



## The Development Higher Order Thinking Skill (Hots) As Questions In Chemistry Study (Solubility And Solubility Product Constant)

Faucut Sarah<sup>1</sup>, Ibnu Khaldun<sup>2</sup>, Abdul Gani<sup>\*2</sup>

<sup>1</sup> Postgraduate Program of Natural Science Education Department of Syiah Kuala University, Banda Aceh, Indonesia, 23111

<sup>2</sup> Chemistry Education Department in Faculty Teacher Training and Education of Syiah Kuala University, Banda Aceh, Indonesia, 23111

\*abdulgani051266@gmail.com

### Article history

Submission : 2021-01-04  
Revised : 2021-02-16  
Accepted : 2021-03-08

### Keywords

First keyword  
Second keyword  
Third keyword  
Fourth keyword  
Fifth keyword

### Abstract

The development of higher order thinking skill (HOTS) questions on chemistry subject takes to drive students critical thinking skills. This research purpose to analyze the quality of HOTS chemistry questions and also to observe the teachers and students responses. The research method used research & development. The sample was determined by random sampling technique in total 104 students from several high school on Banda Aceh. Research procedures are collecting data and informations, product design, product development, validation by experts, initial product revisions, small-scale trials, product revisions, large-scale trials, and final product revisions. The data analysis technique was through calculating the percentage score of the assessment quality and analyzing the question items quantitatively using proanaltes program. The results showed that the quality of HOTS questions developed for Ksp (*solubility product constant*) lesson was 98.1%, means had very feasible criteria. In terms of quantitative analysis question items reached 95% valid, reliability test 0.740 high category, the level of difficulty includes 95% medium category and 65% difference good category. All calculation above plus the teachers and students positive responses attests the development of the WQC-based HOTS test instrument on Ksp lesson help improve students critical thinking skills significantly.



This work is licensed under  
a Creative Commons Attribution 4.0 International License  
©2021 Jurnal Pendidikan Sains Universitas Muhammadiyah Semarang

## 1. INTRODUCTION

Implementation of Curriculum 2013 recent years in Indonesia requires students emphasis learning on create, evaluate, and analyzing. To the educators, the curriculum focus on skills in developing HOTS assessment instruments, which is an evaluation tool that train students critical and creative thinking processes (Khaldun, et al., 2019). Furthermore, education in the 21st century is education that integrates knowledge, skills, attitudes and mastering technology (Phito, et al. 2019). The reasons bring it is necessary to prepare students think more critic and creative, doing good in teamwork, communicate well and understand computer literacy. This is consistent with the statement of Janssen et al. (2019).

The results of question items review conducted by the Directorate of Senior High School Development (2019), USBN assistance for 26 subjects in 136 referral schools across 34 provinces during 2018 to 2019, showed that of the 1,779 items analyzed, most were at Level- 1 and Level-2. From the 136 reference schools, there only 27 schools compiled high-order thinking skills questions as much as 20% of

all the USBN questions compiled, 84 are under 20%, and the rest still struggling on HOTS issues. Based on this research, it can be seen that the ability of educators in Indonesia comprehend the HOTS questions are still low, while the curriculum demanded more.

Critical thinking is an ability to analyze ideas logically, reflectively, systematically and productively to understand and evaluate information with the aim of whether the information is accepted, rejected or its judgment is suspended (Indah, 2020). Critical thinking also a major and important topic in modern education (Damayanti, et al., 2017). To measure the achievement of critical and creative thinking, an assessment is necessary. This assessment has an important role in determining the HOTS of students. HOTS is the ability to use the mind broadly to get something new and want someone to apply new information or previous knowledge in solving problems in dynamic situations (Kusumastuti, et al., 2019). According to Jainal and Louise (2019), HOTS in chemistry learning is one of the competencies that students must have. To find out a person's HOTS, indicators that can measure this ability are needed (Kurniati, et al., 2016). This is where HOTS questions items most critical acting devices.

According to Ghani, et al. (2017), the use of HOTS questions in exams can spur on students to think deeply how to answer the questions in learning and help increase students learning motivation, because HOTS questions bring stimulus that come from surrounding environment. Lee and Choi (2017) also argue that the HOTS issue has an important role in the education success, in academics and also in the workplace. It because HOTS makes students better prepared for challenges and more creative in solving problems. In addition, HOTS questions can also improve the achievement of student learning outcomes so that they can be competitive both nationally and internationally (Fanani, 2018).

