ISSN: 0215-9643 e-ISSN: 2442-8655

Development Of Assessment Instrument Based on Higher Order Thinking (Hot) Using Quizizz Application on The Subject of Reaction Rate for XI Grade High School Students

Radjawali Usman Rery[1,*], Anisa Wulan Sangputri[2], Sri Wilda Albeta[3]

Universitas Riau, Kampus Bina Widya KM. 12,5, Simpang Baru, Kota Pekanbaru, Riau 28293, Indonesia
[1]usmanrery1959@gmail.com, [2]anisawulansp@gmail.com, [3]wilda.albeta@lecturer.unri.ac.id*
**corresponding author*

ARTICLE INFO

Article history
Received 11, 3, 2023
Revised 8, 10, 2023
Accepted 6, 10, 2023

Keywords
Assessment Instrument
HOT
Quizizz, Reaction Rate
Plomp Model

ABSTRACT

This study aims to produce an assessment instrument based on HOT using Quizizz application about a valid reaction rate by the validator, as well as to determine the validity, reliability, level of difficulty, discriminatory power, and user response. This study uses research and development (R&D) design with the Plomp model. The subjects in this study were 2 chemistry teachers and 20 class XII students at SMAN 8 Pekanbaru and SMAN 2 Pekanbaru. Data collection techniques were carried out by interviews, literature studies and field studies. For data analysis techniques used validation by experts, as well as user trials. The validation results for the material validator obtained an average based on material aspects of 99.33%, construction aspects of 94.61% and language aspects of 95.16%. According to the media validator based on the substance aspect of the content is 97.77%, learning design is 91.25%, display (visual communication) is 96.67% and software utilization is 100%. The results of the purification of the item analysis data obtained that 15-item multiple choice HOT questions the valid criteria, had very high reliability criteria, 10 items were obtained in the "moderate" category and 5 items were obtained in the "difficult" category which had good discriminating power overall matter accepted. The user response score was 81.90% for teachers and 87.93% for students with good criteria. Therefore, the HOT-based assessment instrument helps the Quizizz application about class XI SMA/MA equivalent reaction rates which are developed to be valid according to validator material and media and get responses from both teachers and students.

I. Introduction

The 21st century learning system is a learning transition in which the currently developed curriculum requires educational institutions to change teacher-centered learning approaches to student-centered learning approaches. In simple terms, it can be interpreted as learning that provides 21st century skills to students, namely 4C which includes: (1) Communication (2) Collaboration, (3) Critical Thinking and problem solving, and (4) Creative and Innovative. So it takes learning that is oriented to higher order thinking (Larson & Miller, 2012). High-level thinking or what is also known as High Order Thinking (HOT) requires more complex thinking activities. As said by Budsankom et al. (2015) that the definition of HOT thinking involves the transformation of information and ideas. Higher Order Thinking Skills is defined including critical thinking, logic, reflective, metacognitive, and creative (Wang & Wang, 2014). All those skills will be active when someone faces an unusual problem, uncertainty, question and choice. The successful applying from these skills contained in explanation, decision, appearance and a valid product. The application is inappropriate with the context from the knowledge and experience as well as advanced developing or another intellectual ability.

The way that can be done to achieve the goal of HOT ability in students is that it re-quires the application of HOT-based learning. HOT-based learning is a learning interaction between students and teachers, or students and students who are oriented towards higher order thinking skills. Changwong et al. (2018) state that problem-solving requires the ability to think critically, and it is stressed here that the ability to analyze and create is at the heart of critical thinking. Thus, analytical skills must be mastered and sharpened by students. The success of developing higher order thinking in chemistry learning will also be determined by the assessment instruments used by the teacher in the classroom. If the assessment instrument used is high-order thinking, students will have a greater

opportunity to develop these abilities (FitzPatrick & Schulz, 2015). In addition, Barnett & Francis, (2012) states that the higher order thinking questions may encourage students to think deeply about the subject matter.

