



Development and Validation of an Integrated Assessment Instrument to Assess Students' Analytical Thinking Skills in Chemical Literacy

Eka Ad'hiya

Department of Chemistry Education, Post Graduate Programme, Yogyakarta State University, Yogyakarta, Indonesia, *eka.adhiya2016@student.uny.ac.id*

Endang W Laksono

Assoc. Prof., Department of Chemistry Education, Post Graduate Programme, Yogyakarta State University, Yogyakarta, Indonesia, *endang_widjajanti@uny.ac.id*

The research focuses on the development and validation of an integrated assessment instrument to assess students' analytical thinking skills and chemical literacy in chemical equilibrium. The research consists of 5 phases ie analysis, design, development, implementation, and evaluation. The research participants were 174 students of grade 12th senior high schools in Yogyakarta, Indonesia. The participants were selected using stratified random sampling. The Integrated assessment instrument to assess students' analytical thinking skills and chemical literacy consists of 20 open-ended multiple-choice questions developed from 12 integrated indicators. Expert validation result shows that 20 questions are declared valid with Aiken's V value greater than 0.86. The result was collated and coded for analysis with Rasch model approach. 3 items (15%) were included in the very difficult category, 3 items (15%) were included in the difficult category, 8 items (40%) were included in the moderate category, 5 items (25%) were included in the easy category, and 1 question (5%) was included in the very easy category. It turns out that this instrument is capable to measure students' analytical thinking skill and chemical literacy.

Keywords: integrated assessment, analytical thinking, Chemical literacy, ADDIE model, Rasch model, chemical equilibrium

INTRODUCTION

One of the education goals is to make students have a system of knowledge, skills to be independent, effective and creative (Petrovska & Veselinovska, 2013). The objective is aligned with the educational goals proposed by Temel (2014), which teaches students on how to think and solve problems well so that students become individuals who can think, ask, research and formulate solutions to a problem. The 21st-century challenge was put forward by Vanada (2010) stating that students need to improve collaborative

Citation: Ad'hiya, E., & Laksono, E. W. (2018). Development and Validation of an Integrated Assessment Instrument to Assess Students' Analytical Thinking Skills in Chemical Literacy. *International Journal of Instruction*, 11(4), 241-256. <https://doi.org/10.12973/iji.2018.11416a>

skills and creative, analytical and practical thinking skills. Moreover, the role of high-level thinking skills in modern times is not only a progression of facts and concepts, but it must be a person's ability to solve problems as his contribution to society (Lateef, Dahar & Latif, 2016).

Anwar & Mumthas (2014) state that one aspect of a person's intelligence success is the analytical thinking skill. In fact, however, the score of student achievement in Indonesia is still at a medium and low level. This is in line with the results of Trends in International Mathematics and Science Study (TIMSS) in 2015, which shows that the average score of learning achievement on the subjects of mathematics and science students in Indonesia is 397, lower than the average TIMSS score overall country of 500. The students' achievement in math and science subjects in Indonesia is at a very high level (1%), High (6%), Medium (24%) and low (55%). This shows that students in Indonesia are at a medium and low level of achievement and that there is a need to develop students' cognitive abilities to be at higher level.

Celik (2014) stated that changes in knowledge and technology occur quickly and affect society. Therefore, an understanding of the facts of science and the relationship between knowledge, technology, and society is necessary. The ability to apply knowledge to everyday life is called as science literacy. The 2015 result of the Program for International Students Assessment (PISA) shows that the average science literacy score of 15-year-old students is 403, while the average science literacy scores of all countries that follow this assessment program are 493, with the score of Indonesia including 10 countries with the lowest science literacy score. This indicates that assessment instruments for measuring high-level thinking skills are essential. The questions on assessment instruments work not only to evaluate students' understanding, but also stimulate students to think. This corresponds with the statement put forward by Santhitiwanich, Pasiphol, Tangdhanakanond (2013) stating that a specific test is required that teachers will use to measure and assess students' analytical thinking skills.