Based on observations and interviews with high school chemistry teachers in several region in Banda Aceh, it was found that some teachers had never compiled the HOTS category questions in daily tests or semester exams. In the implementation of learning evaluation in class, the teacher only provides evaluations in the form of questions containing the three lowest cognitive levels, it is C1, C2 and C3. The practice questions given to students still use paper and rarely use computer facilities. It also found that the solubility material and the product of the solubility were still difficult for students to understand properly. The average completeness value of students in this material is still less than optimal, proved by their score was under the specified minimum completeness criteria (KKM), 80 point. Based on the fact, it is essential to change the system in the assessment. The instrument developed by the teacher is expected more improve students higher order thinking.

One of the efforts that stakeholder can take of existing computer facilities in the evaluation process requires the right software, which is Wondershare Quiz Creator (WQC). WQC is software for creating questions, quizzes or online tests (web-based). The questions that have been created can be saved in flash format that can privately (stand alone) on the website. With WQC, users can arrange various forms and different levels of questions. Even with WQC, various images (images) and flash files (flash movie) can also be inserted to support students' understanding in working on questions (Hernawati, 2009). WQC is free software, it can be downloaded from the official website <https://downloads.tomsguide.com/Wondershare-QuizCreator,0301-31201.html>.

The results of the development of HOTS-based test instruments have been carried out by several researchers and have shown a positive effect. Research conducted by Kusuma, et al. (2017) show that the HOTS assessment instrument is an effective learning assessment to train HOTS and measure students thinking skills. More, Ramadhan, et al. (2019); Kurniawan and Lestari (2019), also reported in their research that the test instrument developed could be used to measure the HOTS of students. Also, the research by Kusumadani et al. (2016) showed that the HOTS test instrument was effective in measuring the HOTS of students and had good effectiveness.

Research by Sa'adah, et al. (2019) show that the development of HOTS questions with WQC as a display media on stoichiometric material is suitable for use. In addition, the responses of teachers and students to the development of HOTS questions with WQC on stoichiometric material were also high, it is 90.67% and 91.20% respectively with very good criteria. Maifajir, et al. (2019), stated that the questions developed were very feasible to be used as independent training for students to train HOTS, which had positive responses from teachers and students to HOTS questions on chemical bonding material using WQC. Another result, Hidayat et al., (2020) the development of HOTS assessments in Bahasa Indonesia subjects also showed positive responses from teachers and students. However, the use of HOTS-based test instruments is still minimal in developing chemical materials, it is important to develop it further.

## 2. METHOD

The research method used is *Research And Development (RnD)*. This research aims to produce HOTS-based test assessment instrument. The model used is the Borg & Gall development model. The sample in this study consisted of 104 students who were selected randomly from several schools in Banda Aceh, it is SMAN 4, SMAN 5, SMAN 8, and SMAN 11. The research procedure consisted of collecting data and information, product planning, product development, trials, revisions, field trials, and final product revisions. The research instrument consisted of validation sheets, question quality assessment sheets, media quality assessment sheets computer-based test, and student and teacher responses questionnaire sheets. The data analysis technique consisted of analyzing the quality of the instrument on the item, it is the validity, reliability, difficulty level, and differentiation tests that were tested using the Excel program version 6 proanaltes (Khaldun, 2017). Then the instrument feasibility test, result data analysis, teacher and student questionnaire analysis.

## 3. RESULTS AND DISCUSSION

### Development the quality of HOTS in chemistry questions for solubility and solubility products constant chapter

The validation process carried out by material and instrument experts. The process aims to determine the quality of HOTS in chemistry questions developed for the solubility and solubility product constant subject. The results of expert validation / material expert and WQC-based test instrument media can be seen in Table 1.

Table 1. The result analysis of HOTS assessment quality in chemistry questions by experts

No	Aspect	Percentage	Category
1	Material	100	High
2	Construction	98,4	High
3	Practice/ Language	96	High
Avarage score		98,1	High

Based on the table, it can be concluded that the results of the HOTS assessment of chemistry questions are in the high quality category. The results illustrate that HOTS chemistry questions validated by material experts are suitable to be used as a instrument test to improve high school students critical thinking skills. Previous research Sari, et al. (2018), the input results received from the validators reach of 4.0 so that the multiple choice high order thinking test instrument met the valid or qualified criteria. A question items called valid or has high validity when it can measure the expected competence.

Jannah & Pahlevi's research (2020) says that the results of the validator's assessment, obtained scores 98% by material experts, then 87% from the evaluation experts, and 92% by linguists, all of them are qualified in "very strong" criteria. Based on the results, the average value of expert validation was 92% with "very strong" criteria. A similar study was conducted by Khaldun, et al. (2019) that the feasibility results of the computer-based HOTS category in chemistry questions have been developed using WQC are 85% qualifies as "high valid" criteria.

