JPPIPA 9(1) (2023)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Development of Higher Order Thinking Skills-Based Assessment Instrument on Acid-Base Materials in High School

Nunung Ratna Ningsih[1], Agus Kamaludin[1*]

[1] Chemistry Education, Faculty of Education and Teaching Science, Sunan Kalijaga State Islamic University, Yogyakarta, Indonesia.

Received: March 5, 2022
Revised: January 9, 2023
Accepted: January 25, 2023
Published: January 31, 2023

Corresponding Author:
Agus Kamaludin
aguskamaludin@gmail.com

DOI: 10.29303/jppipa.v9i1.1457

Abstract: Higher Order Thinking Skills (HOTS) are necessary skills to deal with the 21st century. This research aimed to develop a HOTS-based assessment instrument on acid-base for class XI that meets the aspects of validity, reliability, difficulty level and distinguishing power. This research is a Research and Development (R&D) study with a 4D model. The quality of the product was assessed by one material expert lecturer, one media expert lecturer, and four chemistry teachers of high school and tested on 30 science students in eleventh grade. The assessment of product quality obtained from material experts is 97.67% with the Very Good (VG) category; from media expert, it was 97.44% with the Very Good (VG) category; according to chemistry teachers, all items are feasible to be used as a field trial. After conducting the field trial, the product contained 18 valid items and 2 invalid items with the very reliable category, which was 0.899. At the difficulty level, the product had 11 easy items and 9 medium items. At the level of distinguishing power, the product had 1 very bad item, 2 sufficient items, 9 good items, and 8 very good items. Therefore, the resulting assessment questions instrument can be used to measure students' HOTS knowledge.

Keywords: Assessment Instrument; Higher Order Thinking Skills; Acid-Base Solution

Introduction

Higher Order Thinking Skills (HOTS) are skills that are very much required to face the 21st century (Chalkiadaki, 2018). HOTS is generally defined as one's ability to think logically, critically, creatively, reflectively, and metacognitively, which is the basic ability to be able to answer challenges, matters, and new problems in everyday life (Marshall & Horton, 2011; Thomas & Thorne, 2009). According to Bloom's revised taxonomy, thinking skills are divided into two, namely Lower Order Thinking Skills (LOTS) and Higher Order Thinking Skills (HOTS). LOTS have three levels. They are remembering (C1), understanding (C2), and applying (C3). Furthermore, HOTS include analyzing (C4), evaluating (C5), and creating/creativity (C6) (Anderson & Krathwol, 2010; Hanifah, 2019). One of the international institutions that measures students' HOTS is Program for International Student Assessment (PISA). Based on PISA evaluation in 2018, Indonesia was ranked 72 out of 77 participating countries (OECD, PISA: 2019). This shows that students in Indonesia are still in the category of Lower Order Thinking Skills (LOTS) (Itsnawati et al., 2019).

One of the efforts made by the government to improve the quality of students' HOTS is to implement the curriculum 2013 which is focused on improving two major parts of the curriculum, namely content standards and assessment standards (Kemendikbud, 2017). The content standard is designed so that students are able to think critically when receiving various types of information, think creatively when solving problem by using their knowledge, and make decisions in complex situations (Saputra, 2016; Saido et al., 2015; Yen & Halili, 2015). Furthermore, the assessment standard is created by adapting international standard assessment models in which the evaluation process focuses more on HOTS. In term of implementing HOTS in the assessment standard, HOTS-based question is apllied in the national exam (known with UN) (Kemendikbud, 2018). However, the average score of the 2019 National Exam, especially in chemistry, was still below average, which was 50.29 (Puspendik, 2019).

