

Development of Digital-based Minimum Competency Assessment (AKM) Type Test Instruments to Analyze Student Problem-Solving

Nurkintan Aprilia*, Sigit Priatmoko, Endang Susilaningih, Sudarmin Sudarmin, Noor Rifaan

Universitas Negeri Semarang, Indonesia

*Corresponding Author: nurkintanapriliala@students.unnes.ac.id

Abstract. The Covid-19 pandemic has given a new nuance to the world of education, learning which is usually done face-to-face has become based on digital technology. However, online learning has an unfavorable effect in improving students' literacy and numeracy skills. Both abilities are needed by students to be able to solve problems well. The purpose of this study was to develop a two-tier multiple choice complex test, Minimum Competency Assessment type, including digital-based literacy and numeracy indicators, for analyzing problem-solving students' level 11 in reaction rate topic. This research is an R&D research with ADDIE model. The test subjects 54 students of class XI High School 1 Welahan. The instrument has been tested for content validity, construct, and reliability. Content validity tested by six expert validators, with a score of 40.83 out of a total of 44, very valid. The construct validity, and reliability were tested using Rasch model analysis. The results show, from the 20 items developed, there are 3 items that are not valid, namely number 13, 12 and 19. The reliability test showed very good results, with index value of 0.91. Digital-based test instruments are considered more practical than paper-based tests, because they make it easier to assess and provide feedback. The results of the practicality test obtained as many as 92% (50 of 54) students gave a positive response. The existence of this test instrument makes it easier for teachers to train students' problem-solving processes and analyze problem-solving abilities more quickly through the analysis of the Rasch.

Key words: Digital Test; Minimum Competency Assessment and Student Problem-Solving

How to Cite: Aprilia, N., Priatmoko, S., Susilaningih, E., Sudarmin, Rifaan, N. (2022). Development of Digital-based Minimum Competency Assessment (AKM) Type Test Instruments to Analyze Student Problem-Solving. *ISET: International Conference on Science, Education and Technology* (2022), 167-172.

INTRODUCTION

The Covid-19 pandemic has given a new nuance to the world of education, learning that used to be done face-to-face has now become based on digital technology. All levels of education from elementary schools, Junior High Schools, and Senior High Schools must carry out digital-based learning online (Putri *et al.*, 2020). Many problems occurred during the COVID-19 pandemic. Distance learning creates gaps in learning (Madya & Abdurahman, 2020). This learning gap has an impact on students' literacy and numeracy abilities. Where literacy and numeracy skills are very important for students to have in the Covid-19 pandemic era, the two abilities include reasoning and argumentation skills, ability to solve problems, ability to provide perception and interpretation as well as the ability to improve insight related to life skills. in the face of global challenges (Nusantara *et al.*, 2020).

Problem-solving is inevitable in human life and is very important for human survival (Donnarumma *et al.*, 2016). Problem-solving is the most important skill needed by children today everywhere, including learning at school and outside school, namely its implementation in

everyday life. (Rahman, 2019). In solving problems, these two skills are needed, namely literacy and numeracy. Students who have high literacy and numeracy skills tend to be able to solve problems well, from simple to complex problems (Akyuz & Pala, 2010).

These two abilities will not appear if students are not used to working on PISA standard questions. PISA-standard questions can now be developed with a Minimum Competency Assessment (AKM) type question model that contains literacy and numeracy indicators (Delima *et al.*, 2022). The AKM model questions will train students' literacy and numeracy skills, as well as help teachers to analyze students' problem solving abilities (Fauziah *et al.*, 2021). Therefore, it is very necessary to have an AKM type question model in certain subjects such as Chemistry. The purpose of this study was to develop a test instrument in the form of a multiple choice complex two-level Minimum Competency Assessment (AKM) type, including digital-based literacy and numeracy indicators, which are useful for analyzing problem solving for high school students level 11 material reaction rates.

METHODS

This research is a Research and Development (RnD) research with ADDIE research design. The ADDIE research design includes five stages, namely Analysis, Design, Development, Implementation, and Evaluation. This ADDIE research design is one of the systematic learning design models and is based on the theoretical foundation of learning design in an effort to solve learning problems related to learning resources that are in accordance with the needs and characteristics of students (Widyastuti & Susiana, 2019). The study involved 54 11th grade students of SMA 1 Welahan. Prior to testing, the instrument was tested for content validity by 6 expert validators, the results showed the instrument was very valid with a score of 40.83 out of a total of 44. The construct validity and reliability were also tested using the Rasch model analysis. In addition, students' problem solving analysis skills were also analyzed using the Rasch model.

