

# The Development of an Integrated Assessment Instrument for Measuring Analytical Thinking and Science Process Skills

**Abstract.** This research aims to develop instrument and ~~know~~ determine the characteristics of an integrated assessment instrument. This research uses 4-D model, which includes define, design, develop, and disseminate. The primary product is validated by expert judgment, tested its readability by students, and assessed its feasibility by chemistry teachers. This research involved 246 students of grade XI IPA of four senior high schools in Yogyakarta, Indonesia. Data collection techniques include interview, questionnaire, and test. Data collection instruments include interview guideline, item validation sheet, users response questionnaire, instrument readability questionnaire, and essay test. The results show that the integrated assessment instrument has Aiken validity value of 0.95. Item reliability was 0.99 and person reliability was 0.69. Teachers response to the integrated assessment instrument is very good. Therefore, the integrated assessment instrument is feasible to be applied to measure the students analytical thinking and science process skills.

## INTRODUCTION

Assessment is one of important parts of learning process. Assessment relates to the objectives and learning process. A learning process includes planning, implementation of learning process, assessment of learning outcomes, and students follow up assessment [1]. A good learning will not be successful without a good assessment. Assessment is expected to reflect the ability of students as a whole, both in terms of knowledge, attitudes, and skills [2-3] and able to stimulate the students to optimize their potential [4-5]. Therefore, learning process and learning outcomes assessment are two components that cannot be separated.

One of the learning activities which can train the students ability in developing those three competencies is internship activities in the laboratory [6]. Practicum activities in laboratory can motivate students to develop a number of important processes skills. One of the most important skills for students in the 21<sup>st</sup> century is science process skills [7]. Moreover, in order to acquire laboratorial experience, practicum cannot be separated from chemistry learning, investigation experience, improvement attitude toward chemistry, and science process skills [8].

Science process skill is a cognitive skill that can be measured by using written test [9]. This condition arises because science process skills relates to the cognitive domains of the students which includes the analytical thinking ability [10]. One of the assessments which uses written test is essay. The advantage of essay test is that they are capable of measuring various aspects of the science process skills of the students, hence the essay has the potential to be developed as an instrument of science process skills assessment. Achievement test is able to provide information which describes the ability of the students [11].

On the other hand, assessment of the students learning outcomes on cognitive and psychomotor aspects often presents various problems. One of the problems which often occur is that the teacher chooses the form of classroom assessment without considering of how they assess and what will be assessed [12]. In addition, an assessment of the



science process skills has not been optimally implemented by the teachers. The ability of teachers to prepare and conduct an assessment is still lacking; in fact, there are many teachers who do not yet have sufficient understanding of the scoring system [13]. Moreover, the learning outcome instruments used by the teacher for daily test or used by the school for general tests are not yet qualified as a good test [14].

Generally, the teacher presents a question which dominantly uses short answers and only covers narrow scale of remembering, understanding, and applying, making the students ability only used to solve on lower level cognitive problems [15]. In relation to the assessment of the students ability, the short answer test at the level of lower level cognitive only capture a small portion of students skill and intelligence [16]. Whereas in the 21<sup>st</sup> century, the students are required to master not only lower order thinking skills, but also to the higher order thinking skills.

One of higher order thinking skills is the analytical thinking. Analytical thinking is not all about memorizing facts, but efforts to improve complex and creative understanding to achieve a better learning-learning objective. Analytical thinking ability is a higher level of cognitive which can be achieved by the students after mastering lower level cognitive domain such as remembering, understanding, and applying. To evaluate the level of analytical thinking, it is needed a standardize learning outcome instrument which can also be utilized to measure science process skill through integrated laboratory activity. This condition needs to be carried out since effective laboratory activity can develop intellectual ability, scientific inquiry, and problem solving ability of the students [17].

Based on these descriptions, it is necessary for the preparation of an integrated assessment instrument that can be used to measure analytical thinking objectively, as well as cover various aspects of science process skills in chemistry practicum learning in laboratory.

## RESEARCH METHODS

This type of research is an education research and development (R&D). This research uses 4-D (four-D) model which was developed by Thiagarajan, Semmel & Semmel [18]. 4-D model consists of four main stages, which are define, design, develop, and disseminate. Education research and development is a method used to produce products in the field of education in the form of an evaluation tool, and then to test the validity, practicality, and effectiveness of the product. This research model is a procedural development research, which is a descriptive research, showing the steps which have to be followed to produce the final product.