How to Cite:
Ningsih, N. R., & Kamaludin, A. (2023). Development of Higher Order Thinking Skills-Based Assessment Instrument on Acid-Base Materials in High School. *Jurnal Penelitian Pendidikan IPA*, *9*(1), 13–19. https://doi.org/10.29303/jppipa.v9i1.1457

The low average score of chemistry exam using HOTS question requires teachers to be able to compile HOTS question in order that students are able to train their high-order thinking skills (HOTS) (Abosalem, 2015). In term of compossing HOTS question, teachers have to be able to present various information with stimulus in the form of text, images, graphics, tables, and others that contain information based on real life (Kwangmuang et al., 2021). The stages of arranging HOTS question must be followed systematically in order to obtain more effective questions. The stages consist of analyzing basic competencies, compiling question grids, selecting interesting and contextual stimulus, writing question that fits with the grid, and making up scoring guidelines or answer keys (Fanani & Kusmaharti, 2016). However, the main problem in reality is that teachers have not understood and mastered how to compose and develop HOTS questions (Salirawati et al., 2017). Based on interviews with high school chemistry teachers in Yogyakarta, there are few references to HOTS question, so that teachers only arrange textual and simple question according to the material that has been taught. As a result, the question cannot be used as a HOTS instrument with the good quality (Hasanah, 2018).

The instrument which has good quality must have criteria of validity, reliability, and practicality so that it can be used as an accurate instrument to measure the achievement of student learning outcomes. The correct procedure to compose the HOTS instrument is to analyze the instrument qualitatively based on considerations of substance, construction, and language. Then, it is also analyzed based on empirical data or based on the question that has been tested. Furthermore, according to the results of qualitative and quantitative analysis, the right items are selected (Setiadi, 2016). So far, there is rarely any quality test on instrument composed by teacher so that the instrument being used is almost similar (Salirawati et al., 2017). By the development of students' mindsets which is getting more advanced, it should be followed by teacher's ability to make up instruments that can reveal high-level cognitive aspects such as applying and reasoning question (Lawe et a., 2019). One of the subjects whose HOTS-based test instrument is still limited is chemistry (Iskandar & Senam, 2015).

Chemistry is a science that has an important role and it is very closely related to daily life (contextual). Chemistry is also often referred to as a core science because it has very wide application in various fields (Ihsani et al., 2020). Everything that exists in nature, even in the human body, cannot be separated from chemical substances. Studying chemistry is meant to find out about phenomena in the surrounding environment, so that the learning process does not only learn about cognitive mastery in the form of theory. Here, students also need to sharpen their thinking and reasoning process skills (Nurkholik & Yonata, 2020). Therefore, during learning chemistry, high-order thinking skills (HOTS) is required to solve some problems in the form of theories, concepts, laws, and facts (Danggus, 2014). However, most students consider chemistry as a difficult subject to understand (Yakina, Kurniati, & Fadhilah, 2017) because chemistry deals with reactions, calculations, and abstract concepts (Ismawati, 2017).

One of chemistry materials is acid-base whose phenomena are easy to find in everyday life (Sari & Helsy, 2018). To understand acid-base concept, integration with other concepts such as the particle nature of matters, the nature and composition of solutions, structure of atom, ionic and covalent bonds, symbols, chemical reactions and equations, ionization and chemical equilibrium in the aqueous phase (Sheppard, 2006). In addition, the concept of acid-base emphasizes two components, namely algorithmic and conceptual. Algorithmic understanding is required when determining the concentration of acid-base solution, pH or pOH, looking for Ka and Kb, and the percent ionization. Conceptual understanding involves an explanation of various acid-base phenomena in life (Drechsler & Schmidt, 2005). The learning process of acid-base must be contextual, creative, and critical to explore the deeper knowledge. However, so far, the learning process of acid-base has just involved understanding theories, reactions, formulas and exercises with routine procedure. Therefore, students barely have a chance to practice their HOTS. Based on PAMER application, the percentage of mastery of the materials for chemistry national exam based on basic competences of acid-base in the 2018/2019 academic calendar is 44.91% at the national level (Puspendik, 2019). The lack of availability of HOTS-based test instruments also leads to the lack of students' cognitive mastery on acid-base materials.

