



How are the Characteristics of Integrated Assessment Instruments to Measure Critical Thinking Skills and Scientific Attitudes of High School Students?

Mir'atul Hasanatin^{1*}, Eli Rohaeti²

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

² Chemistry Education Study Program, Yogyakarta State University, Indonesia

DOI: [10.29303/jppipa.v7iSpecialIssue.1082](https://doi.org/10.29303/jppipa.v7iSpecialIssue.1082)

Article Info

Received: November 1st, 2021

Revised: December 20th, 2021

Accepted: December 25th, 2021

Abstract: This research aimed to develop an integrated assessment instrument and to study the instrument's characteristics. This research used 4-D development model of which the stages included defining, designing, developing, and disseminating. The primary product is validated by expert judgment. This research involved 274 mathematics and natural science eleventh-grade students of Senior High Schools in Sleman district, Yogyakarta. The polytomous data generated by the instrument trial was analyzed using the WINSTEPS 3.73 program and the Partial Credit Model 1-Parameter Logistic (PCM 1-PL) approach. Data collecting techniques included interviews, questionnaires, and tests. Data collecting instruments included interview guidelines, a question item validation sheet, and the integrated assessment instrument an essay test. The result of the research shows that the integrated assessment instrument has 0.93 of Aiken validation value. All tryout test items were stated to fit with the PCM 1-PL based on the criteria for the lowest and the highest limit outfit MNSQ 0.5 and 1.5. Item reliability was 0.96 and person reliability was 0.71. The difficulty level of question items was good, because of the range from -1.05 to +1.01 logit. Therefore, the integrated assessment instrument is suitable to be applied to measure the student's critical thinking skills and scientific attitudes.

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

Citation: Hasanatin, M., & Rohaeti, E. (2021). How are the Characteristics of Integrated Assessment Instruments to Measure Critical Thinking Skills and Scientific Attitudes of High School Students? *Jurnal Penelitian Pendidikan IPA*, 7(SpecialIssue), 359–364. <https://doi.org/10.29303/jppipa.v7iSpecialIssue.1082>

Introduction

Assessment is the process of obtaining a numerical or quantitative description of the level of characteristics that a person has with certain rules according to Gronlund (Subali, 2019). With regard to learning activities, a teacher needs to take measurements before, during, and after the learning process (Subali, 2019). According to Cantos, et al., (2015), assessment is expected to be able to reflect the

learners' ability entirely in terms of knowledge, attitude, and skills.

Learners are required to have the necessary competencies to meet global demands. The competence is referred to as global competence. Global competence covers three aspects, namely empirical-based knowledge and skills, high-order cognitive and metacognitive skills, and global disposition, perspectives, and attitudes (Balistreri, et al., 2012). These three aspects are cognitive, affective, and

*Email: miratulhasanatin.2019@student.uny.ac.id

psychomotor aspects that are the product of learning. This is due to advances in science and technology.

Challenges and the rapid change and the exchange of knowledge lead to a shift in focus on educational institutions, from an initial focus on developing theoretical knowledge to developing thinking skills, such as critical thinking that are one of the main focuses in the world of education today (Stephenson & Sadler-McKninght, 2016; Qing, et al., 2010). Although it has been stated that improving learners' thinking skills is an important goal of education, but there seem to be many problems faced to achieve this goal (Aliakbaria & Sadeghdaghighi, 2013). Regarding Sanjaya (Mahanal, et.al., 2016), one of the issues is that learning that takes place in schools today is less demanding for learners to develop thinking skills. Students are required to memorize information, but memorization only dominates to master science without attitudes developing and thinking skills of students (Sudarmini, et al., 2015)

One of the thinking skills focused on recently is critical thinking skills. It is defined as creative thinking, problem-solving, data analysis, and communication (Stein, et al., 2007). It is also the abilities requiring learners to apply information in new situations and solve problems. A marked thought process requires creative concepts, application of analysis, synthesis, and evaluation based on information gathered from observation and experience as a guide to taking action (Espinosa, et al., 2013). Critical thinking skills are very important to be trained continuously, students are accustomed to problem-solving and finding solutions to these problems so that better learning outcomes are obtained and can overcome life problems according to the experience (Ismayawati, et al., 2016).

In addition to thinking skills in facing global demands, learners should also have a scientific attitude. Attitude is the behavior of a person (learners) in receiving material, such as giving ideas or opinions and even giving reactions in the form of questions (Olasehinde & Olatoye, 2014). According to Pitafi & Farooq (2012) says that a scientific attitude is an action or scientific thinking of a person that includes 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 so as to give rise to new discoveries. Purwanti & Manurung (Priska, et al., 2021) explained that scientific attitudes can encourage students to be actively involved in the learning process because of their curiosity. Students must have a scientific attitude consisting of curiosity, patience, openness, criticalness, objectivity, honest, humble and sensitivity to social life. In fact, the scientific attitude of learners in the learning process has not been properly

demonstrated, such as the lack of curiosity of learners. This is in line with research conducted by Dewantara, et al., (2018) that the scientific attitude score of learners in one of the high schools in Indonesia is still in the average range of 1.42 from the highest score of 3. The same was revealed by Ozden & Yenice (2014) that the scientific attitude score of learners is still at the average level of 139.7 of the highest score of 200. Therefore, teachers must think scientifically to know how the scientific attitude of learners.

