



ANALYSIS OF FEASIBILITY OF INTEGRATED ASSESSMENT INSTRUMENTS TO MEASURE CRITICAL THINKING SKILLS AND SCIENTIFIC ATTITUDES OF HIGH SCHOOL STUDENTS ON ACID-BASE TITRATION MATERIALS

Mir'atul Hasanatin¹,

Eli Rohaeti²ⁱ

¹Chemistry Education Study Program Graduate School,
Yogyakarta State University,
Indonesia

²Chemistry Education Study Program,
Yogyakarta State University,
Indonesia

Abstract:

This research aims to develop and determine the feasibility of an integrated assessment instrument to measure the critical thinking skill and scientific attitude of senior high school students on acid-base titration material. This research used development model 4-D with define, design, develop and disseminate stage. The product was assessed the legibility by 151 students and was assessed the feasibility by 7 chemistry teachers in Sleman Regency, Yogyakarta. The polytomous data generated by the instrument trial was analysed using the WINSTEPS 3.73 program and the Partial Credit Model 1-Parameter Logistic (PCM 1-PL) approach. The technique of analysing data consists of interviews, questionnaires and tests. Data collection analysis used the guidance of interview, questionnaire instrument user response, questionnaire of agility instrument, integrated assessment instrument in the form of description questions. The research result showed that the level of students' legibility has a reliability value 0.80, and the user instrument response has a reliability value 0.94. The teachers' response to the integrated assessment instrument was stated very good. Therefore, an integrated assessment instrument is feasible to be used to measure the critical thinking skill and science attitude of senior high school in on acid-base titration material.

Keywords: integrated assessment instrument; critical thinking skills; scientific attitudes; polytomous

ⁱ Correspondence: email mci.mitra@gmail.com

1. Introduction

Chemistry is one of the sciences included in the Natural Sciences (IPA) family. Agogo & Onda (2014) stated that chemistry has almost the same characteristics as science, so it can be considered as the centre of science as well as the mother of all sciences. Chemistry lesson includes an understanding of macroscopic, microscopic, shapes, and symbolic images, such as symbols, formulas, reaction equations, and graphs (Sirhan, 2007). Chemistry lesson aims to involve students in studying chemistry (Prins, 2010).

Chemistry is knowledge obtained and developed based on the experiment finding the answer to the question; what, why and how natural phenomena occur. Therefore, chemistry is a knowledge learning all things about substances including composition, structure, properties, transformations, dynamics, and energetics of substances that involve skills and reasoning (Mustofa, Pikoli & Suleman, 2013). There are two things about chemistry that cannot be separated, such as chemistry as a product (chemical knowledge in the form of facts, theories, principles, laws) scientific findings, and processes (scientific work) (Nasution, 2014).

One of the chemistries lessons learned by senior high school students is acid-base titration. An acid-base titration is one of the chemical materials that closely relate to daily life. However, students assume that the acid-base titration material classifies into difficult material due to many concepts and examples in acid-base titration material and all they learn is only memorizing the material. Aydm (2013) stated that the concept application in chemistry should not only be memorized from the theory, but they should be able to connect the theory with its application. Therefore, the teacher should have many strategies in learning including practising the students to think critically. Heng *et al.* (2012) confirm that science learning should cultivate scientific skills including process skills, creative thinking skills and critical thinking skills, and scientific attitude.

Critical thinking is creative thinking, problem-solving, data analysing, and communication (Stein *et al.*, 2007). Critical thinking skill is an ability requiring students to apply the information in a new situation and solve the problem. The process of thinking underlined needs a creative concept, analysing application, synthesis, and evaluation based on the information gathered from the observation and experience as guidance to take an action (Espinosa *et al.*, 2013).

In cultivating scientific skills also should have a scientific attitude. An attitude is someone's (students) behaviour in accepting a material, such as giving the idea or argument and even giving a reaction such as a question (Olasehinde & Olatoye, 2014). According to Pitafi & Farooq (2012) that a scientific attitude is someone's scientific action or thinking including curiosity, rationality, open mind, critical thinking, objectivity, honesty and humility. High curiosity will ask questions, read to find information, and immediately initiate and take action to lead to new discoveries.