This study aims to develop a valid and reliable HOTS item on acids and bases in multiple-choice form. The HOTS question assessment instrument is expected to improve students' higher-order thinking skills in working on HOTS-type questions. In addition, it can also be an alternative for teachers in measuring students' HOTS abilities in acid-base materials.

Method

This research used research and development (R&D) method. According to Sugiyono (2014), research and development method is a research method used to create certain product and test its effectiveness. This research also used a 4D development model consisting of define, design, develop and disseminate. The stage of defining consists of needs analysis and curriculum analysis. The stage of design aims to design the instrument to be developed. The development stage involves validation from material expert lecturer, media expert lecturer, peer reviewers (students majoring in chemistry education), chemistry teachers of high school, and a limited trial for science students in eleventh grade. A limited trial was conducted on 30 science students in eleventh grade of SMA Negeri 8, Yogyakarta.

The research instruments used to develop the product were observation sheets, quality assessment sheets, and HOTS-based question on acid-base. Data analysis technique was done by converting qualitative and quantitative data into descriptive form based on the results of assessment. The assessment conducted by material expert lecturer and media expert lecturer used 5 point Likert scale. The assessment by chemistry teachers of high school used Likert scale that was processed based on the Aiken index.

$$V = \sum S / \left[n(c-1) \right] \quad (1)$$

Description:
V = validity index
c = highest validity score
n= a number of assessors
s = r – Io
r = the score given by the assessor
Io = lowest validity score

The results of Aiken index for an item or device can be categorized based on its index. Items can be accepted if the index obtained is greater than 0.88 (Azwar: 2012).

The data from the limited trial were analyzed by using IBM SPSS 20 and ANATES programs. Validity and reliability were analyzed by using SPSS while level of difficulty and distinguishing power were analyzed by using ANATES.

In this research, validity was calculated with the help of the IBM SPSS 20.0 program. The validity test of the instrument used Pearson Product-Moment Correlation Coefficient with a significance level of 5% as follows (Arikunto, 2018).

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sum X^2 - (\sum X)^2 \left\{ N\sum Y^2 - (\sum Y)^2 \right\}} \quad (2)$$

Description:
r_{xy} = Correlation coefficient between variables of X and Y
X = Score of each item
Y = Total score
$\sum XY$ = The number of multiplications between variables of X and Y
$\sum X^2$ = The number of squares of X variable
$\sum Y^2$ = The number of squares of Y variable
N = The number of subjects

If the decision making is based on the value of rcount (Corrected Item-Total Correlation) > rtable, the question can be considered to be valid. The criteria for classifying the level of validity can be seen in Table 1 below (Arifin, 2013):

Table 1. Classification of Validity Level

Coefficient Interval	Validity Level
0.81-1.00	Very high
0.61-0.80	High
0.41-0.60	Fair
0.21-0.40	Low
0.00-0.20	Very low

The reliability of instrument was measured by using the Cronbach's Alfa method as follows (Arikunto, 2018):

$$r_{11} = \frac{k}{(k-1)} 1 - \frac{\sum S_i}{S_t} \quad (3)$$

Description:
r_{11} = The value of realibility
$\sum S_i$ = The number of variance scores for each item
S_t = Total variance
k = The number of items

This scale was measured based on the Cronbach's Alfa scale, which was 0 to 1. If the scale is grouped into five classes with the same range, the stability measurement of Alpha can be interpreted as follows (Sudjiono, 2015):

a. Cronbach's alpha value of 0.00-0.20 = less reliable
b. Cronbach's alpha value of 0.21-0.40 = quite reliable
c. Cronbach's alpha value of 0.41-0.60 = moderately reliable
d. Cronbach's alpha value of 0,61-0,80 = reliable
e. Cronbach's alpha value of 0,81-1,00 = very reliable

The difficulty index can be determined by using the formula 4.

$$P = \frac{B}{JS} \quad (4)$$

Description:
P = Difficulty index of question
B = The number of students who answer the question correctly
JS = The number of students taking the test