The next progress was validating the WQC media. The validation process carried out by instrument experts to determine the validator's assessment of WQC quality. The results of the validation can be seen in Table 2.

Tabel 2. The result analysis of WQC assessment quality by experts

No	Aspect	Percentage	Criteria
1	Utilization	87,5	High
2	Regulation	93,8	High
3	Access Capability	87,5	High
4	Access Speed	100	High
Avarage		92,2	High

Conformance scores obtained from the results above shows that the media quality validation results gained an average of 92.2%. It is proves that the four aspects measured are in the very feasible category, with conclusion this WQC media are approved to be applied. According to research conducted by Adah and Qosyim (2019) that the results of the media validation as a whole obtained a percentage of 95.48 or >80, qualified as very good criteria.

Meanwhile, the results of the WQC-based HOTS questionnaire assessment results carried out by the teacher, can be seen more detail in Table 3.

Tabel 3. The result analysis of assessment HOTS question instrument questionnaire  
 WQC-based teachers views

No	Aspect	Percentage	Criteria
1	Saving time and effort	100	High
2	Functionality <i>web-based</i>	100	High
3	Ease of use	87,5	High
4	Proportional view	95,8	High
Rata-rata		95,8	High

Based on Table 3 results, it proves that the teacher assessment items are in the qualified category. The validity also supports that the WQC-based HOTS instrument will bring positive impact in chemistry learning.

### Results of Items Test in a Small Scale

The trial was carried out on a small sample consisting of 27 students with 25 item questions. The results can be seen in following table.

Table 4. The results of items test in a small scale

Small Scale	Validity Test				Reliability Test	Level of Difficulty			Discrimination Power Test			
	ST	T	R	SR		Su	Se	M	BS	B	C	K
	32	48	8	12	0,874	8	84	8	44	40	12	4

The result shows that the HOTS analysis items test in a small sample reach the high category validity test were 48% greater than the other categories. Then the reliability test results obtained 0.874 qualified in the high and reliable category. The level of difficulty test results the highest score in the medium category in 84 percentage. In accord with research of Khaldun et al. (2019) that items belonging the middle category are maintained and should be recorded in the question bank. The difficult and easy category questions must be corrected again and replaced with questions where some students are able to answer and the questions sentence should more complex so students use to takes more effort and think critically. Furthermore, the discrimination power reach highest percentage qualified as good and excellent category. It bring to the hypothesis that, the HOTS instrument test can also be used on a large sample. The difference power test carried out to determine the difference between students who have mastered the concept and who did not. Conform to the opinion expressed by Rahayu & Djazari (2016) that the calculation of distinguishing power is a measurement of the extent to which an item can differentiate between students who have mastered competence from students who have not. The distinguishing power can be seen by looking at the size of the discrimination index of the questions.

Based on the analysis of the items consisting of validity, reliability, difficulty level, and differential power tests in the early, 20 questions in the good category, 4 discarded questions and 1 revised question were obtained. It can be concluded that of the 25 questions that were tested , there were 20 questions that were worthy of being tested to the final stage.

### Results of Items Test in a Large Sample

The large group trial was carried out in both groups, it is the experiment group and the control group. HOTS question instruments test in the experiment group using WQC media while the control group was done manually or without media. The test question instrument consisted of 20 multiple choice

questions designed according to HOTS criteria, it is analyzing (C4), evaluating (C5), and creating (C6). The large scale sample consisted of 104 students in the experiment group and 53 students in the control group. The difference in the number of samples taken was caused by the Ramadhan national holiday and the end of students semester exams. The test results of large sample items includes:

### 1. Validity Test

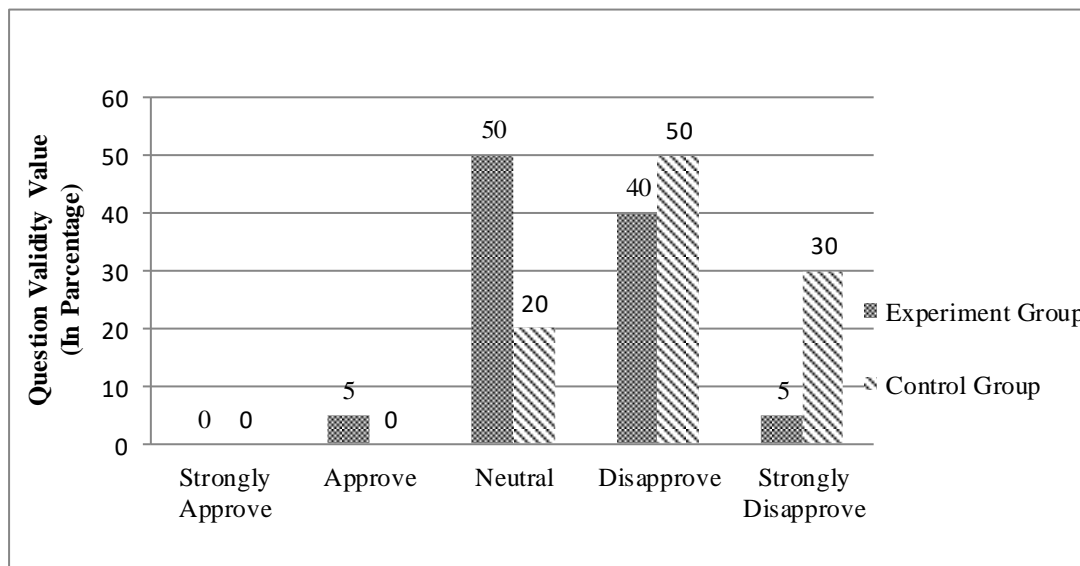


Chart 1. The results of validity test of the HOTS question instrument in large sample

The chart indicates that the experiment group reach the highest score in the neutral category 50% and the control group also shows similar score. The calculation in this validity test called the empirical validity test. In accord with the research by Riyani et al. (2017) that empirical validity contains the word "empirical" which means "experience". An instrument said it is qualified to use when it empirical validity has been tested.

### 2. Reliability test

The reliability test of HOTS instrument on a large scale shows that the experiment group with a coefficient of  $r_1$  of 0.704 is in the high and reliable category. While the reliability results in the control group with a coefficient  $r_1$  of 0.090 spot in the very poor and unreliable category.

### 3. Level of Difficulty Test

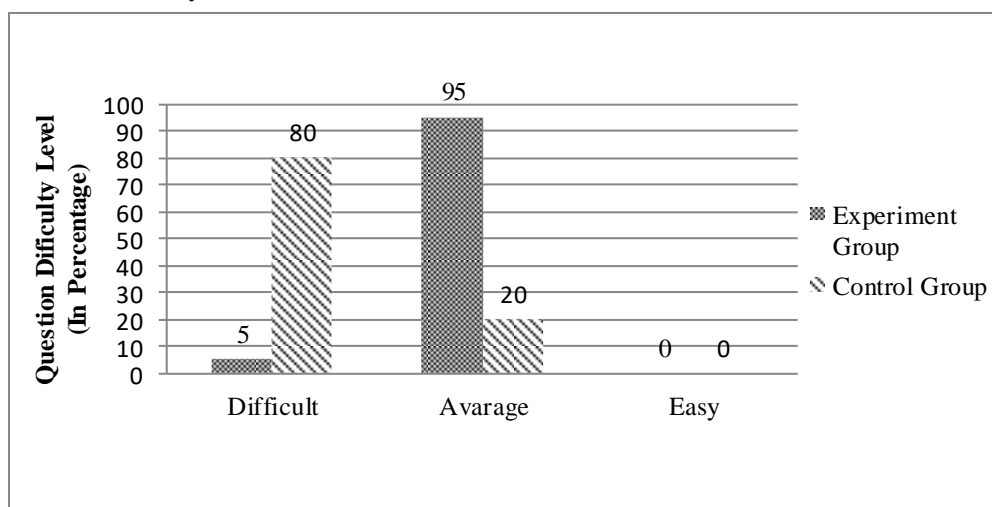


Chart 2. The results of HOTS questions difficulty level test in large sample

Found that questions in the experiment group are in the medium difficulty category while the control group in the difficult category. According to Solichin (2017) opinion, a question called qualified when the quality are in avarage, not too easy to answer and not difficult enough. Easy question do not stimulate students to enhance their efforts in solving. On the other hand, question that are too difficult will cause students to become discouraged and have no enthusiasm to try because they are out of reach. Rahayu & Djazari (2016) also stated that the level of difficulty calculation is proportion between students who can correctly answer an item with the total number of test participants. Rahmainsi & Taufiq (2018) says that the number shows the question easy or difficult called the difficulty index, which is in the formula symbolized by the letter P. The greater the question difficulty index, the easier question can be or the opposite.