Besides practising students' ability, the teacher needs to assess and evaluate. The thing that should be viewed and taken an attention in assessing is an instrument made

by the teacher. However, the question made still dominant used a short answer. Limited to knowledge, understanding, and application will make students' abilities only used to solve low-level cognitive problems (Sastrawati, Rusdi & Syamsurizal, 2011). Gardner (Murtono & Miskiyah, 2014) said that the assessment of short-answer tests at the low-level cognitive level only captures a small part of the skills and intelligence of students. Even though at this time, students are required not only to have low-level thinking skills but also to have higher-order thinking skills (Istiyono, Mardapi & Suparno, 2014).

Mardapi, Kumaidi & Kartowagiran (2011) stated that until today there are still many instruments of learning result both used by teachers for daily tests and used by the school for a general test that is still not meet the standard for a good test. Moreover, the teacher still prioritizes the assessment to the knowledge aspect while the assessment to the attitude is based on the teacher's subjective assumption. This case is in line with the statement of Zoller (2001) that the process of learning that is going on all this time only emphasize the cognitive aspect. Cantos *et al.* (2015) proposed that the assessment expected could reflect students' overall ability both in knowledge, attitude and skill side. This case is corresponding with the Minister of Education and Culture Regulation No. 66 of 2013 about the standard of educational assessment that stated the assessment in learning should include three aspects in learning such as knowledge, attitude and skill.

Based on the description, one of the alternatives that could be done by the teacher to measure students' critical thinking skills and scientific attitudes simultaneously is by integrated assessment. Integrated assessment is a process of students' learning results merging from various topics to be one series of efficient assessments (McPhun, 2010). The instrument of integrated assessment could help the teacher in doing the assessment to the students' learning result after following the learning.

2. Method

The type of this research is a Research and Development (R&D). The research method of Research and Development is a research method used to produce a certain product and examine the effectiveness of the product. This research model is a procedural model, i.e., descriptive research showing the steps that should be followed in producing a product. The development model in this research used development 4-D (four-D) which was developed by Thiagarajan, Semmel, & Semmel (1974). This model consists of four stages of development such as defining, designing, developing and disseminating (4-D). the development model was chosen since the sequence activities were designed and developed systems to create and develop the learning product such as assessment instrument.

This research and development were conducted at a senior high school in Sleman Regency, Yogyakarta, Indonesia. The legibility test of the integrated assessment instrument involved 151 students of class XI MIPA. Determination of the test subject for

the legibility of the integrated assessment instrument used the purposive sampling technique.

The type of data used is instrument feasibility data. Types of data collection instruments include interview guidelines, instrument legibility questionnaire sheets, instrument user response questionnaire sheets, and integrated assessment instruments in the form of description tests. The data analysis technique used is qualitative data analysis and quantitative data analysis.

The data were analysed by descriptive quantitative approach then it was converted to be a category. The product was assessed by the chemistry teacher by using Scale Linkert as presented in Table 1 below.

Table 1: Scoring Rule

No.	Category	Score
1	Very good	5
2	Good	4
3	Sufficient	3
4	Less	2
5	Very less	1

Furthermore, the final score was converted to be the feasibility of product as presented on Table 2.

Table 2: Ideal Assessment Criteria

No.	Interval Score	Category
1	$4,20 \leq$	Very Good
2	$3,40 \leq IQ < 4,20$	Good
3	$2,60 \leq IQ < 3,40$	Sufficient
4	$1,80 \leq IQ < 2,60$	Less
5	$< 1,80$	Very Less

After assessing the user's response to the product in the form of instrument feasibility by the chemistry teacher, the next step is to conduct a legibility test for students. The final score of the legibility test was converted into the eligibility category. The instrument legibility data was obtained from the students' responses which are shown in Table 3.

Table 3: Terms of Qualitative Score Conversion

No.	Answer Choice	Score
1	Very Agree	4
2	Agree	3
3	Disagree	2
4	Very Disagree	1

The next is the final score was converted into the feasibility category of product as presentind on Table 4.