Montaku, Kaittikomol & Tiranathanakul (2012) stated that low student scores are due to low intentions in conducting tests and exams using analytical thinking-oriented questions. It is in line with the statement of Fensham & Bellochi (2013) that the most of the assessment of learning outcomes in science still focus on low-level thinking skills, where the questions given by teachers to students are still memorizing without thinking. This is consistent with the result of preliminary observation in high school students, which shows the result that 19.3% of students stated strongly agree, 57.89% of students stated agree and 22.8% of students stated disagree that the problem given by the teacher is in the form of memorization. And the results of the questionnaire also showed that 17.5% of students stated strongly disagree, 78.9% of students said disagree, 3.5% of students agreed that the evaluation of the teacher given in the form of a matter of reasoning. It means that teachers in schools only use assessment instruments that measure rote or other low-level thinking skills. This suggests that an assessment instrument that can measure high-level thinking skills is necessary.

Shwartz, Benz-vi & Hofstein (2006) explains that science literacy is one of the science Education goals, and it requires a new standard of content, pedagogy, and assessment.

Based on this, it can be concluded that a measuring instrument is needed to measure students' chemical literacy. But in reality, there are still a few teachers who provide the evaluation questions related to everyday life. This is consistent with the result of preliminary observation on high school students, showing the result of 52.63% student express disagree, 3.5% student express disagree, 38.6% agree and 5.2% student expressly agree that evaluation question was given by the teacher related to the problems of everyday life. It means that teachers only use theoretical assessment instruments without relating to real life. This indicates that the assessment instrument that contains literacy is necessary. Because of the lack of an instrument that teachers can use, teachers often use only an assessment instrument that does not know the level of ability to be measured.

There is a relationship between analytical thinking skills and chemical literacy, as Cigdemolu & Geban (2015) suggests, which states that a person with chemical literacy will be able to increase the frequency of asking, searching for information and using it when needed, and can analyze the advantages and disadvantages associated with the position in the debate. Based on the above, an integrated measuring instrument is needed to measure both aspects of analytical thinking and chemical literacy. The integrated assessment in this study is a form of assessment that can measure the skills of analytical thinking and also aligned can measure the level of student's chemical literacy.

The development of integrated assessment instruments in this study focused on chemical equilibrium materials. The selection of this material is based on the fact that this material is related to everyday life, but a matter used in very little judgment that has linked matter with life. In addition, the question that many used in the form of calculations that constitute the cognitive domain of the application has not yet reached the higher cognitive domain of analysis. Based on this, the researchers felt it was important to develop an integrated assessment instrument that could measure students' analytical thinking and chemical literacy skills on chemical equilibrium materials.

LITERATURE

Integrated Assessment

El-Yassin (2015) states that integrated assessments should be made in educational institutions, as these assessments are important ways to improve student learning outcomes. Moreover, McPhun (2010) explained that integrated assessment is a process of combining and integrating various learning outcomes into a series of simple, realistic and based cases. It is reinforced with Bourke and Mentis (2014) statement that the integrated assessment framework shows how different approaches can be aligned with specific goals, and further SAQA (2005) explains that there are several key elements in the integrated assessment, ie need to demonstrate the competencies applied, including the provision of evidence of results; Link assessment to define goals (standards and qualifications); The context in which the assessment should be done is key; practicality and efficiency of assessment are important; integrated assessment should add value to student learning; Relating theory to practice makes integrated judgments more authentic and coherent to be used as a method of assessment of learning; integrated assessment

should make a variety of assessment instruments in which more than one model of learning is assessed; integrated assessment is an appropriate way to assess the extent to which learners can integrate concepts, ideas and actions; assessment of integration is the assessment of knowledge, skills, and personal qualities.

Analytical Thinking Skills

The definition of analytical thinking is explained by Art-in & Sitthipon (2012), analytical thinking is a competency in identifying and classifying different aspects that can be objects, stories or events, and find relationships of those aspects. Meanwhile, Chonkaew (2016) states that students with analytical thinking skills can express their concepts understanding and implement their knowledge. Other definition, analytical thinking skills are a process of the mental formation used to determine the solution of a problem (Anwar & Mumthas, 2014). Furthermore, the indicators for analytical thinking skills explained by Ramirez & Ganaden (2008) that are differentiating, organizing and attributing. Differentiating is differentiate a relevant sections with irrelevant and unimportant sections, organizing is determined how an element fits with its structure, and attributing is determined the point of view, the value or purpose underlying the existence of an element.

Based on these matters, it is concluded that the analytical thinking skill is the competence in differentiating, organizing and relating an object, theory, problem or event, and can determine the relation of those aspects based on certain reason, principle or function.