The results of interviews with researchers at SMAN 2 Pekanbaru and SMAN 8 Pekanba-ru with two chemistry teachers in December 2021, obtained information that teachers have implemented HOT learning as the Discovery Learning model or the Problem Based Learn-ing models can stimulate students because with this model students are given stimulus or problem that requires students' high-order thinking skills. In addition to using HOT learn-ing models, the teacher must also provide an evaluation in the form of HOT questions, but the HOT questions at SMAN 8 Pekanbaru and SMAN 2 Pekanbaru which are applied to the Semester Final Examination in the question of reaction rate only include one HOT question. This is because the preparation of HOT ques-tions is not an easy thing. The resulting study from Jensen et al. (2014) showed that in writing the test level of HOT is a challenging task for teachers, and it needs to be improved because it really will help students in obtaining the deep understanding toward the materials thought. The statement above is same with the argument from (Ong et al., 2016) which state that how essential of teachers'role in helping students to build their scientific ideas and their reflective thinking skill. Therefore, it is necessary to develop HOT questions which can be used to train students' higher order thinking skills. Based on the analysis of assessment instruments in schools, it shows that there is still a lack of higher-order thinking questions.

The material chosen for the development of this HOT-based assessment instrument is reaction rate. Reaction rate is a complex material where simple concepts are needed to build these complex concepts. So, it takes a thought process that is more than just memorizing to understand the concepts of this reaction rate. Vong & Kaewurai, (2017) reveal that identify-ing and investigating are just a few of the keys to successful learning. These two parts are certainly part of the ability to analyze students. The majority of students are suspected of being unable to differentiate and relate the concepts of reaction rate and the factors. The results of interviews with two chemistry teachers obtained information that, the students' understanding of the reaction rate material was still not understood by the students, this could be seen from the average value of the students' Minimum Completeness Criteria (KKM) which was still relatively low, namely 75 on the material rate reaction.

The arrival of the Covid-19 pandemic also had an impact on the learning process so that it was carried out using an online system (e-learning). Along with the development of science and technology, this is very influential in the learning process. This is demonstrated by the development of online computer based evaluation questions. One of the technology based applications that teachers can easily use to conduct online evaluations is Quizizz. Through the Quizizz application, students are more enthusiastic about getting better at learn-ing, because this application is tournament-based so that students are motivated to become winners in tournaments. ICT-based tournaments using Quizizz can improve learning outcomes and motivation and students feel happy, excited and triggered to master the material so they can be the best in tournaments (Ju & Adam, 2019; Rahayu & Purnawarman, 2019; Zhao, 2019). Besides that, with the tournament, students feel happy and not bored in learning (Albeta et al., 2020).

There are several studies related to the development of HOT-based assessment instruments, especially in chemistry learning, namely research conducted by (Afriani et al., 2018; Firmansyah et al., 2018; Hutapea et al., 2020; Sa'adah et al., 2019). Among the four studies regarding the development of HOT-based assessment instruments, the similarities between the research above and this research are that they both test students' related abilities in solving HOT questions in chemistry subjects and use the research method used, namely a mixture of descriptive quantitative and qualitative. While the differences between the research above and this research are the research subjects, the development model and the formulation of the problem, there are slight differences. The research above also uses media, but in this study the media used is a more innovative medium so that it further increases the interest of students, namely using the Quizizz application.

Based on these various statements, the purpose of this study was to develop an assessment instrument based on Higher Order Thinking (HOT) that can be used to measure students' knowledge of higher order thinking on the cognitive aspects of the Reaction Rate material.

II. Method

The type of research used is research and development (R&D) using the Plomp development model which consists of several phases, namely the initial investigation phase, the design phase, the realization/construction phase, the validation phase, trials, and revisions as well as the implementation phase. This research was conducted only until the validation, trial, and revision phases. The research was conducted at the Chemistry Education Study Program, Faculty of Teacher Training and Education (FKIP), Riau University, Pekanbaru in the even semester of 2021/2022 and trials were carried out at SMAN 8 Pekanbaru and SMAN 2 Pekanbaru. The following describes the activities of each phase of the Plomp development model:

A. Initial Investigation Phase

This phase is carried out for information retrieval within the scope of product development. In this phase,

various analyzes were carried out, namely front-end analysis, students, competencies, and materials.

B. Design Phase

The design phase aims to prepare developed assessment instruments that meet the feasibility of HOT-based assessment instruments assisted by the quizizz application. The activities carried out at this design stage are (a) designing product prototypes, (b) designing assessment instruments, in the form of validation sheets and user response questionnaires.

C. Realization/Construction Phase

The realization/construction phase aims to produce prototypes and instruments as a realization of the design that has been designed. At this stage a prototype was produced as a realization of the results of the design questions including storyboard questions, grid questions, HOT-based questions assisted by the Quizizz application on Reaction Rate material for Class XI SMA/MA equivalent.