RESULTS AND DISCUSSION

Distinctiveness of the Two Tier Digital Test Complex (TTDTC) Instrument

This instrument is a combination of a diagnostic multiple choice instrument in the form of two-level multiple choice and a complex AKM type instrument that contains more than one correct answer at the first level (Mendikbud, 2020; Nahadi *et al.*, 2020). The instrument was developed through several stages including, the analysis of the chemistry syllabus for grade 11 curriculum 2013, and the preparation of the right grid, namely basic competency indicators, question indicators, problem solving indicators and minimum competency assessment indicators which include reading literacy and numeracy indicators (Siswaningsih *et al.*, 2019; Yunita & Dewi, 2020). Instruments are presented in digital form by utilizing the Google Form platform. Figure 1 presents an example of a developed instrument displayed via the Google Forms platform.

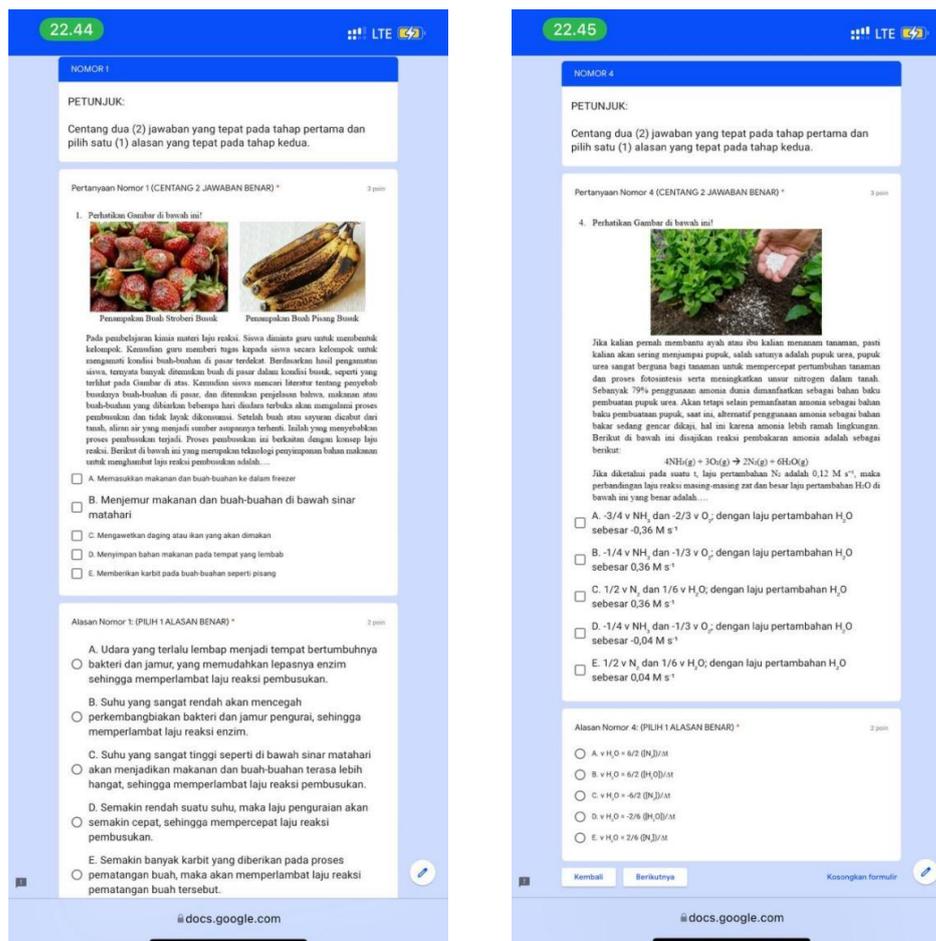


Figure 1. Display of TTDTC Questions on the Google Form Digital Platform

Figure 1 above shows that each of the questions from item 1 to item 20 are all separate questions in their respective sections. The advantage in a presentation like this is that students can focus on working on one question without panicking when they see other questions, students also don't feel confused in working because one question contains reading literacy which is quite long, when between one question and another question is not made a separate part it is feared that it will make students do not focus on work. However, the lack of a presentation like this is that students cannot work on questions that they think are easy to do first, but in answering them students must sequentially work from number 1 to number 20 without being able to pass the questions that they consider difficult.