The classification of the difficulty level of question is presented in Table 2 (Arikunto, 2018):

Table 2. The Classification of Difficulty Level

Difficulty Index (P)	Criteria
$0.00 \le P \le 0.30$	Difficult
$0.30 < P \le 0.70$	Moderate
$0.70 < P \le 1.00$	Easy

The distinguishing power of multiple choice questions can be analyzed by using the following formula 5 (Arikunto, 2018):

DP = = PA – PB (5)

Description:

DP : Distinguishing power index

PA : The proportion of upper students who answer correctly

PB : The proportion of lower students who answer correctly

Furthermore, the classification of distinguishing power can be seen in Table 3 below (Arikunto, 2018):

Table 3. The Classification of Distinguishing Power

Distinguishing Power (DP)	Category
$DP \leq 0.00$	Very bad
$0.00 < DP \leq 0.20$	Bad
$0.20 < DP \leq 0.40$	Fair
$0.40 < DP \leq 0.70$	Good
$0.70 < DP \leq 1.00$	Very good

Result and Discussion

The development of HOTS-based assessment instrument on acid-base solutions used a 4-D model involving stages of define, design, develop, and disseminate. The stage of define was carried out by observing the results of the chemistry national exam and interviews with chemistry teachers of high school in Yogyakarta. Based on observation, high school students' mastery on chemistry materials, according to the results of National Exam in the academic year 2018/2019, in which contained HOTS question, were still relatively low. This shows that there are still students who have not been able to apply their higher-order thinking skills when working on question. In addition, based on interviews with teachers, the availability of HOTS-based assessment instruments was still little. One of them is acid base material. Acid-base material is very important so that students can understand about buffer solutions and hydrolysis solutions.

Curriculum analysis was conducted by examining the content standards contained in the curriculum 2013 in which includes a study of core competencies and basic competencies of acid-base materials. The core competence (known with KI) used in the HOTS-based assessment instrument was KI 3 while the basic competencies (known with KD) were KD 3.10 and KD 3.11. Based on KI and KD, indicators, grid question, and HOTS items were then determined.

The design stage aimed to create HOTS-based assessment instrument on acid-base materials. The instrument was 20 multiple choice questions that meet characteristics of HOTS question. Those characteristics were stimulus, contextual, thinking levels of C4-C6, and uncommon question.

In the development stage, validation was conducted by expert and chemistry teachers of high school. Material expert and media expert assessed the instrument by using quality assessment sheet. The data of quality assessment which were obtained from material expert and media expert were in the form of qualitative data which were then converted into quantitative data by using 5 point Likert scale. The results of the quality assessment by material expert can be seen in Table 4.

Table 4. The Quality Assessment by Material Expert

Assessment Aspects	Ideal Percentage (%)	Category
Content	95.33	Very Good
Characteristics of HOTS Question	100.00	Very Good
Role of HOTS Question	100.00	Very Good
Total	97.67	Very Good

Furthermore, the results of quality assessment by media expert can be seen in Table 5.

Table 5. The Quality Assessment by Media Expert

Assessment Aspects	Ideal Percentage (%)	Category
Language	99.25	Very Good
Ghrapics	100.00	Very Good
Design	93.33	Very Good
Total	97.44	Very Good

Validation data, in the form of qualitative data, was obtained based on assessment conducted by chemistry teachers. It was converted into quantitative data by using 5 point Likert scale and it was processed as well as analyzed by using Aiken index. The data of quality assessment by chemistry teachers can be seen in Table 6.

The quality assessment by chemistry teachers consisted of 6 aspects including content, characteristics of HOTS question, role of HOTS question, language, graphics, and design. Those aspects were elaborated into 15 indicators. Based on chemistry teachers' assessment, it can be concluded that the Aiken index of 20 question is greater than 0.88. All HOTS question about acid-base which have been developed are acceptable and can be used for limited trial.