Ross (Chebii, et al., 2012) stated that science is a process of understanding and finding solutions to a problem of natural symptoms that involves testing the solutions offered. Chemistry is a science that applies critical thinking skills in every material and also plays an important role in the development of students' critical thinking skills (Zhou, et al., 2013). Heng, et al., (2012) affirm that science learning should foster scientific skills that include process skills, creative thinking and critical thinking skills, and scientific attitudes.

Based on the description above, then one alternative to measuring critical thinking skills and scientific attitudes simultaneously is with integrated assessment. Integrated assessment is the process of combining learners' learning outcomes from various topics into a series of effective and efficient assessments (McPhun, 2010).

Method

The type of this study was research and development (R&D). A research and development method was a research method used to result in a particular product and test the effectiveness of the product. The research model was a procedural research model, which was descriptive research showing the steps to be followed in producing a product. The development model for the study employed a 4-D (development model) Four-D) developed by Thiagarajan, Semmel, and Semmel (1974). The model consisted of 4 stages of development called Defining, Designing, Developing, and Disseminating (4-D). The development model was chosen because it was systematically designed in order of activities and developed to create or develop learning products such as assessment instruments.

The research and development were conducted at a senior high school in Sleman Regency, Yogyakarta, Indonesia. A trial attempted an integrated assessment instrument involving 274 eleventh-grade students in a science major. The subject of the integrated assessment instrument trial used a stratified random sampling

technique. Trial sampling technique was based on school rankings (high, medium, and low) according to the results of the National Examination (UN) of chemical studies in the 2019/2020 academic year. The time allotted was 90 minutes.

The type of data used was characteristic data. Types of data collection instruments included interview guidelines, item validation sheets, and integrated assessment instruments in the form of description tests. The data analysis techniques used were qualitative and quantitative data analysis.

The qualitative analysis included consideration of the validity of contents and constructs by expert judgment. Aspects that were considered in qualitative analysis of items were between indicators of learning materials, indicators of critical thinking skills, and indicators of scientific attitudes. Quantitative analysis of data to determine the characteristics of instruments used in research as a whole based on empirical data.

The data was analyzed with a quantitative descriptive approach which was then converted into several categories. The initial data obtained was data from validation of items obtained from expert judgment. The result of the validation of the problem item used the score conversion rule as presented in Table 1.

Table 1. Score Conversion

Interval Score	Categories
≥ 3.00	Very Good
$3.00 > X \geq 2.50$	Good
$2.50 > X \geq 2.00$	Pretty Good
< 2.00	Not Good

Furthermore, empirical validity was carried out through instrument trials to learners to determine the characteristics of items. Validity was empirically proven by fit items on the partial credit model (PCM) using the Winsteps 3.73 program. Item fit was important information in the analysis using RASCH modeling. The item conformity index determined whether an item was functioning optimally and met the requirements as a good measurement tool or not. The items were stated to be fit with the model if they met the criteria of outfit means-square value, z-standard outfit, and point measure correlation (Bond & Fox, 2015).

According to Sumintono and Widhiarso (2015: 72) if the test item met one of the criteria for receiving the MNSQ, ZSTD, or Pt Mean Corr limits, then the items could be declared appropriate to the model and appropriateness of further analysis. The guidelines for assessing item conformity criteria according to Boone, et al., (2014) were as follows: (1) Outfit Mean Square (MNSQ) value received: $0.5 < MNSQ < 1.5$; (2) Z-

Standard Outfit Value (ZSTD) received: $-2.0 < ZSTD < +2.0$; (3) Point Measure Correlation value received: $0.4 < Pt\ Measure\ Corr < 0.85$.

The next stage was to find out the difficulty level of the items. The difficulty level aimed to find out whether the details of the item level were simple, moderate, or difficult. A good item was not too easy and also not too difficult. The difficulty level of each stage in the item should be complete and sequential. A good item had a difficulty index range of more than -2.0 or less than +2.0 logit. The items belonging to the difficulty level under the -2.0 logit were categorized as too easy, while the items belonging to the difficulty level above +2.0 logit were categorized as too difficult. According to Hambleton & Swaminathan (1985) and Baker (2001), an item was said to be good if it had a difficulty index ranging from -2.0 to +2.0 logit.