#### 4. Differential Power Tests

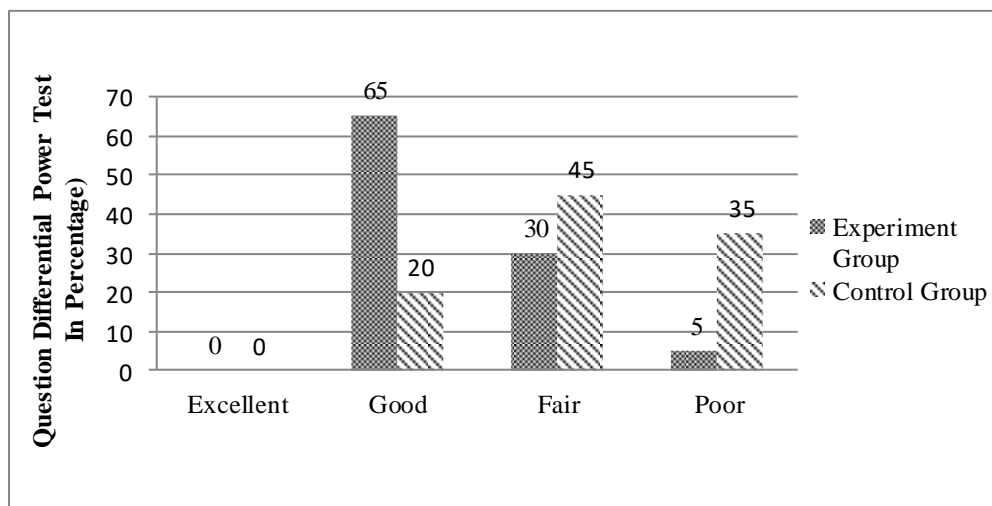


Chart 3. The results of HOTS differential question power in large sample

The differential question power in the experiment group are more good be compared to the control class score. The result proves that the questions that were tested in the experiment group have good differential power and the questions are more widely used. Stated by Susdelina, et al. (2018) the question differential power is the ability to distinguish between students who have a high level of ability and students who have low abilities. Supriadi, et al. (2018) the index of question differential power denoted by the letter D (stands for discriminatory power), and the index of question differential power ranges from -1.00 to +1.00. The higher the distinguishing power of a question, the better it get. Dewi, et al. (2018) said that the analysis of differential power is very important to do to determine the level of success of the questions in measuring students actual ability. When the question has bad distinguishing power, it not be able to measure the students ability, and when the question has good differential power, the better the question can measure.

#### **Analysis of Students Higher Level Thinking Ability Test Results**

Based on the results from the final trial in the experiment group and control group, students higher order thinking abilities can be measure. The results can be seen in Chart 4 and 5.

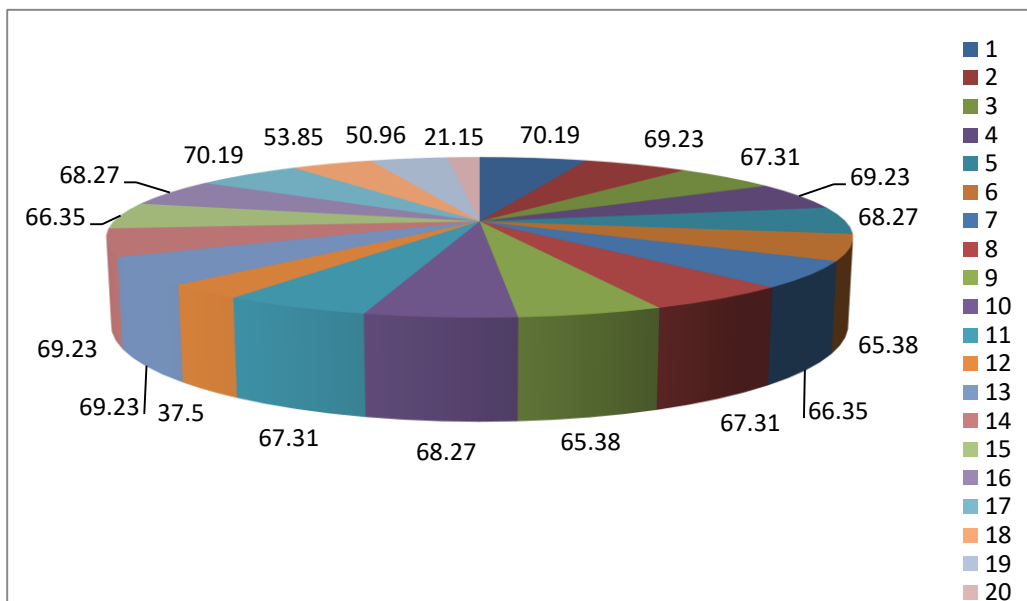


Chart 4. The result of high-order thinking skills percentage in the experiment group

The result of high-order thinking skills percentage analyzed on 20 HOTS instruments test in the experiment group gained 17 questions in good criteria, 2 items medium criteria and 1 item poor criteria. The good criteria range of values between 51-76, medium criteria 26-50, and poor criteria 1-25. The calculation proves that students thinking skills in the experiment class are qualified to the good criteria. Beside, the result of high-order thinking skills percentage in the control group can be seen in Image 5.