The limited trial was conducted in eleventh grade of IPA 7, at SMA Negeri 8 Yogyakarta. It involved 30 students. The results of data analysis were in the form of validity, reliability, level of difficulty, and distinguishing power.

The validity of items was tested by using the SPSS 20.0 program. Decision making was based on the value of rarithmetic (Corrected Item-Total Correlation) > rtable. The rtable for respondents (students) was 30, df =

N-2 = 30-2 = 28 is 0.361; α = 0.05 meant that the item was considered to be valid. The results of item validity can be seen in Table 7.

Table 6. The Data of Quality Assessment by Chemistry Teachers

No.	Average of Aiken index	Conclusion
1	1.054	acceptable item
2	1.072	acceptable item
3	1.035	acceptable item
4	0.971	acceptable item
5	0.997	acceptable item
6	1.035	acceptable item
7	1.032	acceptable item
8	1.061	acceptable item
9	1.125	acceptable item
10	1.081	acceptable item
11	1.061	acceptable item
12	1.096	acceptable item
13	1.075	acceptable item
14	1.067	acceptable item
15	1.085	acceptable item
16	1.078	acceptable item
17	1.042	acceptable item
18	1.051	acceptable item
19	1.019	acceptable item
20	1.077	acceptable item
Average	1.056	acceptable item

Table 7. The Results of Item Validity

No. Item	*(Pearson Correlation)*	Correlation Probability [sig. (2-tailed)]	Conclusion
1	0.579	0.001	Valid
2	0.623	0.000	Valid
3	0.551	0.002	Valid
4	0.542	0.002	Valid
5	0.713	0.000	Valid
6	0.809	0.000	Valid
7	0.609	0.000	Valid
8	0.636	0.000	Valid
9	0.488	0.006	Valid
10	0.809	0.000	Valid
11	0.817	0.000	Valid
12	0.825	0.000	Valid
13	0.615	0.000	Valid
14	0.107	0.574	Invalid
15	0.441	0.015	Valid
16	0.550	0.002	Valid
17	0.562	0.001	Valid
18	0.565	0.001	Valid
19	0.707	0.000	Valid
20	0.358	0.052	Invalid

Based on the results of the validity test by using SPSS 20.0, 18 items were declared valid. Items number 14 and 20 have r_{count} (Corrected Item-Total Correlation) < r_{table} so that those were declared invalid.

An analysis of reliability test was viewed from the value of Cronbach's alpha gained from reliability statistics, an analysis by using the SPSS program. The result of item reliability analysis can be seen in Table 8.

Table 8. The Result of Item Reliability Analysis

Reliability Score	Criteria
0.899	Very reliable

The value of Cronbach's alpha which was obtained in this research was 0.899. It meant that the reliability of test instrument which was aimed to measure students' HOTS on acid-base material in class XI IPA was categorized as very reliable.

The difficulty level of HOTS-based multiple choice question in this research was analyzed by using 4.0.9 version of ANATES application. Based on the results of analysis conducted on 20 multiple choice question, it can be seen that there are 9 items in the medium category and 11 items in the easy category. The distribution of difficulty level of question can be seen in Table 9 below.

Table 9. The Distribution of Difficulty Level of Question

Index of Difficulty Level	Category	Question Number	Total	Percentage (%)
0.70 < P ≤ 1.00	Easy	1, 3, 5, 6, 7, 10, 11, 12, 13, 15, 17	11	55.00
0.30 < P ≤ 0.70	Medium	2, 4, 8, 9, 14, 16, 18, 19, 20	9	45.00

Distinguishing power of HOTS–based multiple choice question on acid-base materials was analyzed by using ANATES application. Based on an analysis conducted on 20 multiple choice question, it was known that 1 item was in the very bad category, 2 items were in the sufficient category, 9 items were in the good category, and 8 items were in the very good category. The distribution of distinguishing power of question can be seen in Table 10.

