to determine the level of understanding of learners. The four-tier diagnostic test was conducted on 120 learners from 2 high schools in North Aceh Regency, and the result was a misconception on the acid-base material experienced by the learners at most by 20%, namely on the point of problem 21, and at least on the point of problem 6 and 27, each as much as 3.33%. On average, overall learners received 44.47% who Scientific conception, 44.03% did not lack of knowledge and 11.50% who experienced misconceptions in acid and base materials.

ACKNOWLEDGEMENT

This Editorial reflects contribution from many past and present members of the Chemistry Education group at the Malikusslaeh University, and Institute for Research and Community Service (LPPM), Malikussaleh University for the support of the 2021 PNBP funding service.

REFERENCES

Abd-El-Khalick, F., & Akerson, V. (2009). The influence of metacognitive training on preservice elementary teachers' conceptions of nature of science. *International Journal of Science Education*, *31*(16), 2161–2184. https://doi.org/10.1080/09500690802563324

Aiken, L. R. (1985). Three Coefficients For Analyzing The Reliability And Validity of Ratings. *Educational and Psychological Measurment*, 45.

Bevir, M. (2003). Notes Toward an Analysis of Conceptual Change. *Social Epistemology*, *17*(1),

55–63. https://doi.org/10.1080/0269112032000114831

Diani, R., Alfin, J., Anggraeni, Y. M., Mustari, M., & Fujiani, D. (2019). Four-Tier Diagnostic Test with Certainty of Response Index on the Concepts of Fluid. *Journal of Physics: Conference Series*, *1155*(1). https://doi.org/10.1088/1742-6596/1155/1/012078

Eggen, P., & Kauchak, D. (2004). *Educational Psychology: Windows, Classroom*. Upper Saddle River: Pearson Prentice Hall.

Firdaus, A., Samhati, S., & Suyanto, E. (2014). *Analisis Kelayakan Isi Buku Teks Bahasa Indonesia Terbitan Erlangga Kelas VII SMP/MTs*. *1*, 1–12.

Georghiades, P. (2000). Beyond conceptual change learning in science education: Focusing on transfer, durability and metacognition. *Educational Research*, *42*(2), 119–139. https://doi.org/10.1080/001318800363773

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

Jubaedah, D. S., Kaniawati, I., Suyana, I., Samsudin, A., & Suhendi, E. (2017). Pengembangan Tes Diagnostik Berformat Four-Tier Untuk Mengidentifikasi. *Prosiding Seminar Nasional Fisika (E-Journal) SNF 2017*, *VI*(October), 35–40.

Kang, S., Scharmann, L. C., & Noh, T. (2004). Reexamining the role of cognitive conflict in science concept learning. *Research in Science Education*, *34*(1), 71–96. https://doi.org/10.1023/B:RISE.0000021001.77568.b3

Kartikasari, Y., Widodo, M., & Karomani. (2015). Kelayakan Isi Dan Bahasa Pada Buku Teks Bupena Bahasa Indonesia Kelas VII. *Jurnal Kata (Bahasa, Sastra, Dan Pembelajarannya)*, 1–9.

Leach, J., & Scott, P. (2003). Individual and Sociocultural Views of Learning in Science Education. *Science and Education*, *12*(1), 91–113. https://doi.org/10.1023/A:1022665519862

Lestari, D., Handayani, D., Darussyamsu, R., & Armen, A. (2019). Identification Students' Misconceptions of Class VIII SMPN 21 Padang in the Skeletal Systems of Organism by Using CRI Technique. *Jurnal Atrium Pendidikan Biologi*, *4*(1), 135. https://doi.org/10.24036/apb.v4i1.5470

Macbeth, D. (2000). On an actual apparatus for conceptual change. *Science Education*, *84*(2), 228–264. https://doi.org/10.1002/(SICI)1098-237X(200003)84:2<228::AID-SCE6>3.0.CO;2-3

Martin, R., Sexton, C., Franklin, T., Gerlovich, J., & McElroy, D. (2002). *Teaching Science for all Children: Methods for Conctructing Understanding*. Allyn and Bacon.

McDermott, L. C. (2006). Preparing K-12 teachers in physics: Insights from history, experience, and research. *American Journal of Physics*, *74*(9), 758–762. https://doi.org/10.1119/1.2209243

Mudau, A. V. (2013). Teaching Difficulties from Interactions and Discourse in a Science Classroom. *Journal of Educational and Social Research*, *3*(3), 113–120. https://doi.org/10.5901/jesr.2013.v4n3p113

Novick, S., & Nussbaum, J. (1981). Pupils' understanding of the particulate nature of matter: A cross-age study. *Science Education*, *65*(2), 187–196. https://doi.org/10.1002/sce.3730650209

Qurrota, A., & Nuswowati, M. (2018). Analisis Miskonsepsi Siswa Menggunakan Tes Diagnostic Multiple Choice Berbantuan Cri (Certainty of Response Index). *Jurnal Inovasi Pendidikan Kimia*, *12*(1), 2108–2117.

Suparno. (2013). *Miskonsepsi dan Perubahan Konsep Pendidikan Fisik*. Grafindo.

Suwarto. (2013). *Pengembangan tes diagnostik dalam pembelajaran*. Grafindo.

Venville, G. (2004). Young children learning about living things: A case study of conceptual change from ontological and social perspectives. *Journal of Research in Science Teaching*, *41*(5), 449–480. https://doi.org/10.1002/tea.20011

Wardany, K., & Anjarwati, S. (2020). Kelayakan Instrumen Penilaian Higher Order Thinking

Skills pada Materi Lingkungan. *IISEJ : Ndonesian Science Education Journal*, *1*(3), 226–237.

Wilantika, N., Khoiri, N., & Hidayat, S. (2018). Pengembangan Penyusunan Instrumen Four-Tier Diagnostic Test Untuk Mengungkap Miskonsepsi Materi Sistem Ekskresi Di SMA Negeri 1 Mayong Jepara. *Jurnal Phenomenon*, *08*(2), 200–214.