D. Validation, trial, and revision phases

This phase is carried out to get assessments and suggestions from the validator team and users on the prototypes that have been compiled.

The validity of the Assessment Instrument was assessed through validation activities with 4 validators, namely 2 material expert validators and 2 media experts. Based on the provisions of the University of Riau, the validator must at least have received a doctoral degree or a master's degree and at least have held the position of associate professor. Validation data uses a 1-5 scale assessment rubric.

Where P is the percentage of the validation score expressed in percent (%), where n is the total score obtained and N is the maximum total score. The results of the validation score percentage are converted into qualitative values as presented in Table 1.

Table 1. Criteria for the validity of the validator's assessment

Percentage (%)	Information
80.00 – 100	Good/Valid/Decent
60.00 – 79.99	Good Enough/Valid Enough/Decent Enough
50.00 – 59.99	Less Good / Less Valid / Less Feasible
0.00 – 49.99	Not Good/ Invalid/ Inadequate

[a.] (Riduwan, 2012)

The HOT-based assessment instrument which was declared valid by the validator was tested beforehand to obtain validity, reliability, difficulty level and discriminatory scores.

1) Test the validity of the items

$$rxy = \frac{N\sum XY - (\sum Y)}{\sqrt{[N\sum X^2 - (\sum X^2)][N\sum Y^2 - (\sum Y^2)]}}$$

[b.] (Arikunto, 2016)

Information:

r xy = The correlation coefficient sought
N = Number of students
X = Variable value X
Y = Value of variable Y

If the value of r count > r table, then the item is said to be valid.

2) Reliability Test

Calculated by the Spearman Brown formula as follows:

$$\text{ri} = \frac{2\text{rb}}{1 + \text{rb}}$$

Information:

r = internal reliability of all instruments
rb = correlation product of moments between the first and second splits

Table 2. Reliability Criteria

Reliability	Criteria
0.81–1.00	Very high
0.61– 0.80	Tall
0.41– 0.60	Enough
0.21– 0.40	Low
0.01– 0.20	Very low

[c.] (Arikunto, 2016)

3) Difficulty Level

With the formula: P = B/JS

Information:

P = Difficulty index
B = The total number of students who answered the questions correctly
JS = Total number of students

Table 3. Difficulty Rating Criteria

Difficulty Level	Criteria
0.00 – 0.30	Hard
0.31 – 0.70	Currently
0.71 – 1.00	Easy

[d.] (Arikunto, 2016)

4) Discriminating Power

The item's discriminating power was calculated using the following formula:

$$\text{DP} = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Information:

DP = Different index
J = Number of test takers
JA = Total number of upper group participants

JB = Total number of lower group participants
BA = Number of test takers who answered correctly in the upper group
BB = Number of test takers who answered correctly in the lower group
PA = Proportion of upper group participants who answered correctly
PB = Proportion of lower group participants who answered correctly

The research criteria are based on the criteria for discriminating power with a discriminating power value of ≥ 0.3, so the item is accepted; if 0.10 < discriminating power value < 0.29 then the question is revised; and the value of discriminating power <0.10 then the item is rejected (Surapranata, 2015).

After trying out the test items so that the questions match the characteristics of the questions, then one-on-one trials are carried out at SMAN 8 Pekanbaru with 3 students who have high, medium, and low abilities. Then a limited trial (small group) was carried out on 10 students at SMAN 8 Pekanbaru and 10 students at SMAN 2 Pekanbaru. Selected students are considered students to have different abilities. Each student represents a group of students with high ability, medium ability group, and ability group low. This is meant to be able to identify product deficiencies from students with different thinking abilities. User responses were obtained by carrying out trial activities using instruments in the form of user response questionnaires.

Where P is the percentage of user validation/response scores expressed in percent (%), n is the total score obtained and N is the maximum total score. The results of the percentage of user response scores are converted into qualitative values as presented in Table 4 for user response criteria.

Table 4. User Response Criteria

Percentage (%)	Information
80.00 – 100	Very good
60.00 – 79.99	Well
50.00 – 59.99	Not good
0.00 – 49.99	Not good

III. Results and Discussion

A. Initial Investigation Phase (Preliminary Investigation)

At this stage, several analyzes were carried out, including analysis of the front end, students, competencies, and materials. The results of the pre-research findings and detailed literature review results can be seen in Table.