Table 10. The Distribution of Distinguishing Power

Index of distinguishing power	Category	Question Number	Total	Percentage (%)
DP ≤ 0.00	Very Bad	14	1	5.00
0.00<DP≤0.20	Bad	-	-	0.00
0.20<DP≤ 0.40	Sufficient	1, 15	2	10.00
0.40<DP≤ 0.70	Good	3, 6, 7, 8, 10, 13, 16, 17, 20	9	45.00
0.70<DP≤ 1.00	Very Good	2, 4, 5, 9, 11, 12, 18, 19	8	40.00

This research produced a product, HOTS-based assessment instrument on acid-base material for eleventh grade, which can be used as an assessment instrument. HOTS-based assessment instrument consists of 20 multiple choice questions with 5 answer choices in each question. The assessment instrument, which has been developed, consists of stimulus, C4 (analyzing), C5 (evaluating), C6 (creating), multi-concept, unusual, and contextual question that can be used by teachers to measure students' higher order thinking skills (HOTS).

HOTS-based assessment instrument was assessed by material expert lecturer, media expert lecturer, and reviewers (chemistry teachers). Based on quality assessment by material expert, the product gained a percentage of 97.67% and it was categorized as Very Good (VG). Furthermore, according to quality assessments by media expert, a percentage of 97.44% was obtained and it was categorized as Very Good (VG). Hereafter, based on quality assessments by reviewers/ chemistry teachers, all items were acceptable and suitable for use in field trial. Overall, HOTS-based assessment instrument has good validity and reliability, easy and medium difficulty levels, and quite good distinguishing power. Research by Widhiyani et al. (2019) states that valid and reliable HOTS questions can be used to measure students' abilities in learning. In addition, valid and reliable question instruments can be used in further research (Hayati & Lailatussaadah, 2016)

Conclusion

Based on the results, it can be concluded that development of HOTS-based assessment instrument on acid-base material uses a 4D model that meets the aspects of feasibility, validity, reliability, level of difficulty, and distinguishing power. The quality of product is assessed by material expert and it gets a percentage of 97.67% with the Very Good (VG) category; from media expert, it gets a percentage of 97.44% with the Very Good (VG) category; from chemistry teachers of high school, it contains acceptable and suitable items which can be used for field trial. Meanwhile, based on a limited trial, 18 valid items and 2 invalid items are obtained in the product; at the reliability level, the product is categorized as very reliable with Cronbach's alpha value of 0.899; at difficulty level of items, the product contain 11 easy items and 9 medium items; at the level of distinguishing power, it has 1 very bad item. 2 sufficient items, 9 good items, and 8 very good items. Therefore, the resulting assessment questions instrument can be used to measure students' HOTS knowledge.

References

Abosalem, Y. 2015. Assessment techniques and students' higher-order thinking skills. *ICSIT 2018 - 9th International Conference on Society and Information Technologies, Proceedings, March,* 61–66. https://doi.org/10.11648/j.ijsedu.20160401.11

Anderson, L. W., & Krathwohl, D. R. (2010). *Kerangka landasan untuk pembelajaran, pengajaran, dan asesmen: Revisi taksonomi bloom.* Yogyakarta: Pustaka Pelajar.

Arifin, Z. (2013). *Evaluasi pembelajaran*. Bandung: PT. Remaja Rosdakarya.

Arikunto, S. (2018). *Prosedur penelitian: Suatu pendekatan praktik.* Jakarta:Rineka Cipta.

Azwar, S. (2012). *Reliabilitas dan validitas.*Yogyakarta: Pustaka Pelajar

Chalkiadaki, A. (2018). A systematic literature review of 21st-century skills and competencies in primary education. *International Journal of Instruction,* 11(3), 1–16. https://doi.org/10.12973/iji.2018.1131a

Danggus, G. (2014). Meningkatkan hasil belajar materi polimer melalui penerapan model pembelajaran kooperatif tipe numbered heads together pada siswa kelas XII IPA SMAN 2 Pontianak. *Jurnal Pendidikan Matematika dan IPA, 5*(2), 9-20. https://doi.org/10.26418/jpmipa.v5i2.13335