This research and development was conducted on 5 September - 25 December 2015. The research sites were four senior high schools (SMA) in Yogyakarta, Indonesia. Research location selection was determined by purposive sampling techniques based on school ranking (high, medium, and low) according to the results of chemistry National Examination (UN) 2015. The subjects involved 246 students of class XI IPA academic year of 2015/2016. The subjects covers the students of SMA 2 Yogyakarta, SMA 4 Yogyakarta, SMA 6 Yogyakarta, and SMA 10 Yogyakarta. The allocation of time used in the empirical test is 90 minutes.

The types of data are feasibility and characteristics instrument data. Feasibility data of the integrated assessment instrument were obtained from validation by expert judgment, feasibility of chemical products by the chemistry teacher, and the instrument readability test by the students. Whereas the instrument characteristic data were obtained from the empirical test. The types of data collection instruments include interviews guideline, validation items sheet, users response questionnaire, instrument readability questionnaire, and essay test. Data analysis techniques are qualitative and quantitative.

Data were analyzed using quantitative descriptive approach in accordance with the statistical methods used then interpreted. Initial data obtained were the items validation result data by expert judgment. Data analysis of the validation item result used rules of score conversion as seen in Table 1 [19].

**Table 1.** Score Conversion

No.	Score
1	$X \geq X + 1.Sbx$ Very Good
2	$\bar{X} + 1.Sbx > X \geq X$ Good
3	$\bar{X} > X \geq X - 1.Sbx$ Poor
4	$X < X -$

**Comment [DA1]:** Applying is considered as Low order thinking skills in Science?

**Comment [DA2]:** Application or applying is better right?

**Comment [DA3]:** How about the research design?

**Comment [DA4]:** This is the research design. I suggest to put this before saying on educational research and development

**Comment [DA5]:** Reference?





%).

The other results from items validity is in the form of improvement suggestion for the instrument from expert judgment. Instrument improvement was mainly ~~applied~~ applied on writing technique and language used which can a rise

After the items are validated and assembled into the test, the next step is scoring product by the chemistry teacher using Likert scale as seen in Table 2.

confusion and difficulty to the students in understanding the question. It is important since sentence mistake is a factor that can influence test performance.

### The Validity of Instrument Readability Questionnaire Data

The validity of instrument readability questionnaire was obtained from assessment by using ~~validity~~ validity sheet of instrument readability questionnaire which was conducted by content and evaluation specialist. Validation results of instrument readability questionnaire is shown in ~~Table~~ Table 6.

**Table 6.** Readability Questionnaire Validation Results

No.	Expert Judgment	Score	Category
1.	Evaluation Specialist	3 . 8 7	Very Good
2.	Content Specialist	4 . 0 0	Very Good
	<b>Average</b>	<b>3 . 9 3</b>	<b>Very Good</b>

Max. Score = 4.00

The validity of instrument readability questionnaire was calculated by V Aiken formula. Based on analysis results of content and evaluation specialist, it is concluded that all items have very high validity (100 %). The calculation result shows that the highest and the lowest validity are 1.00 and 0.83. Validity index average of instrument readability questionnaire is 0.95 which is very high.

### The Validity of UsersResponse Questionnaire Data

The validity of usersresponses questionnaire is obtained from scoring sheet of responses questionnaire which was ~~applied~~ applied by content and learning evaluation specialist. The validation was conducted by giving ~~accuracy~~ accuracy score in every statement item. The validation results of usersresponses questionnaire is shown in Table 7.

**Table 7.** UsersResponse Questionnaire Validation Results

No.	Expert Judgment	Score	Category
1.	Evaluation Specialist	3 . 8 8	Very Good
2.	Content Specialist	4 . 0 0	Very Good
	<b>Average</b>	<b>3 . 9 4</b>	<b>Very Good</b>

Max. Score = 4.00

The validity of usersresponses questionnaire was calculated by using V Aiken formula. Based on result analysis from content and evaluation specialist, it is concluded that every item in response questionnaire has a very high validity (100 %). The result shows that the highest and lowest validity are 1.00 and 0.83. Validity index average of instrument readability questionnaire is 0.98 which is very high category.