Activity	Findings
Interview with two chemistry teachers	The HOT questions have been applied to the Semester Final Examination but only one HOT question is included in the reaction rate question. This is because the preparation of HOT questions is not an easy thing. Lack of teacher skills in compiling HOT questions so that the questions made are still in the C1-C3 category and only one question can be said to be C4. Therefore, it cannot be used to train students' higher-order thinking skills. The questions in schools are not only developed by the teachers themselves but also come from teaching materials, textbooks, the internet or practice questions which have not been able to improve students' higher order thinking. Indicators of learning that have not been completed in the matter of reaction rate, namely in the indicator of calculating the pH value of the reaction order. Applications in providing evaluations to students used by teachers during online learning have not fully utilized technology-based applications that can increase student motivation to work on evaluations.
Distribution of questionnaires at SMAN 8 Pekanbaru, and SMAN 2 Pekanbaru	The dissemination of school questionnaire data shows that, as much as 70% of students are not well acquainted with questions with higher order thinking skills. (Singaravelu, 2017) revealed that students' analytical skills are still weak, so it is suggested that teachers can emphasize more problem-based exercises and develop students' problem-solving.
Results of literature review	Based on the problems that have been obtained from interviews with two chemistry teachers, a HOT-based evaluation is needed on the reaction rate material. The ability to analyze is a logical basis that relates the picture, nature, and relationships of each section, which can be separately identified into a new entity. The majority of students are suspected of being unable to differentiate and relate the concepts of reaction rate factors (Politsinsky et al., 2015).

B. Design Phase (Design)

At this stage the results obtained were: (a) design of an assessment instrument based on HOT with the stages of preparing tests, test grids, and answer keys and scoring guidelines (b) initial design of the Quizizz application prototype an assessment instrument based on HOT (c) design of research instruments, namely sheets and grid sheets of material validation and response questionnaire media to teachers and students.

C. Realization or Construction Phase

At this stage a prototype of the developed assessment instrument was produced and a material and media validation sheet instrument as well as a response questionnaire to teachers and students.

Jurnal Ilmu Pendidikan (JIP)
Vol. 29, Issue 1, June 2023, pp. 16-24
ISSN: 0215-9643
e-ISSN: 2442-8655

D. Validation, Test, and Revision Phase (Evaluation, Test, and Revision)

At this stage, the results are: (a) the validation score of an assessment instrument based on HOT using Quizizz application about Reaction Rate (b) the validity, reliability, difficulty level, and distinguishing power scores (c) the percentage value of the user's response to the HOT-assisted assessment instrument Quizizz app about Reaction Rate. Validation activities were carried out by 4 validators which included lecturers from the Chemical Education Study Program at the University of Riau and the Chemistry Education Study Program at the Islamic University of Riau as material validators and two lecturers from the Department of Informatics at Sultan Syarif Kasim Riau State Islamic University as media validators.

Material validation is assessed based on 3 aspects, namely material aspects, construction aspects and language aspects. The product validation results obtained constructive suggestions from the material validator, including improving the writing of the questions, improving the editorial questions, formulating the questions, improving the concept of the material in the questions. Correction of errors in writing/typing in the items to avoid typos so that students do not have multiple interpretations of the questions. All questions developed are in accordance with the assessment component, but in the construction aspect it has a low percentage, this is because the questions are still not in the Hot category and the stimulus is still not appropriate. But the overall assessment of the 2 validators is 95.58% with a very valid category. The overall average score of the percentage of material validation based on material aspects.

Table 5. Results of validity Test by the Learning Material Experts

Aspect	Assessment component indicators	I-th validation	Category	II validation	Category
Material	Truth	60	Valid	99.33	Very Valid
construction	Suitability	71.33	Valid	87,87	Very Valid
	Legibility	50	Valid Enough	99.33	Very Valid
	Completeness	76,67	Valid	96.65	Very Valid
Language	Indonesian Rules	88	Very Valid	90.65	Very Valid
	Question Presentation	99.67	Very Valid	99.65	Very Valid
Total Score		445.67	Valid	573.55	Very Valid
Percentage (%)		74,27		95.58	

The average percentage of material validation was 95.58% based on the material, construction and language aspects according to the eligibility criteria (Riduwan, 2012)meet the very valid category. The percentage of the overall score from the material validation of an assessment instrument based on Higher Order Thinking (HOT) using Quizizz application about reaction rates based on aspects of material, construction, and language can be seen in the percentage diagram of the average score of the validity of various aspects by the material validator presented in Picture 1.