Results and Discussion

Data of Item Validity Results

Reviewing the items was conducted by assessing the first product to the peer reviewer. Peer reviewers include friends who conducted similar research. Peer reviewers may be also referred to as language correctors.

Furthermore, the items were assessed to expert judgments, which were a lecturer in learning materials, a learning evaluation expert, and three chemistry teachers who have been proficient in the preparation of items. The purpose of item validation is to determine the conformity between the items developed with learning indicators, critical thinking skills, and scientific attitudes. Indicators of critical thinking skills include learners being able to determine solutions, evaluate opinions, select opinions, express opinions, categorize and draw conclusions. Indicators of scientific attitudes include curiosity, open mind, objectivity, opinion, rationality, and recommendation. Good items are appropriate to use by Expert Judgment for field trials to learners.

The validity of the item was calculated using Aiken's V formula. The calculation results show that the highest and lowest validity are consecutively 1.00 and 0.80. The average validity index of test content is 0.93 with a very good category. Based on the analysis of the results of the study, experts concluded that there are 11 items with very good validity criteria (85%) and 2 items with good validity criteria (15%). The validation results of the problem item are shown in Table 2 as follows.

Table 2. The Results of ITEM Validity

Expert Judgment	Score	Category
Evaluation Expert	3.57	Very Good
Learning Material Expert	3.36	Very Good
Practicioners	3.55	Very Good
Average	3.49	Very Good

Max. Score = 4.00

Other results obtained in the form of advice from expert judgment include the use of language, technical writing, and considering the number of items with a predetermined allocation of time. It is done so that learners do the items well without any confusion and other disturbing factors. It is in line with the opinion of Irwanto, et al., (2017) that instruments need to be considered especially in the technical writing and use of evolving languages so that learners do not have difficulty in understanding items. It is important because sentence mistake is one of the factors that can affect test performance.

Data of Instrument Test Result

The instrument trial was conducted on 274 learners. The results of the trial were further analyzed using the partial credit model response theory of 1-logistic parameter by calibration using Winsteps to determine the problem item that is fit with the model that has been set.

Based on an analysis of 13 points about field trial results, all points of the problem were declared fit with the model with estimated sample reliability of 0.96. Estimation reliability based on the item separation index is very high. Therefore, all the problems in the final product are expressed empirically reliable.

The results of the item analysis by using Winsteps also resulted in an estimate of the reliability of the test of 0.71. Reliability estimates calculated based on tests are called person separation indexes. The results of the analysis showed that the reliability estimate based on the test separation index was relatively high. In addition, the analysis of the problem using the Winsteps program also produced an Alpha Cronbach coefficient based on classical test theory (CTT) of 0.73. The difficulty level of the problem item is relatively good because it is in the range of -1.05 to +1.01 logit. These conditions show that measurement activities provide relatively consistent results.

Integrated assessment instruments can train students' abilities to think openly and be scientific because it is not only based on the results when students take a test but can be seen from the process. This is in line with the opinion of Wilsa et al. (Ilfiana, et al. 2021) that something has to open mind, both conceptually and procedurally can develop students'

abilities, especially critical thinking skills. Someone who thinks critically when obtaining information first analyzes, observes, and evaluates before accepting or rejecting the information (Arisa, et al., 2021).

Conclusion

Based on the results of data analysis and discussion, it can be concluded that the development and characteristics of integrated assessment instruments using the 4-D model, the results show an Aiken validity value of 0.93. All test items (13 items) are declared *fit* with PCM 1-PL based on the criteria of the lowest and highest limit outfit *MNSQ* 0.5 and 1.5. Item reliability of 0.96 and test reliability of 0.71. The difficulty level of the problem item is relatively good because it is in the range of -1.05 to +1.01 logit. Therefore, integrated assessment instruments are appropriate to measure the critical thinking skills and scientific attitudes of learners.

References

- Aliakbari, M., & Sadeghdaghighi, A. (2013). Teachers' perception of the barriers to critical thinking. *Procedia - Social and Behavioral Sciences*, 70, 1-5. <https://doi.org/10.1016/j.sbspro.2013.01.031>
- Arisa, N.S., Khaldun, I., & Safrida. (2021). The effect of search, solve, create and share learning models to improve students' critical thinking skills on acid and basic titration materials. *Jurnal Penelitian Pendidikan IPA*, 7(2), 191-195.
- Baker, F. B. (2001). *The basic of item response theory*. New York: ERIC Clearinghouse on Assessment and Evaluation.
- Balistreri, S., Tony, F., Giacomo, D., Noisette, I., & Ptak, T. (2012). Global education: connections, concepts, and careers. *The College Board*, 33. Retrieved from: www.collegeboard.org
- Bond, T.G., & Fox, Ch.M. (2007). *Applying the rasch model: fundamental measurement in the human science (2nd Ed)*. New Jersey: Lawrence Erlbaum Associates.
- Boone, William. J., Staver, John. R., Yale, & Melissa. S. (2014). *Rasch analysis in the human sciences*. London: Springer. <https://doi.org/10.1007/978-94-007-6857-4>
- 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