Table 4: Score Conversion

No.	Interval Score	Category
1	$\geq 3,00$	Very Good
2	$3,00 > X \geq 2,50$	Good
3	$2,50 > X \geq 2,00$	Less
4	$< 2,00$	Not Good

3. Result and Discussion

3.1 Data Validity of Legibility Instrument Questionnaire

The validity of the legibility instrument questionnaire was obtained through assessment used the questionnaire validation of legibility instrument sheet conducted by material expert lecturers and learning evaluation expert lecturers. The study was carried out by giving an accuracy score for each item in the questionnaire statement. After obtaining the data from the analysis, the validity of the instrument legibility questionnaire was calculated by using equation 1 below.

$$A = \frac{\sum x}{n} \dots\dots\dots(1)$$

Where,

A = score of legibility questionnaire

$\sum x$ = the number of maximum scores

n = the number of questionnaire statement item

After calculating the average content validity by two experts, the final score of the validation result was converted into a feasibility category. Based on the analysis result of material experts and evaluation experts, it was concluded that the average content validity index of the instrument readability questionnaire was 3.87 with a maximum score 4.00 which was included in a very good category. The validation result of the instrument legibility questionnaire is presented in Table 5 as follows.

Table 5: Validation Result of Legibility Instrument Questionnaire

No.	Expert Judgment	Average	Category
1	Material Expert	3,80	Very Good
2	Evaluation Expert	3,93	Very Good
	Sum	3,87	Very Good

Note: Max. Score = 4,00

Based on Table 5, the result indicates that the instrument legibility questionnaire is feasible to be used for field trials. Other results obtained in the form of suggestions from expert judgment include the use of language, pictures, tables, graphs, materials and instrument displays. This case is done to make students do the questions well without any confusion and other disturbing factors. This case is in line with the opinion of Irwanto et al. (2017) that the instrument needs to be considered, especially in the use of developing language, so that students do not feel difficulties in understanding questions. This is important since language error is one of the factors that can affect test performance.

3.2 Data of Validation Result of Instrument User Response Questionnaire

The validity of the instrument user response questionnaire was obtained through an assessment using the instrument user response questionnaire validation sheet which was done by material expert lecturers and learning evaluation expert lecturers. The analysis was done by giving an accuracy score for each item in the questionnaire statement. After obtaining the data from the results of the analysis, then the validity of the instrument user response questionnaire was calculated using equation 1.

After calculating the average content validity by two experts, the final score of the validation result was converted into a feasibility category. Based on the analysis result of material experts and evaluation experts, it was concluded that the average validity index of the instrument user response questionnaire content was 3.88 with a maximum score of 4.00 which was included in a very good category. The results of the instrument user response questionnaire validation are presented in Table 6 as follows.

Table 6: Validation Results of Instrument User Response Questionnaire

No.	Expert Judgment	Average	Category
1	Material expert	3,79	Very good
2	Evaluation expert	3,96	Very good
	Sum	3,88	Very good

Note: Max. Score = 4,00

Based on Table 6, the result shows that the instrument user response questionnaire is feasible to be used for field trials.

3.3 Data of Instrument Legibility Result

The instrument feasibility level is also supported by the result of the legibility test questionnaire given to 151 students. The legibility test of the instrument was obtained through the questionnaire consisting of 15 questions. The legibility test aims to determine students' understanding of the language use, pictures, tables, graphs, materials, and instrument displays. The results of the instrument legibility test are presented in Table 7 as follows.

Table 7: The Result of Instrument Legibility Test

No.	Aspect	Score	Category
1	Didactic	3,23	Very Good
2	Construction	3,24	Very Good
3	Technical	3,27	Very Good
	Average	3,25	Very Good

Max. Score = 4,00

Table 7 shows that most of the students responded strongly agree to all aspects measured in the instrument readability questionnaire with didactic, construction, and technical requirements obtaining a very good category. These results indicate that the language use, pictures, tables, graphs, and the material presented in the instrument is easy for students to understand and it is feasible to be tested in the field.