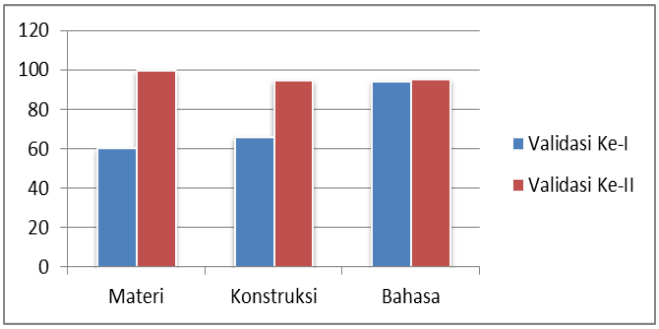

Fig. 1. Average Score of Material Validation from Various Aspects

Media validation is assessed based on 4 aspects, namely aspects of content substance, design, appearance (visual communication) and utilization. The results of the validation on the substance aspect of the content meet the very valid criteria of 97.77%, this means that Quizizz media has included a cover, complete instructions for use, has synchronized the terms in the instructions for use with the terms in the application, has feedback, and the identity of the compiler that has been according to the validator's suggestion. The validation results on the design aspect meet the very valid criteria of 91.25%, this means that Quizizz media has presented interesting covers and questions. The validation results on the display aspect (verbal communication) meet the very valid criteria of 96.67%, this means that the quizizz media is equipped with a setting time that functions properly in working on 15 items and a neat color appearance, so as not to interfere with readability. The results of the validation on aspects of software utilization meet the very valid criteria of 100%, this means that Quizizz media is easily accessible anytime and anywhere and can respond to user commands. The results of the assessment by the two media validators can be seen in Table 6.

Table 6. Validation Results by Media Validators

Assessment Aspects	I-th validation	Category	II validation	Category
Content Substance	70	Valid	97.77	Very Valid
Design	66,25	Valid	91.25	Very Valid

Assessment Aspects	I-th validation	Category	II validation	Category
Display (Visual Communication)	79,67	Valid	96.77	Very Valid
Software Utilization	100	Very Valid	100	Very Valid
Average total percentage	78.98	Valid	96.4225	Very Valid

The average percentage of media validation was 96.42% according to eligibility criteria (Riduwan, 2012) declared very valid. The average percentage of media validation based on aspects of the substance of the content, design, display (verbal communication) and software utilization can be seen in Figure 2 diagram of the average percentage score of media validity.

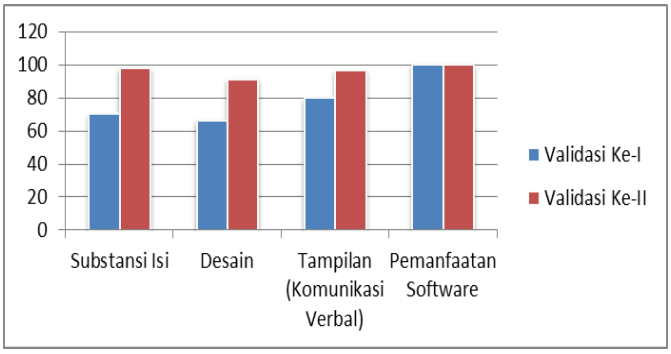

Fig. 2. Graph of Average Scores of Various Aspects of Media Validation

The next stage is the item test aimed at knowing the value of validity, reliability, level of difficulty and discriminating power. Based on the analysis of the validity of the 15 items, it was found that all questions were valid so that they were able to measure students' HOT abilities. Where all questions meet the criteria where if the value of r count > r table then the question is declared valid (Arikunto, 2016). Processing the validity test data obtained 15 items declared valid. It can be said that all items are positively correlated with the assessment instrument so that they can measure students' high-level thinking abilities. Test the validity of the 15 items seen in Table 7.