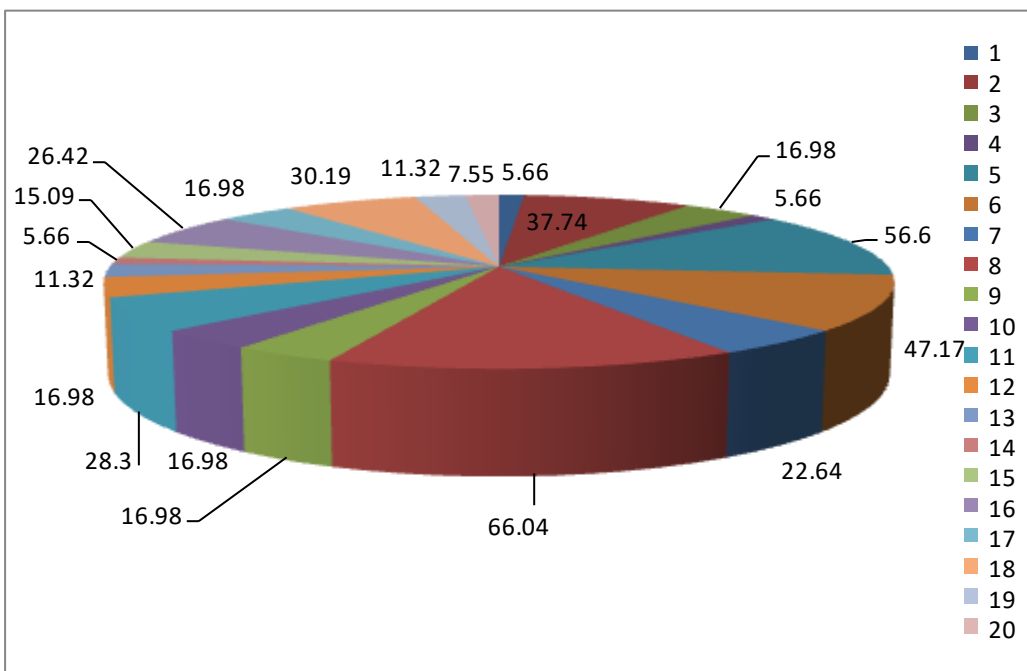


Chart 5. The result of high-order thinking skills percentage in the control group

The results of the percentage of students high-order thinking skills in the control class measured by 20 HOTS question instruments obtained 13 poor questions criteria 5 fair questions criteria and, 2 good questions criteria. The results prove that the high order thinking skills of students in the group are still lack of critic and creative.

**Analysis of Teachers and Students Responses to the development of the HOTS instrument test through the Wondershare Quiz Creator software**

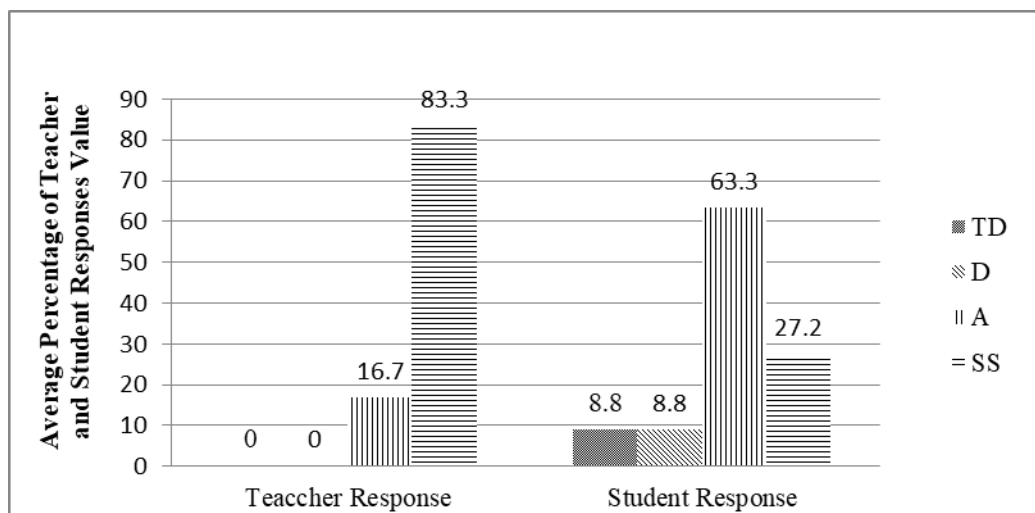


Chart 6. The average percentage of teacher and student responses on the development of the HOTS instrument test through the Wondershare Quiz Creator software

Based on the analysis, teacher responses to the development of the HOTS instrument question using the WQC are in the strongly agree and agree category. The analysis percentage value reached 16.7% for agree category and 83.3% for strongly agree category. The results of the student response analysis obtained percentage score of 27.2% strongly agree, 63.3% agree, 8.8% disagree and 0.7% strongly disagree. Regarding to the result, it proves that majority students support on the development of the WQC-based HOTS test instrument.