Drechsler, M. & Schmidt, H. (2005). Upper secondary school understanding of models used in chemistry to define acids and bases. *Science Education International. 16*(1), 39-53. https://doi.org/10.5617/nordina.384

Fanani, A., & Kusmaharti, D. (2016). Pengembangan pembelajaran berbasis HOTS di sekolah dasar kelas V. *Jurnal Pendidikan Dasar,* 9(1), 1-11. https://doi.org/10.21009/JPD.091.01

Hanifah, N. (2019). Pengembangan instrumen penilaian higher order thinking skill (hots) di sekolah dasar. *Conference Series*, 1(1), 1–8. Retrieved from http://ejournal.upi.edu/index.php/crecs/article/view/14286

Hasanah, U. (2018). Pengembangan soal kimia berbasis *higher order thinking skills* (HOTS) materi kesetimbangan kimia kelas XI SMA/MA. *Skripsi,* tidak diterbitkan, Universitas Negeri Malang, Malang.

Hayati, S., & Lailatussaada. (2016). Validitas dan reliabilitas instrumen pengetahuan pembelajaran aktif, kreatif, dan menyenangkan (PAKEM) menggunakan model Rasch. *Jurnal Ilmiah Didaktika, 16*(2), 169-179. http://dx.doi.org/10.22373/jid.v16i2.593

Ihsani, N., Idrus, A. A., & Jamaludin, J. (2020). Perangkat pembelajaran biologi berbasis masalah terintegrasi nilai-nilai islami untuk meningkatkan penguasaan konsep peserta didik. *Jurnal Pijar Mipa*, 15(2), 103–109. https://doi.org/10.29303/jpm.v15i2.1326

Iskandar, D., & Senam, S. (2015). Studi Kemampuan Guru Kimia Sma Lulusan Uny Dalam Mengembangkan Soal Uas Berbasis Hots. *Jurnal Inovasi Pendidikan IPA,* 1(1), 65. https://doi.org/10.21831/jipi.v1i1.4533

Ismawati, R. (2017). Strategi REACT dalam pembelajaran kimia SMA. *Indonesian Jurnal of Science and Education, 1*(1), 1-7. https://dx.doi.org/10.31002/ijose.v1i1.413

Itsnawati, S., Budiyanto, M., & Purnomo, A. R., (2019). Validitas lembar kerja siswa berbasis higher order thinking skills dengan tema pencemaran lingkungan untuk meningkatkan keterampilan pemecahan masalah. *Jurnal Pendidikan Sains, 7*(3), 377-381. Retrieved from https://ejournal.unesa.ac.id/index.php/pensa/article/view/32297.

Kemendikbud. (2018). *Modul penyusunan soal higher order thinking skills (HOTS).* Jakarta: Kementerian Pendidikan dan Kebudayaan.

Kubiszyn, T., & Borich, G. (2013). *Educational testing and measurement: Classroom application and practice* (10th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Kwangmuang, P., Jarutkamolpong, S., Sangboonraung, W., & Daungtod, S. (2021). The development of learning innovation to enhance higher order thinking skills for students in Thailand junior high schools. *Journal of humanities and social science*, *7*(6). https://doi.org/10.1016/j.heliyon.2021.e07309.

Lawe, Y. U., Laksana, D. N. L., Wewe, M., Rawa. N. R., & Kua. M. Y. (2020). Pelatihan soal pengembangan matematika dan ipa berbasis higher order thinking dalam upaya meningkatkan kompetensi pedagogik guru SD di kecamatan golewa barat. *Jurnal Abdimas Ilmiah Citra*, *1* (1) Diambil dari: https://doi.org/10.38048/jailcb.v1i1.74

Marshall, J. C., & Horton, R. M. (2011). The relationship of teacher faciliated, inquiry-based instruction to student higher order thinking. *School Science & Mathematics, 111*(3), 93-101. https://doi.org/10.1111/j.1949-8594.2010.00066.x.