Table 7. Validity's Results of Item Questions

Question Items	r count	r table	Category
1	0.4616	0.4438	Valid
2	0.6205	0.4438	Valid
3	0.5743	0.4438	Valid
4	0.6741	0.4438	Valid
5	0.4786	0.4438	Valid
6	0.4547	0.4438	Valid
7	0.8825	0.4438	Valid
8	0.6422	0.4438	Valid
9	0.5920	0.4438	Valid
10	0.7506	0.4438	Valid
11	0.5526	0.4438	Valid
12	0.5953	0.4438	Valid
13	0.5099	0.4438	Valid
14	0.4762	0.4438	Valid
15	0.7538	0.4438	Valid

Data processing reliability value obtained a number of 0.91 fulfilling very high criteria, this means that this assessment instrument is reliable. The extent to which measurement results using the same object will produce the same data (Sugiyono, 2016). Processing the analysis of the level of difficulty of the data obtained about the difficult category and the medium category of 5 questions and 10 questions. The more difficult the questions, the lower the index of difficulty and vice versa. A good question is a question that is not too easy because it will not stimulate students to increase their efforts in solving problems (Wantoro et al., 2019). The difficulty index of question number 10 gets the smallest value because very few students answer correctly. The level of difficulty of the 15 questions can be seen in Table 8.

Table 8. Results of Calculation of the Level of Difficulty of the Question

Question Number	Difficulty Level	Criteria
1	0,65	Currently
2	0,55	Currently
3	0,5	Currently
4	0,6	Currently
5	0,5	Currently
6	0,5	Currently
7	0,3	Hard
8	0,35	Currently
9	0,15	Currently
10	0,25	Hard
11	0,3	Hard
12	0,45	Hard
13	0,55	Currently
14	0,2	Currently
15	0,35	Hard

Data processing analysis of the discriminating power of questions is carried out to determine the ability of the questions to distinguish students who have high abilities from students who have low abilities. Calculation of discriminating power analysis obtained all questions accepted and appropriate to use refer to (Surapranata, 2015), if the discriminating power value is ≥ 0.3 then the item is accepted. The value of the highest discriminating power is found in question number 9, because the higher the discriminating power of the item the more students from the high group can answer correctly and the fewer students from the lower group who answer correctly

(Wantoro et al., 2019). The calculation of the discriminating power of 15 items is shown in Table 9.

Table 9. Calculation of Item Distinguishing Power Calculations

Question Number	Difficulty Level	Criteria
1	0.5	Question accepted
2	0.5	Question accepted
3	0.4	Question accepted
4	0.6	Question accepted
5	0.4	Question accepted
6	0.2	Question accepted
7	0.6	Question accepted
8	0.5	Question accepted
9	0.7	Question accepted
10	0.3	Question accepted
11	0.3	Question accepted
12	0.4	Question accepted
13	0.3	Question accepted
14	0.7	Question accepted
15	0.4	Question accepted

After trying out the items, the questions match the characteristics of the questions, one-on-one trials and limited trials can be carried out. One-on-one trials were carried out with the aim of obtaining comments and suggestions as well as responses from the user's side of the assessment instrument. (Thaneerananon et al., 2016) emphasize that tests can help develop students' abilities in reflecting and developing critical thinking skills. After that, the student response questionnaire sheet was developed as a supporting instrument to determine student responses to HOTS items. One-on-one trials were carried out on 3 students with different abilities, namely high, medium, and low. Students must do quizizz and then to find out the students' responses, an interview is conducted. Students in working on quizizz require different times. The time the results of students' work are presented in Table 10.

Table 10. Time for Working on Assessment Instruments by Students

No	Name	Ability	Minute
1	Raihani Fakhruddin	High	55
2	Mufidah Mansour	Currently	69
3	Shafiq Abdullah H.	Low	78

Students during the one-on-one trials are difficult to work on the questions because these questions are not often encountered and are foreign to students and students are grade 12 students so that the material regarding reaction rates has not been studied by students for a long time. written by students on the paper provided, they are able to make creative solutions in their own language, this already fulfills one of the creative aspects, namely flexibility (Puspitasari et al., 2018). The value of working on the questions on Quizizz obtained the highest student scores, namely 12 questions correct in 55 minutes. The next stage is conducting interviews to provide suggestions and comments in the form of improvements to the discourse on question number 11 to simplify the sentences so that they are easier to understand so that the next trial stage can be carried out, namely testing the user's response to teachers and students. According to Abosalem, (2015) regarding assessment techniques in students' higher-order thinking ability which shows that using the HOTS assessment will assist students in deriving and evaluating it thinking skills such as using multiple choice tests or essay tests.