This study experienced a few problems when collecting data, namely there was one student who experienced a trouble connection during the test, the student was asked to refresh the google form page. However, this obstacle was well

resolved, this is because the researchers had anticipated the obstacles that would occur. This google form system has the ability to save automatic answers. So that students' answers are not lost.

TTDTC Instrument Practicality Test

The practicality assessed in this study is about the practicality of using a digital test. Practicality indicators are included in the student and teacher response questionnaires. The practicality points listed in the student questionnaire include 6 points, namely, 1) the speed and efficiency of the digital test compared to the written test; 2) the practicality of the problem to work on; 3) the accuracy of the time allocation given; 4) the rapid process of assessing test results after completion of work; 5) easy access to the Google Form digital platform, both on laptops and cellphones; 6) paperless. The results of the practicality test based on the results of the student and teacher questionnaire responses are shown in Figure 2.

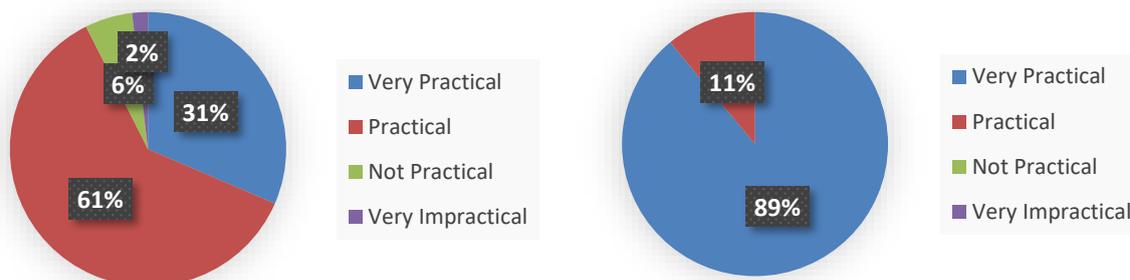


Figure 2. Results of Practical Recapitulation of Student and Teacher Response Questionnaires

It can be concluded that as many as 92% of students, namely 50 out of a total of 54 students, gave a positive response related to the practicality of the TTDTC test instrument developed, this is because the TTDTC instrument has made it easier for students in its implementation. Students can immediately see their scores when finished working and students can work anytime and anywhere regardless of place and time (Setiyoningtyas & Kasmui, 2020).

While the teacher response questionnaire, the results showed very good results, namely 89% as many as 8 of the total 9 chemistry teachers stated that the TTDTC instrument was very practical to use as a problem solving assessment tool, because in answering questions it was necessary to carry

out problem solving analysis and high concept understanding.

Summary Statistic Analysis

The Cronbach-alpha score gives us the reliability of the test, but does not tell us, if we have a problem with the student or the item. However, using the Rasch measurement model, it can inform us about student reliability and item reliability (Alagumalai *et al.*, 2005). Summary statistics provide information about reliability and dissociation index (Bichi *et al.*, 2019). Table 1 below presents the information obtained through the output summary statistics based on the analysis of the Rasch model.

Table 1. Output Results Summary Statistics Based on Rasch Model Analysis

Student Reliability Value	Criteria	Item Reliability Value	Criteria	Cronbach-Alpha	Criteria	Student Separation Index	Grain Separation Index
0.87	Good	0.90	Good	0.91	Very good	2.61	2.95

Table 1 above can be interpreted that, student reliability and item reliability get good results, meaning that the items developed are good in measuring student abilities, items can reach individuals with various levels of ability. Good results on student and item reliability analysis will certainly affect reliability results on cronbach-alpha, cronbach-alpha results get a very good reliability score.

The value of the separation index of students got good results, namely with a value of more than 2 (Mohamad et al., 2015). If the results are rounded up, each rounding approaches the number 3, meaning that there are 3 different groups that can be categorized, both in terms of

student ability and item difficulty level.