- Chebbi, R., Wachanga, S., & Kiboss, J. (2012). Effects of science process skills mastery learning approach on student's acquisition of selected chemistry practical skills in school. *Creative Education*, 3(8), 1291-1296. <http://dx.doi.org/10.4236/ce.2012.38188>
- Dewantara, D., Hasan, S. M., & Annur, S. (2018). The development of student worksheet by using guided inquiry learning model to train student's scientific attitude. *Unnes Science Education Journal*, 7(1), 19-26. <https://doi.org/10.15294/usej.v7i1.15799>
- 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. <http://dx.doi.org/10.1155/2013/834584>
- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: principles and applications*. New York: Kluwer Nijhof Publishing. <https://doi.org/10.1007/978-94-017-1988-9>.
- 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
- Ilfiana, A., Widodo, W., & Setiarso, P. (2021). The Improvement of student's critical thinking skills through the development of science learning material based socioscientific issues with interactive multimedia-assisted on gadget. *Jurnal Penelitian Pendidikan IPA*, 7(4), 496-501. <https://doi.org/10.29303/jppipa.v7i4.764>
- 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>
- Ismayawati, B., Purwoko, A.A., & Muntari. (2016). Pengaruh model pembelajaran berbasis masalah (Pbm) dalam setting pembelajaran kooperatif tipe TGT dan GI terhadap keterampilan berpikir kritis dan hasil belajar kimia peserta didik SMAN 1 Aikmel. *Jurnal Penelitian Pendidikan IPA*, 2(1). <https://doi.org/10.29303/jppipa.v2i1.33> [Indonesian]
- Mahanal, S., Zubaidah, S., Bahri, A., & Dinnurriya, M. S. (2016). Improving students' critical thinking skills through Remap NHT in biology classroom. *Asia-Pacific Forum on Science Learning and Teaching*, 17(2). Retrieved from: <https://www.researchgate.net/publication/318093694>
- 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/pdf>
- 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. <https://doi.org/10.5901/jesr.2014.v4n1p445>
- Ozden, B., & Yenice, N. (2014). An analysis of the secondary education students' scientific attitudes*. *International Journal of Contemporary Educational Research*, 1(2), 86-97. Retrieved from www.ijcer.net
- Pitafi, A.I. & Farooq, M. (2012). Measurement of scientific attitude of secondary school students in Pakistan. *Academic Research International*, 2(2), 379-392. Retrieved from: www.savap.org.pk%5Cnwww.journals.savap.org.pk
- Priska, M., Peni, N., & Wao, Y.P. (2021). Development of acid-base devices integrating ARCS motivation strategy in problem-solving learning model scientific attitude and critical thinking skills of students. *Jurnal Penelitian Pendidikan IPA*, 7(SpecialIssue). <https://doi.org/10.29303/jppipa.v7iSpecialIssue.1126>
- Qing, Z., Ni, S., & Hong, T. (2010). Developing critical thinking disposition by task-based learning in chemistry experiment teaching. *Procedia - Social and Behavioral Sciences*, 2(2), 4561-4570. <https://doi.org/10.1016/j.sbspro.2010.03.731>
- 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. https://doi.org/10.1007/978-1-4020-6262-9_14
- Stephenson, N. S., & Sadler-Mcknight, N. P. (2016). Developing critical thinking skills using the Science Writing Heuristic in the chemistry laboratory. *Chemistry Education Research and Practice*, 17(1), 72-79. <https://doi.org/10.1039/c5rp00102a>
- Subali, B. (2019). *Prinsip asesmen dan evaluasi pembelajaran edisi ketiga*. Yogyakarta: UNY Press. [Indonesian]
- Sudarmini, Y., Kosim., & Hadiwijaya, A.S. (2015). Pembelajaran fisika berbasis inkuiri terbimbing dengan menggunakan LKS untuk meningkatkan keterampilan berpikir kritis ditinjau dari sikap ilmiah siswa Madrasah Aliyah Qamarul Huda

- Bagu Lombok Tengah. *Jurnal Penelitian Pendidikan IPA*, 1(1), 35-48.
<https://doi.org/10.29303/jppipa.v1i1.4>
[Indonesian]
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Jakarta: Trim Komunikata. [Indonesian]
- Thiagarajan, S., Semmel, D., & Semmel, M. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Minneapolis: University of Minnesota.
- Zhou, Q., Huang, Q., & Tian, H. (2013). Developing student's critical thinking skills by task-based learning in chemistry experiment teaching. *Creative Education*, 4(12A), 40-45.
<http://dx.doi.org/10.4236/ce.2013.412A1006>