### Instrument Readability Test Data

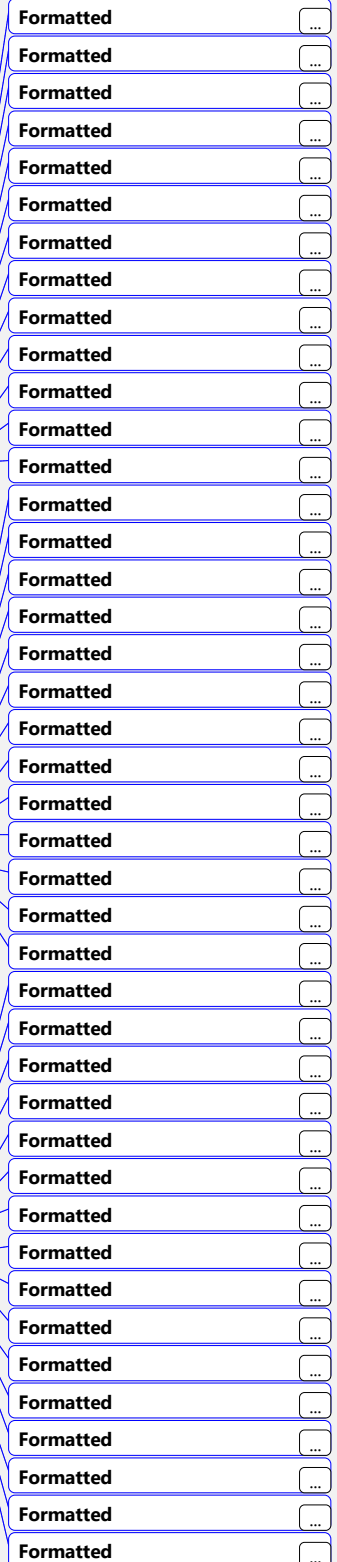
The Readability test involved 58 students of XI IPA 5 SMAN 2 Yogyakarta and XI 4 SMAN 10 Yogyakarta. The aim of the readability test is to determine the understanding level of the students to the use of language, picture, table, graphic, content, and instruments display. In summary, the results of instrument readability test is shown in ~~Table~~ Table 8.

**Table 8.** Instrument Readability Test Results

No.	Components	Score	Category
1	Didactic	3 . 3 1	Very Good
2	Construction	3 . 1 1	Very Good
3	Technical	3 . 1 7	Very Good
	<b>Average</b>	<b>3 . 1 4</b>	<b>Very Good</b>

Max. Score = 4.00

Table 8 shows that most of the students give strongly agree respond to all aspect which was measured in instrument readability questionnaire. It shows that the use of language, picture, table, graphic, or content presented in instrument can be understood easily by the students and feasible to be field tested.











6. D. Donnelly, J. O'Reilly, and O. McGarr, *Journal of Research in Science Education* **43**, 1571—1592 (2013).
7. S. Areesophonpichet, *Proceeding of the Asian Conference on Education*, Osaka, Japan, 2013, pp. 1-15.
8. E. Susilaningsih, *Jurnal Penelitian dan Evaluasi Pendidikan* **16**, 234—248 (2012).
9. B. Subali, *Jurnal Cakrawala Pendidikan* **15**, 130—144 (2011).
10. S. Özgelen, *Eurasia Journal of Mathematics, Science & Technology Education* **8**, 283-292 (2012).
11. S.C. Bauer, *Journal of Education Policy Analysis Archives* **46**, 1-18 (2000).
12. I.H. Wenno, *Jurnal Cakrawala Pendidikan* **29**, 176—188 (2010).
13. S. Suwandi, *Model Asesmen dalam Pembelajaran* (Yuma Pustaka, Surakarta, 2011), p.2.
14. D. Mardapi, Kumaidi, and B. Kartowagiran, *Jurnal Penelitian dan Evaluasi Pendidikan* **15**, 326—341 (2011).
15. B.D. Kulkarni and V.D. Tendolkar, *Journal of Health Sciences Education* **2**, 43—47 (2015).
16. Murtono and E. Miskiyah, *Jurnal Inovasi dan Pembelajaran Fisika* **1**, 1-12 (2014).
17. E.M. Stone, *CBE—Life Sciences Education* **13**, 90—101 (2014).
18. S. Thiagarajan, D. Semmel, and M. Semmel, *Instructional Development for Training Teachers of Exceptional Children: A Sourcebook* (University of Minnesota, Minneapolis, 1974), p.5.
19. D. Mardapi, *Pengukuran, Penilaian, dan Evaluasi Pendidikan* (Nuha Medika, Yogyakarta, 2012), p.162.
20. E.P. Widoyoko, *Evaluasi Program Pembelajaran* (Pustaka Pelajar, Yogyakarta, 2009), p.239.