3.4 Product Assessment by Chemistry Teacher

The product assessment was done by seven chemistry teachers. The aim is to get suggestions and input from the chemistry teacher to complete the question that has been developed. The instrument assessment by the reviewer includes five components of feasibility which were described into 24 criteria. The product feasibility component consists of substance, construction, language, validity, and practicality aspect which are adapted from the provisions of the Minister of Education and Culture Number 66 of 2013 concerning educational assessment standards. The user response data for the integrated assessment instrument is presented in Table 8 as follows.

Table 8: Product Feasibility Category

No.	Aspect	Score	Category
1	Substance	4,68	Very good
2	Construction	4,19	Good
3	Language	4,09	Good
4	Validity	4,19	Good
5	Practicability	4,31	Very good
	Average	4,29	Very good

Note: Max. Score = 5,00

Table 8 shows that most chemistry teachers give very well responses to all aspects measured in the instrument user response questionnaire. These results indicate that the integrated assessment instrument is feasible for field trials.

3.5 Estimation of Instrument Reliability

All research instruments, besides it, has met the requirements of logical validity by expert judgment, it also needs to have empirical reliability. Azwar (2011) states that reliability means how far the results of measurement have trustworthiness, reliability, consistency,

stability which can be trusted. The measurement results can be trusted if in several measurements of the same subject group, obtained the relative same result.

The instrument of this research includes user responses and an instrument readability questionnaire. The results of the reliability analysis using the Winsteps program are presented in Table 9, while the details are presented in Figure 1 and Figure 2.

Table 9: The Reliability of Research Instrument

No.	Instrument	Reliability	Category
1	Questionnaire of Response	0,94	Very High
2	Questionnaire of Legibility	0,80	High

The estimation of reliability of the research instrument presented on Table 9 shows that the response questionnaire and the instrument legibility questionnaire have sample reliability of 0.94 (very high) and 0.80 (high). Therefore, the instrument user response questionnaire and the instrument legibility questionnaire have met the requirements of empirical reliability.

4. Conclusion

Based on the result of data analysis and discussion, it can be concluded that the integrated assessment instrument development used model 4-D, the result showed that students' legibility level based on didactic, constructive and technical aspects have an average score 3.25 includes in very good category and also have the estimation of reliability 0.80 which classify as high. The chemistry teacher response to the integrated assessment instrument is very good with an average score 4.25 based on the substance, construction, language, validity and practicability. The estimated reliability of the instrument user response questionnaire is classified as very high with a value 0.94. Therefore, the integrated assessment instrument is appropriate to be used to measure senior high school students critical thinking skills and scientific attitude on acid-base titration material.

Conflict of Interest Statement

The authors declare no conflicts of interest.

About the Authors

Mir'atul Hasanatin is a graduate student in chemistry education, faculty of mathematics and natural sciences, Universitas Negeri Yogyakarta, Indonesia.

Eli Rohaeti is a Lecturer in Department of Chemistry Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta, Indonesia. Research interest in polymer chemistry and chemistry education. Orcid Id : orcid.org/0000-0002-0930-732X, Scopus Id: 57195958933