The results of one-on-one trials were obtained from students, the results of comments and suggestions as well as positive responses were obtained so that they could proceed to further trials, namely teacher response tests and small group trials. Teacher and small group response tests were carried out in order to find out the user's response to the assess-ment instrument which was developed based on aspects of attractiveness, effectiveness and practicality. The trial to the teacher obtained comments and suggestions in the form that the assessment instrument developed was based on HOT which was presented through quizizz which was easy to operate and understand and this instrument would train students' higher order thinking. The teacher's response questionnaire has 3 indicators in the form of indicators of effectiveness, attractiveness and practicality with an average percentage of the overall score of 81.9% with very good crite-ria. However, in statements number 8 and 10, they get a score of 3, because the position of the picture/table on Quizizz is located above the question, so it is not in accordance with the instructions on the question and students cannot work on the problem randomly. Student response tests obtained positive comments as seen from the questionnaire filled out by students, almost all of them agreed with an average percentage of 87.93% ful-filling the good category.

Research conducted by Firmansyah et al. (Firmansyah et al., 2018) and Sa'adah et al. (Sa'adah et al., 2019) both used media as display media, but in this study the media used was more innovative media so that it further increased students' interest, namely using the Quizizz app. According to Albeta et al. (Albeta et al., 2020), ICT-based Quizizz can be operated easily by students and has an attractive appearance so that students are excited while learning. Quizizz is an educational app that brings multiplayer activities and makes the app interactive and fun. The characteristics of the quizizz application game are in the form of themes, avatars, music and memes that are entertaining and spark the enthusiasm of students during the learning process (Mulyati & Evendi, 2020).

IV. Conclusion

The developed assessment instrument based HOT using Quizizz application about reaction rate for eleventh-grade senior high school students had 15 valid items, as

suggested from the validity test involving learning material and media experts. In the material validity test, the assessment instrument obtained scores of 99.33% for the content aspect, 94.61% for construction aspects, and 95.16% for language aspects. In the media validity, it obtained a 99.77% score for content substance aspects, 91.25% for learning design, 96,67% for display (visual communication), and 100% for software utilization. In the validity analysis, we concluded that 15 items had valid criteria, and very high reliability, with difficulty level for 10 items in the "moderate" category and 5 items in the "difficult" category, as well as good discriminating power. Thus, all items were accepted. The teacher's response suggested that the assessment instrument-based HOT had excellent effectiveness and attractiveness, with a percentage of 95.71% and 80%, respectively. Meanwhile, in the practicality indicator, the teacher felt this instrument was in a good category with a percentage of 70%. Further, the students also found that the effectiveness, attractiveness, and practicality of this instrument was in the very good category, with a score of 87.8%, 89%, and 87%, respectively.

References

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

Afriani, A., Setyarini, M., & Efkar, T. (2018). Pengembangan instrumen asesmen pengetahuan berbasis HOTS pada materi elektrolit dan non elektrolit. *Jurnal Pendidikan Dan Pembelajaran Kimia (JPPK)*, *7*(2), 1–12.

Albeta, S. W., Nofianti, N., & Rahmandani, S. (2020). Peranan turnamen berbasis ICT dengan aplikasi quizizz terhadap pembelajaran kimia. *Jurnal Pendidikan Kimia Universitas Riau*, *5*(1), 11–15.

Arikunto, S. (2016). *Metodologi penelitian suatu pendekatan proposal*. PT. Rineka Cipta.

Barnett, J. E., & Francis, A. L. (2012). Using higher order thinking questionsto foster critical thinking: a classroom study. *Ducational Psychology: An International Journal of Experimental Educational Psycholog*.

Budsankom, P., Sawangboon, T., Damrongpanit, S., & Chuensirimongkol, J. (2015). Factors affecting higher order thinking skills of students: A meta-analytic structural equation modeling study. *Academic Journal*, *10(19)*(2), 2639–2652. https://doi.org/10.-5897/ERR2015

Changwong, K., Sukkamart, A., & Sisan, B. (2018). Critical thinking skill development: Analysis of a new learning management model for Thai high schools. *Journal of International Studies*, *11*(2), 37–48. https://doi.org/10.14254/2071-8330.2018/11-2/3

Firmansyah, A., Hadiarti, D., & Sartika, R. P. (2018). Pengembangan instrumen penilaian (assesment) menggunakan wondershare quiz creator pada materi konsep MOL siswa kelas X SMK negeri 7 Pontianak. *Jurnal Ilmiah*, *4*(2), 1–23.