#### 4. CONCLUSION

Based on the research on the development of HOTS instruments test based on wondershare quiz creator, it can be concluded that the quality of HOTS developed in chemistry questions for solubility material and the solubility product in terms of qualitative analysis of items from the validation test results obtained an average score of 98, 1%, then the HOTS test questions are very viable to use. In terms of the quantitative analysis of the items, the validity of the questions was 95% valid and 5% invalid, the reliability test was 0.740 in the high category. The difficulty level of the questions covers 95% medium questions and 5% difficult questions. The difference power of the questions obtained for each category is 65% good, 30% fair and 5% poor. The response of teachers and students to the development of the computer-based HOTS question instrument developed using WQC reach positive responses. The results of the teacher's response with the category agree 16.7% and 83.3% strongly agree. While the response of students in the category of strongly agree 27.2%, 63.3% agree, 0.7% disagree, and 8.8% strongly disagree.

#### REFERENCES

- Adah, L. S., & Qosyim, A. (2019). Kelayakan tes HOTS berbasis komputer pada materi sistem peredaran darah manusia kelas VII. *Pensa E-Jurnal : Pendidikan Sains*, 7(3), 382–387.
- Damayanti, R. S., Suyatna, A., Warsono, W., & Rosidin, U. (2017). Development of Authentic Assessment instruments for Critical Thinking skills in Global Warming with a Scientific Approach. *International Journal of Science and Applied Science: Conference Series*, 2(1), 289.
- Dewi, S. S., Hariastuti, R. M., & Utami, A. U. (2019). Analisis Tingkat Kesukaran Dan Daya Pembeda Soal Olimpiade Matematika (Omi) Tingkat Smp Tahun 2018. *Transformasi : Jurnal Pendidikan Matematika Dan Matematika*, 3(1), 15–26.
- Supriadi, S.O.W., Rahim, U., & Zamsir. (2018). Kualitas tes sumatif mata pelajaran matematika kelas VIII semester genap SMP Negeri 20 Kendari tahun pelajaran 2016/2017. *Jurnal Pendidikan Matematika*, 6(3), 85-98.