References

- Agogo, P. O., & Onda, M. O. (2014). Identification of student's perceived difficult concepts in senior secondary school chemistry in Oju local government area of Benue State, Nigeria. *Global Educational Research Journal*, 2(4), 44-49.
- Aydm, A. (2013). Representation of science process skill in the chemistry curricula for grades 10, 11, and 12/Turkey. *International Journal of Education and Practice*, 1(5), 51-63
- Azwar, S. (2011). Reliabilitas dan Validitas. Yogyakarta: Pustaka Pelajar [Indonesia]
- Cantos, A. E., Alday, M. G. K. A., Alog, K. J. A., Asi, K. J. G., Calacal, R. H. U., & Britiller, M. C. (2015). Changing learning needs of student nurses: input to the nursing curriculum. *Asia Pacific Journal of Multidisciplinary Research*, 3(3), 108-119. Retrieved from www.apjmr.com
- Espinosa, A. A., Monterola, S. L. C., & Punzalan, A. E. (2013). Career-oriented performance tasks in Chemistry: effects on students' Critical Thinking Skills. *Education Research International*. 1-10. doi: <http://dx.doi.org/10.1155/2013/834584>
- Heng, Y. C., Joo, C.E., Basri, A. A. M., Leng, H. H., Bari, N. A., Suleiman, R., Som, A. M., Mustafa, S., Mohamed, S. H. O, Yusof, Z. M., Yazid, Z., & Majid, Z. A. (2012). *Integrated curriculum for secondary school (curriculum specification: science form 2)*. Kuala Lumpur: Ministry of Education Malaysia
- Irwanto, Rohaeti, E., Lfx, E. W., & Suyanta, S. (2017). The development of an integrated assessment instrument for measuring analytical thinking and science process skills. *AIP Conference Proceedings*, 1847. <https://doi.org/10.1063/1.4983907>
- Istiyono, E., Mardapi, D., & Suparno. (2014). Pengembangan tes kemampuan berpikir tingkat tinggi fisika (PysTHOTS). *Jurnal Penelitian dan Evaluasi Pendidikan*, 18(1), 1-12.
- Mardapi, D., Kumaidi, & Kartowagiran, B. (2011). Pengembangan instrumen pengukur hasil belajar nirbias dan terskala baku. *Jurnal Penelitian dan Evaluasi Pendidikan*, 15(2), 326-341.
- McPhun, H. (2010). Integrated assessments-engaging ways to enhance learner outcomes. *Good Practice Publication Grant*, 1-9. Retrieved from <https://ako.ac.nz/assets/Knowledge-centre/Integrated-assessments-Engaging-ways-to-enhance-learner-outcomes/Integrated-assessments-Engaging-ways-to-enhance-learner-outcomes.pdf>
- Mendikbud. (2013). *Peraturan menteri pendidikan dan kebudayaan republik Indonesia nomor 66 tahun 2013 tentang standar penilaian pendidikan*. [Indonesia]
- Murtono & Miskiyah, E. (2014). Pengembangan instrumen evaluasi dengan teknik simulasi sebagai asesmen alternatif dalam pembelajaran fisika materi mekanika fluida SMA kelas XI. *Jurnal Inovasi dan Pembelajaran Fisika*, 1(1), 1-12. DOI: <https://doi.org/10.36706/jipf.v1i1.1041>

- Mustofa, Pikoli, M., & Suleman, N. (2013). Hubungan antara kemampuan berpikir formal dan kecerdasan visual-spasial dengan kemampuan menggambarkan bentuk molekul siswa kelas XI MAN model Gorontalo tahun ajaran 2010/2011. *Jurnal Entropi*, 8(1), 551-561. Retrieved from <https://ejurnal.ung.ac.id/index.php/JE/article/view/1164/950>
- Nasution, N. (2014). Pengaruh penerapan pembelajaran inquiry terbimbing menggunakan macromedia flash player untuk meningkatkan hasil belajar siswa pada pokok bahasan struktur atom. *Jurnal Pendidikan Kimia*, Universitas Negeri Medan.
- Olasehinde, K. J., & Olatoye, R. A. (2014). Scientific attitude, attitude to science and science achievement of senior secondary school students in katsina state, Nigeria. *Journal of Educational and Social Research*, 4(1), 445–452. doi: <https://doi.org/10.5901/jesr.2014.v4n1p445>
- Pitafi, A. I. & Farooq, M. (2012). Measurement of scientific attitude of secondary school students in Pakistan. *Academic Rearch International*, 2(2), 379–392. Retrieved from www.savap.org.pk%5Cnwww.journals.savap.org.pk
- Prins, G. T. (2010). *Teaching and learning of modelling in chemistry education authentic practices as contexts for learning*. Netherlands: Utrecht.
- Sastrawati, E., Rusdi, M., & Syamsurizal. (2011). Problem-based learning, strategi metakognisi, dan keterampilan berpikir tingkat tinggi siswa. *Jurnal Tekno-Pedagogi*, 1(2), 1-14.
- Sirhan, G. (2007). Learning difficulties in chemistry: an overview. *Journal of Turkish Science Education*, 4, Issue 2, 2-20.
- Stein, B., Haynes, A., Redding, M., Ennis, T., & Cecil, M. (2007). Assessing critical thinking in STEM and beyond. *Innovations in E-Learning, Instruction Technology, Assessment, and Engineering Education*, 79–82. doi: https://doi.org/10.1007/978-1-4020-6262-9_14
- Thiagarajan, S., Semmel, D., & Semmel, M. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Minneapolis: University of Minnesota. Retrieved from <https://files.eric.ed.gov/fulltext/ED090725.pdf>
- Zoller, U. (2001). Alternative assesment as (critical) means of facilitating HOCS promoting teaching and learning in chemistry education. *Chemistry Education Research and Practice in Europe*, 2(1), 9-17. doi: <https://doi.org/10.1039/B1RP90004H>