Item Suitability (Item-Fit) and Item Difficulty Level

Item-fit shows the number of valid items from the items developed. In the item-fit analysis there are several acceptance score criteria from MNSQ, ZSTD, and Pt Measure Corr, this acceptance score follows the rules of Boone *et al.* (2014). Item difficulty is seen from the item-measure output results, this value is indicated by the logit value, and can provide an overview of the relationship between students' abilities in answering the items developed.

Table 2. Result of Item Suitability (Item-Fit) and Item Difficulty Level

Information	Number of Item	Measure	Outfit MNSQ	ZSTD	Pt Measure Corr
The highest item difficulty level	15	1.58	1.59	1.72	0.44
The lowest item difficulty level	9	-1.90	0.47	-0.96	0.61
Unfit items (invalid)	13	-0.73	0.43	-2.13	0.79
	12	-0.39	2.17	3.19	0.37
	19	1.40	1.83	2.44	0.35

Table 2 provides information that there are 3 items that are not valid from a total of 20 items developed. The invalid items are items number 13, 12, and 19. There must be a reason why these items are invalid in this study. As with item number 13, the reason for the invalidity of item

number 13 may be caused by the information presented in the question not being coherent, so it is necessary to change the wording. Improvements to item number 13 are presented in Figure 3.

13. Pada masa pandemi Covid-19, masyarakat diberbagai negara disarankan untuk mengkonsumsi vitamin C. Vitamin C sangat bagus untuk meningkatkan daya tahan tubuh. Vitamin adalah zat yang dapat larut dalam air, sehingga apabila berlebih dan tidak dibutuhkan oleh tubuh maka akan larut dan dibuang melalui air seni. Berkaitan dengan hal ini, seorang siswa mencoba melakukan percobaan sederhana untuk mengetahui faktor yang mempengaruhi laju reaksi dengan cara melarutkan vitamin C dengan massa yang sama ke dalam air dengan beberapa perlakuan sebagai berikut.

Bahan	Gelas 1	Gelas 2	Gelas 3
Volume Air	30 mL	30 mL	30 mL
Vitamin Enervon C	Dalam bentuk bongkahan (utuh)	Dalam bentuk kepingan (diiris menjadi 4 bagian)	Dalam bentuk serbuk (sangat halus)

Prediksikanlah secara tepat apakah yang terjadi pada masing-masing gelas berdasarkan percobaan di atas dan jelaskan faktor yang mempengaruhi laju reaksinya!

- Reaksi yang terjadi pada gelas 3 paling cepat dibandingkan dengan gelas 1 dan 2
- Reaksi yang terjadi pada gelas 2 paling cepat dibandingkan dengan gelas 1 dan 3
- Reaksi yang terjadi pada gelas 1 paling cepat dibandingkan dengan gelas 2 dan 3
- Reaksi yang terjadi pada gelas 1 paling lambat dibandingkan dengan gelas 2 dan 3
- Reaksi yang terjadi pada gelas 3 paling lambat dibandingkan dengan gelas 1 dan 2

(a)

13. Vitamin adalah zat yang dapat larut dalam air, sehingga apabila berlebih dan tidak dibutuhkan oleh tubuh, maka vitamin akan larut dan dibuang melalui air seni. Saat masa pandemi Covid-19 yang terjadi pada tahun 2020, masyarakat diberbagai negara mulai mengkonsumsi vitamin C. Vitamin C sangat baik untuk meningkatkan daya tahan tubuh. Berkaitan dengan hal ini, seorang siswa mencoba melakukan percobaan sederhana untuk mengetahui faktor yang mempengaruhi laju reaksi dengan cara melarutkan vitamin C dengan massa yang sama ke dalam air dengan beberapa perlakuan sebagai berikut.

Bahan	Gelas 1	Gelas 2	Gelas 3
Volume Air	30 mL	30 mL	30 mL
Vitamin Enervon C	 Dalam bentuk bongkahan (utuh)	 Dalam bentuk kepingan (dipotong lebih kecil)	 Dalam bentuk serbuk (halus)

Prediksikanlah secara tepat apakah yang terjadi pada masing-masing gelas berdasarkan percobaan di atas dan jelaskan faktor yang mempengaruhi laju reaksinya!