Nurkholik, M., & Yonata, B. (2020). Implementasi model pembelajaran inkuiri untuk melatihkan higher order thinking skills siswa pada materi laju reaksi kelas XI IPA MAN 2 Gresik. *Jurnal Pendidikan Kimia, 9*(1), 158-164. https://doi.org/10.26740/ujced.v9n1.p158-164

OECD. (2019). *Result from programme for international students assessment (PISA).* Retrieved from https://www.oecd.org/pisa/publications/PISA2019_CN_IDN.pdf

Puspendik. (2019). *Laporan hasil ujian nasional*. Retrieved from *https://hasilun.puspendik.kemdikbud.go.id*

Saido, G. A. M., Siraj, S., Nordin, A. B., & Al-Amedy, O. S. (2015). Higher order thinking skills among secondary school students in science learning. *Malaysian Online Journal of Educational Management*, *3*(3), 16-30. Retrieved from https://eric.ed.gov/?id=EJ1085914

Salirawati, D., Permanasari, L., Purtadi, S., Nugraheni, A. R. E., & Dina. (2017). Pelatihan pengembangan soal HOT (Higher Order Thinking) sebagai peningkatan kompetensi pedagogik guru. INOTEKS *Jurnal Inovasi Ilmu Pengetahuan, Teknologi, dan Sains, 21*(1), 14-25. Retrieved from https://journal.uny.ac.id/index.php/inotek/article/view/13175.

Saputra, H. (2016). *Pengembangan mutu pendidikan menuju era global: Penguatan mutu pembelajaran dengan penerapan HOTS (higher order thinking skills).* Bandung: Smile's Publishing.

Sari, C. W., & Helsy, I. (2018). Analisis kemampuan tiga level representasi siswa pada konsep asam basa menggunakan kerangka DAC (definition, algorithmic, conceptual). *Jurnal Tadris Kimiya, 3*(2), 158-170. https://doi.org/10.15575/jtk.v3i2.3660

Setiadi, H. (2016). Pelaksanaan penilaian pada kurikulum 2013. *Jurnal Penelitian dan Evaluasi Pendidikan, 20*(2), 166-178. https://doi.org/10.21831/pep.v20i2.7173

Sheppard, K. 2006. High School Students' Understanding of Titrations andRelated Acid-Base Phenomena. *Chemistry Education Research and Practice* *7*(1), 32-45. https://doi.org/10.1039/B5RP90014J.

Sudijono, A. (2015). *Pengantar statistik pendidikan*. Jakarta: PT. Raja Grafindo Persada

Sugiyono, (2014). *Metode penelitian pendidikan pendekatan kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.

Thomas, A., & Thorne, G. (2009). How to Increase Higher Order Thinking. (Los Angeles: Center for Development and Learning). http://www.cdl.org/resourcelibrary/articles/HOT.php?type=subject&id=18.

Widhiyani, I. A. N. T., Sukajaya, I. N., & Suweken, G. (2019). Pengembangan soal higher order thingking skills untuk pengkategorian kemampuan pemecahan masalah geometri siswa SMP. *Jurnal Pendidikan dan Pembelajaran Matematika Indonesia*, *8*(2), 161 – 170. https://doi.org/10.23887/jppm.v8i2.2854

Yakina, Kurniati, T., & Fadhilah, R. (2017). Analisis kesulitan belajar siswa pada mata pelajaran kimia kelas X di SMA Negeri 1 Sungai Ambawang. *Jurnal Ilmiah Ar-Razi,* 5(2), 287-297. https://dx.doi.org/10.29406/arz.v5i2.641

Yen, T. S., & Halili, S. H. (2015). Effective teaching of higher order thinking skills (HOTS) in education. *The Online Journal of Distance Education and e-Learning* *3*(2) 41-47. Retrieved from https://tojdel.net/journals/tojdel/articles/v03i02/v03i02-04.pdf