Appendix

TABLE 3.1 HASIL UJI KETERBACAAN INSTRUMEN				ZOU624WS.TXT Aug 15 11:39 2021				
INPUT: 151 Person		15 Item		REPORTED: 151 Person		15 Item 3 CATS WINSTEPS 3.73		
SUMMARY OF 151 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	48.7	15.0	1.35	.61	.96	-.3	.96	-.3
S.D.	4.4	.0	1.50	.06	.72	1.6	.74	1.6
MAX.	58.0	15.0	4.56	.77	4.59	5.1	4.65	4.9
MIN.	38.0	15.0	-2.49	.52	.03	-3.1	.03	-3.1
REAL RMSE	.67	TRUE SD	1.34	SEPARATION	2.01	Person	RELIABILITY	.80
MODEL RMSE	.61	TRUE SD	1.37	SEPARATION	2.25	Person	RELIABILITY	.83
S.E. OF Person MEAN = .12								
Person RAW SCORE-TO-MEASURE CORRELATION = .99								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .84								

Figure 1: Empirical Data Analysis of Instrument Readability Questionnaire Results

TABLE 3.1 HASIL RESPON PENGGUNA INSTRUMEN				ZOU225WS.TXT Aug 15 15:47 2021				
INPUT: 7 Person		24 Item		REPORTED: 7 Person		24 Item 3 CATS WINSTEPS 3.73		
SUMMARY OF 7 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	102.7	24.0	1.43	.47	1.00	.1	1.15	.0
S.D.	9.9	.0	2.07	.05	.19	.6	.70	.7
MAX.	115.0	24.0	4.07	.54	1.40	1.4	2.86	1.6
MIN.	89.0	24.0	-1.30	.43	.82	-.6	.78	-.7
REAL RMSE	.50	TRUE SD	2.01	SEPARATION	4.05	Person	RELIABILITY	.94
MODEL RMSE	.47	TRUE SD	2.02	SEPARATION	4.26	Person	RELIABILITY	.95
S.E. OF Person MEAN = .85								
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .95								

Figure 2: Data Analysis of Instrument User Response Questionnaire Test Results

Mir'atul Hasanatin, Eli Rohaeti
ANALYSIS OF FEASIBILITY OF INTEGRATED ASSESSMENT INSTRUMENTS
TO MEASURE CRITICAL THINKING SKILLS AND SCIENTIFIC ATTITUDES
OF HIGH SCHOOL STUDENTS ON ACID-BASE TITRATION MATERIALS

Creative Commons licensing terms

Author(s) will retain the copyright of their published articles agreeing that a Creative Commons Attribution 4.0 International License (CC BY 4.0) terms will be applied to their work. Under the terms of this license, no permission is required from the author(s) or publisher for members of the community to copy, distribute, transmit or adapt the article content, providing a proper, prominent and unambiguous attribution to the authors in a manner that makes clear that the materials are being reused under permission of a Creative Commons License. Views, opinions and conclusions expressed in this research article are views, opinions and conclusions of the author(s). Open Access Publishing Group and European Journal of Education Studies shall not be responsible or answerable for any loss, damage or liability caused in relation to/arising out of conflicts of interest, copyright violations and inappropriate or inaccurate use of any kind content related or integrated into the research work. All the published works are meeting the Open Access Publishing requirements and can be freely accessed, shared, modified, distributed and used in educational, commercial and non-commercial purposes under a [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](#).