FitzPatrick, B., & Schulz, H. (2015). Do curriculum outcomes and assessment activities in science encourage higher order thinking? *Canadian Journal of Science, Mathematics and Technology Education*, *15(2)*, 136–154.

Hutapea, N., Sudrajat, A., & Situmorang, M. (2020). Pengembangan Instrumen penilaian berbasis kemampuan berpikir tingkat tinggi (HOTS) untuk pembelajaran sifat koligatif larutan pada siswa kelas XII IPA SMA Negeri 12 Medan. *Pros SemNas Peningkatan Mutu Pendidikan*, *1*(1), 470–501.

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test… or testing to teach: exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review*, *26*(2), 307–329.

Ju, S. Y., & Adam, Z. (2019). Implementing Quizizz as game based learning in the Arabic classroom. *European Journal of Social Science Education and Research*, *5*(1), 194–198. https://doi.org/10.2478/ejser-2018-0022

Larson, L. C., & Miller, T. N. (2012). 21st century skills: Prepare students for the future. *Kappa Delta Pi Record*, *47*(3), 121–123.

Mulyati, S., & Evendi, H. (2020). Pembelajaran matematika melalui media game quizizz untuk meningkatkan hasil belajar matematika SMP. *GAUSS: Jurnal Pendidikan Matematika*, *3*(1), 64–73. https://doi.org/10.30656/-gauss.v3i1.2127

Ong, K. K. A., Hart, C. E., & Chen, P. K. (2016). Promoting higher-order thinking through teacher questioning: a case study of a Singapore science classroom. *New Waves Educational Research & Development*, *19*(1), 1–19.

Politsinsky, E., Demenkova, L., & Medvedeva, O. (2015). Ways of students training aimed at analytical skills development while solving learning tasks. *Procedia - Social and Behavioral Sciences*, *206*(November), 383–387. https://doi.org/10.1016/j.sbspro.2015.10.-070

Puspitasari, L., In'am, A., & Syaifuddin, M. (2018). Analysis of students' creative thinking in solving arithmetic problems. *International Electronic Journal of Mathematics Education*, *14*(1), 49–60.

Rahayu, I. S. D., & Purnawarman, P. (2019). *The use of Quizizz in improving students' grammar understanding through self-assessment*. *254*(Conaplin 2018), 102–106. https://doi.org/10.2991/conaplin-18.2019.235

Riduwan, R. (2012). *Dasar-dasar statistika*. Alfabeta.

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. https://doi.org/10.15575/jtk.v4i2.5469

Singaravelu, S. (2017). Problem solving ability of higher secondary chemistry students. *IOSR Journal of Research & Method in Education (IOSR-JRME)*, *7*(4), 19–22. https://doi.org/10.9790/7388-0704041922

Sugiyono, S. (2016). *Statistika untuk penelitian*. Alfabeta.

Surapranata, S. (2015). *Analisis, Validitas, reliabilitas dan interpretasi hasil tes: Implementasi kurikulum 2004*. Remaja Rosdakarya.

Thaneerananon, T., Triampo, W., & Nokkaew, A. (2016). Development of a test to evaluate students' analytical thinking based on fact versus opinion differentiation. *International Journal of Instruction*, *9*(2), 123–138. https://doi.org/10.12973/iji.2016.929a

Vong, S. A., & Kaewurai, W. (2017). Instructional model development to enhance critical thinking and critical thinking teaching ability of trainee students at regional teaching training center in Takeo province, Cambodia. *Kasetsart Journal of Social Sciences*, *38*(1), 88–95. https://doi.org/10.1016/j.kjss.2016.05.002

Wang, S., & Wang, H. (2014). Teaching and learning higher-order thinking. *International Journal of Arts & Sciences*, *7(2)*, 179.

Wantoro, J., Sutama, S., Zuhriah, S., & Hafida, S. H. N. (2019). Pengembangan instrumen penilaian pendidikan profesi guru sekolah dasar bebasis Hots. *Profesi Pendidikan Dasar*, *1*(1), 11–20. https://doi.-org/10.23917/ppd.v1i1.8453

Zhao, F. (2019). Using quizizz to integrate fun multiplayer activity in the accounting classroom. *International Journal of Higher Education*, *8*(1), 37–43. https://-doi.org/10.5430/ijhe.v8n1p37.