- Reaksi yang terjadi pada gelas 3 paling cepat dibandingkan dengan gelas 1 dan 2
- Reaksi yang terjadi pada gelas 2 paling cepat dibandingkan dengan gelas 1 dan 3
- Reaksi yang terjadi pada gelas 1 paling cepat dibandingkan dengan gelas 2 dan 3
- Reaksi yang terjadi pada gelas 1 paling lambat dibandingkan dengan gelas 2 dan 3
- Reaksi yang terjadi pada gelas 3 paling lambat dibandingkan dengan gelas 1 dan 2

(b)

Figure 3. Repair Item Number 13, (a) Before Revision and (b) After Revision

Then items number 12 and 19 were detected as invalid in this test, possibly because the two questions were not able to measure normally on more test subjects. Because in the previous test, with fewer test subjects, these items were included in the category of valid items, so that these two items also needed to be revised so that they became valid and could take measurements normally. The improvements made were related to the reading text information presented in the questions, while the answer choices and reasons did not change.

Table 2 also provides information on the hardest and easiest items in this study, the hardest item is item number 15 with a logit value of +1.58 logit, and the easiest item is item number 9 with a logit value of -1.90 logit. These two items are in

the valid category because only the MNSQ value does not meet the criteria of Boone *et al.* (2014).

Results of Students' Problem-Solving Ability Through Person-Measure

The average value of logit person in this study is 0.59, with a standard deviation of 1.19. The two values are the benchmark for determining the category of student abilities, in this case the problem solving ability (Sumintono & Widhiarso, 2015). The category of students' problem-solving abilities based on the analysis of the Rasch model is divided into 3 categories, this division is based on the results of the analysis of the students' separation index in the statistical summary. The following Table 3 presents the results of the analysis of the level of student problem solving through person-measure analysis.

Table 3. Criteria For Grouping Students' Problem-Solving Abilities

Category of Student Problem-Solving Ability	Logit Measure	Student Number	Total students
High	Greater than 1.19	1-23	23
Medium	1.19 – (-0.59)	24-45	22
Low	Less than (-0.59)	46-54	9

Table 3 shows the overall results of students' abilities, the highest proportion is students with high abilities, as many as 23 students or by 43%. While students with low abilities have very few numbers when compared to other categories, which are only 9 students. Students who have the highest ability are students number 1, 2, and 3, with a logitability value of +2.43 logit, if it is associated with the logit value of the items on the item-measure, then the logitability value of the three students is much greater than the logit value of the most difficult item is item number 15 with a logit value of +1.58 logit.

This means, it is very likely that the three students can answer correctly all the questions presented with ease. While the student who has the lowest ability is student number 54 with a logit value of -2.42 logit. This logit value is lower than the logit value of the easiest item (number 9) on the item-measure, which is -1.90 logit, so it is most likely that student number 54 cannot answer all of the items presented correctly.

CONCLUSION

The results of the development of a two-level multiple-choice complex test instrument containing digital-based literacy and numeracy indicators show that out of the 20 items developed, there are 2 items that are invalid, namely items number 12 and 19. These invalid items are revised again before further testing is

carried out. While the reliability test showed very good results, with a Cronbach-alpha index value of 0.91. Digital-based test instruments are considered more practical than paper-based tests, this is because they are easier to assess and provide feedback and are environmentally friendly. The results of the practicality test based on the student response questionnaire obtained as many as 92% (50 out of 54) of students giving positive responses regarding the practicality of the instrument. The existence of this test instrument makes it easier for teachers to train students' problem-solving processes, and can analyze students' problem-solving abilities on the topic of grade 11 reaction rates more quickly through the analysis of the Rasch model.

ACKNOWLEDGEMENT

Thank you to the Postgraduate of the Universitas Negeri Semarang, for the facility of granting the 2022 UNNES DIPA Fund Grant for Postgraduate Student Thesis Research Grants with the DIPA Number: DIPA-023.17.2.677507/2022, November 17, 2021.

REFERENCES

- Akyuz, G., & Pala, N. M. (2010). The Effect of Student and Class Characteristics on Mathematics Literacy and Problem Solving in PISA 2003. *Elementary Education Online*, 9, 668–678.