- Fanani, M. Z. (2018). Strategi Pengembangan Soal HOTS dalam kurikulum 2013. *Edudeena: Journal of Islamic Religious Education*, 2(1), 57–76.
- Ghani, I. B. A., Ibrahim, N. H., Yahaya, N. A., & Surif, J. (2017). Enhancing students' HOTS in laboratory educational activity by using concept map as an alternative assessment tool. *Chemistry Education Research and Practice*, 18(4), 849–874.
- Hernawati, K. (2009). *Membuat Quiz / Evaluasi Dengan WonderShare Quiz*. 1(November), 1–12.
- Hidayat, D. M. F., Adisaputera, A., & Pramuniati, I. (2020). Development of HOTS (High Order Thinking Skill) based news text assessment instrument for 8th grade students in SMP Muhammadiyah 7 Medan. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, 3(2), 1123–1136.
- Indah, P. (2020). Development of HOTS (High Order Thinking Skill) Oriented Learning Through Discovery Learning Model to Increase The Critical Thinking Skill of High School Students. *International Journal of Chemistry Education Research*, 3(3).
- Jainal, S., & Yosephine Louise, I. S. (2019). Macromedia flash based on guided inquiry in critical thinking skills as learning innovations. *International Journal on New Trends in Education and Their Implications*, 10(3), 21–29.
- Jannah, K., & Pahlevi, T. (2020). Pengembangan instrumen penilaian berbasis higher order thinking skills berbantuan aplikasi " Kahoot!" pada kompetensi dasar menerapkan penanganan surat masuk dan surat keluar jurusan PTKP di SMK Negeri 2 Buduran. *Jurnal Pendidikan Administrasi Perkantoran*, 8(2018), 108–122.
- Janssen, E. M., Meulendijks, W., Mainhard, T., Verkoeijen, P. P. J. L., Heijltjes, A. E. G., Peppen, L. M., & van Gog, T. (2019). Identifying characteristics associated with higher education teachers' Cognitive Reflection Test performance and their attitudes towards teaching critical thinking. *Teaching and Teacher Education*, 84, 139–149.
- Khalidun, I., Hanum, L., & Utami, S. D. (2019). Pengembangan Soal Kimia Higher Order Thinking Skills Berbasis Komputer Dengan Wondershare Quiz Creator Materi Hidrolisis Garam Dan Larutan Penyangga. *Jurnal Pendidikan Sains Indonesia (Indonesian Journal of Science Education)*, 7(2), 132–142.
- Kurniati, D., Harimukti, R., & Jamil, N. A. (2016). Kemampuan berpikir tingkat tinggi siswa SMP di Kabupaten Jember dalam menyelesaikan soal berstandar PISA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 142–155.
- Kusuma, M. D., Rosidin, U., Abdurrahman, A., & Suyatna, A. (2017). The Development of Higher Order Thinking Skill (Hots) Instrument Assessment In Physics Study. *IOSR Journal of Research & Method in Education (IOSRJRME)*, 07(01), 26–32.
- Kusumadani, A. I., Nugroho, R. I. O. T., & Pratiwi, A. I. (2016). *Pengembangan Tes Berpikir Tingkat Tinggi pada Mata Kuliah Mikrobiologi Berbantuan Komputer dengan Microsoft Office dan Visual Basic for Application*. *Bioedukasi*, 9,(21–26.
- Kusumastuti, I., Fauziati, E., Marmanto, S., Maret, U. S., & Java, C. (2019). Revealing teachers' beliefs of higher order thinking skills in teaching reading at junior high school. *ELLiC (English Language and Literature International Conference) Proceeding*, 3, 155–162.
- Lee, J., & Choi, H. (2017). What affects learner's higher-order thinking in technology-enhanced learning environments? The effects of learner factors. *Computers and Education*, 115(January), 143–152.
- Maifajir, R., Khalidun, I., & Maulana, I. (2019). Developing Higher Order Thinking Skill ( HOTS ) Questions on Chemical Bonding in Senior High. *International Journal of Innovation in Science and Mathematics*, 7(2), 117–125.
- Mujianto, S. (2017). Analisis daya beda soal. taraf kesukaran, butir tes, validitas butir tes, interpretasi hasil tes valliditas ramalan dalam evaluasi pendidikan. *Jurnal Manajemen Dan Pendidikan Islam* 2, 2(2), 192–213.
- Direktorat Pembinaan Sekolah Menengah Atas. (2019). *Modul Penyusunan Soal Higher Order Thinking Skill (HOTS) Kimia*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah Kementerian Pendidikan dan Kebudayaan.
- Phito, V., Arief, A., & Roza, M. (2019). Pengembangan instrumen asesmen higher order thinking skill (HOTS) dalam pembelajaran Fisika pada materi hukum newton kelas X SMA/MA. *Natural Science Journal*, 5, 787–799..
- Riyani, R., Maizora, S., dan Hanifah. 2017. Uji validitas pengembangan tes untuk mengukur kemampuan

- pemahaman relasional pada materi persamaan kuadrat siswa kelas VIII SMP. *Jurnal Penelitian Pembelajaran Matematika Sekolah*, 1(1):60-65.
- Rahmaini, A., & Taufiq, A. nur. (2018). Analisis butir soal pendidikan agama islam di SMK N 1 Sedayu tahun ajaran 2017/2018. *Jurnal Mudarisuna*, 8(1), 1–24.
- Ramadhan, S., Mardapi, D., Prasetyo, Z. K., & Utomo, H. B. (2019). The development of an instrument to measure the higher order thinking skill in physics. *European Journal of Educational Research*, 8(3), 743–751.
- Sa'adah, S. I., Rasmiwetti, R., & Linda, R. (2019). Pengembangan soal hots dengan wondershare quiz creator sebagai media display pada materi stoikiometri kelas X. *JTK (Jurnal Tadris Kimiya)*, 4(2), 177–188.
- Sari, D. R. U., Wahyuni, S., & Bachtiar, R. W. (2018). Pengembangan instrumen tes multiple choice high order thinking padapembelajaran fisika berbasis e-learning di SMA. *Jurnal Pembelajaran Fisika*, 7(1), 100.
- Susdelina, Perdana, A.S., & Febrian. (2018). Analisis kualitas instrumen pengukuran pemahaman konsep persamaan kuadrat melalui teori tes klasik dan rasch model. *Jurnal Kiprah*, 6(1): 41-48.