- Alagumalai, S., Curtis, D. D., & Hungi, N. (2005). *Applied Rasch Measurement: A Book of Exemplars*. Dordrecht: Springer.
- Bichi, A. A., Talib, R., Atan, N. A., Ibrahim, H., & Yusof, S. M. (2019). Validation of a developed university placement test using classical test theory and Rasch measurement approach. *International Journal of Advanced and Applied Sciences*, 6, 22–29.
- Boone, W. J., Yale, M. S., & Staver, J. R. (2014). *Rasch analysis in the human sciences*. New York: Springer.
- Delima, N., Kurniasih, I., Tohari, Hutneriana, R., Amalia, F. N., & Arumanegara, E. (2022). *PISA DAN AKM: Literasi Matematika dan Kompetensi Numerasi*. Subang: Unsub Press.
- Donnarumma, F., Maisto, D., & Pezzulo, G. (2016). Problem Solving as Probabilistic Inference with Subgoalting: Explaining Human Successes and Pitfalls in the Tower of Hanoi. *PLoS Computational Biology*, 12, 1–30.
- Fauziah, A., Sobari, E. F. D., & Robandi, B. (2021). Analisis Pemahaman Guru Sekolah Menengah Pertama (SMP) Mengenai Asesmen Kompetensi Minimum (AKM). *Edukatif: Jurnal Ilmu Pendidikan*, 3, 1550–1558.
- Madya, S. A., & Abdurahman. (2020). Online Learning Implementation in the Covid-19 Pandemic. *Advances in Social Science, Education and Humanities Research, Atlantis Press*, 539, 26-31.
- Mendikbud. (2020). *Desain Pengembangan Soal Asesmen Kompetensi Minimum (AKM)*. Jakarta: Pusat Asesmen dan Pembelajaran Badan Penelitian dan Pengembangan dan Perbukuan Kementerian Pendidikan dan Kebudayaan.
- Mohamad, M. M., Sulaiman, N. L., Sern, L. C., & Salleh, K. M. (2015). Measuring the Validity and Reliability of Research Instruments. *Procedia - Social and Behavioral Sciences*, 204, 164–171.
- Nahadi, N., Siswaningsih, W., & Muchtar, H. K. (2020). Misconception Profile of High School Student on Electrolyte and Non-Electrolyte Solution Using Pictorial-Based Two-Tier Multiple Choices Diagnostic Test. *JKPK (Jurnal Kimia Dan Pendidikan Kimia)*, 5, 264–273.
- Nusantara, D. S., Zulkardi, & Putri, R. I. I. (2020). Designing PISA-like mathematics problem in covid-19 pandemic (PISAComat). *Journal of Physics: Conference Series*, 1657, 012057.
- Putri, R. S., Purwanto, A., Pramono, R., Asbari, M., Wijayanti, L. M., & Hyun, C. C. (2020). Impact of the COVID-19 Pandemic on Online Home Learning: An Explorative Study of Primary Schools in Indonesia. *International Journal of Advanced Science and Technology*, 29, 4809 – 4818.
- Rahman, M. M. (2019). 21st Century Skill “Problem Solving”: Defining the Concept. *Asian Journal of Interdisciplinary Research*, 2, 71–81.
- Siswaningsih, W., Nahadi, & Widasmara, R. (2019). Development of Three Tier Multiple Choice Diagnostic Test to Assess Students’ Misconception of Chemical Equilibrium. *Journal of Physics: Conference Series*, 1280, 032019.
- Setiyoningtyas, R., & Kasmui. (2020). Pengembangan Quizizz-Assisted Test Berbasis Literasi Peserta Didik Pada Materi Larutan Elektrolit Nonelektrolit. *Chemistry in Education*, 9, 1–7.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan (Rasch Modeling Application on Educational Assessment)* (1st ed.). Cimahi: Trim Komunikata.
- Widyastuti, E., & Susiana. (2019). Using the ADDIE model to develop learning material for actuarial mathematics. *Journal of Physics: Conference Series*, 1188, 012052.
- Yunita, & Dewi, N. R. (2020). The Development of Two-Tier Multiple Choice Assessment Instrument to Measure Higher Order Thinking Skills (HOTS) of The Students on Excretion System Material. *Unnes Science Education Journal*, 9, 30–39.