Chemical Literacy

Jurecki & Wander (2012) defines scientific literacy as recognizing the natural world, understanding the key concepts and principles of science, possessing scientific reasoning capacity, and being able to use scientific knowledge for personal or social purposes. The definition of literacy of science also stated in the OECD (2013), PISA defines literacy as the ability to connect issues or problems related to science, and ideas in science. Moreover, literacy of science is the ability to understand the process of science and interpret the scientific information contained in everyday life (Fives, Huebner, Birnbaum & Nicolich, 2014).

A person who has the ability of science literacy will be able to understand science and the relationship of science with society and the environment, and can know the concepts, laws, theories and basic principles of chemistry and can use the skills of the scientific process (Turiman, Omar, Daud & Osman, 2012). Based on that, then someone with chemical literacy must understand the basic concepts of science because chemistry is a branch of science and in it, there are theories that are able to explain the natural world, and the knowledge gained from chemistry can be transferred and applied to other topics in science and technology. A person with chemical literacy must know the basic purpose of chemistry. This includes the chemical principles that teach us to understand macroscopic phenomena by interpreting the microscopic and symbolic levels of chemistry. A person that have chemical literacy must be able to use knowledge in everyday life.

Chemical literacy can explain as the ability to understand the concept of chemistry and connect the concepts to explain the relationship between the chemical concepts with phenomena, problems or facts that exist in everyday life, which is shown by three aspects of explaining the phenomenon by using the concept of chemistry, using a chemical understanding of problem solving and analyzing the strategies and benefits of chemical applications.

METHOD

Design and Model

This research design is Research and Development (R & D) and adapting the ADDIE model which consists of five phases; these are analysis, design, development, implementation, and evaluation. This model was chosen because this study aims to develop an integrated assessment instrument and further determine the quality of the instruments that have been developed.

Procedure

The first phase is *analysis*. At this stage, an analysis of the chemical equilibrium material domain in the chemical education curriculum is based on the regulation of the ministry of education and culture. In addition, at this stage also analyzed indicators of analytical thinking skills and indicators of chemical literacy from various literatures. So at this stage, there are three kinds of indicators, namely the indicator of chemical equilibrium material, the indicator of analytical thinking skill and the chemical literacy indicator.

The second phase is *design*. At this stage, the design of the form of assessment instrument will be developed. The design of the instrument starts from the determination of an indicator called as an integrated indicator (integration of chemical equilibrium material indicators, indicators of analytical thinking and chemical literacy indicators). This integrated indicator will serve as a reference in the process of developing an integrated assessment instrument. The integrated assessment instrument must be completed by scoring techniques. The scoring technique design must also be determined in accordance with the shape of the instrument to be developed. The third phase is the *development* of items on integrated assessment instruments and scoring guidelines in accordance with designs that have been determined.

The fourth phase is *implementation*. At this phase, an empirical test of integrated assessment instruments has been developed. Empirical tests are performed for 90 minutes. Each participant was given an integrated assessment instrument sheet on chemical equilibrium material and an answer sheet. Implementation of empirical tests was conducted after classroom learning on chemical equilibrium (Formative test). The last phase is the *evaluation*, the process of determining the quality of integrated assessment instruments based on the results of empirical tests. Guideline for tasks and outputs for each phase are summarized in table 1.

Table 1
Guideline for tasks and outputs of each phase

Phase	Tasks	Outputs
Analysis	- the student needs analysis - standard competence analysis - analysis of analytical thinking- skills - analysis of chemical literacy indicators	- profile of student needs - indicator of chemical equilibrium material - indicators of analytical thinking skills - indicators of chemical literacy
Design	- design the instrument form - design product indicators - design scoring methods	- instrument form - integrated indicators - coring guidelines
Development	- develop a product	- integrated assessment instrument to measure analytical thinking skill and chemical literacy on chemical equilibrium
Implementation	- expert validation - an empirical test	- data validation - student comment and data
Evaluation	- interpret test results	- quality of product

Participant and Sampling Technique

The participants are required to be involved only in the implementation phase, which is the empirical test of integrated assessment instruments. The participant was selected using stratified random sampling based on the school's ranking (high, moderate and low) on the national exam results. The participant is expected to be representative of all levels, so then all students at every level can use the product. Participants consist of 174 12th graders from the senior high school in Yogyakarta, Indonesia.

Data Collection Tools

Data collection tools used in the validation process content is a *validation sheet*. The validation sheet contains integrated indicators and integrated assessment instrument items. This validation sheet also provides a column for validator (expert) to determine the match between the indicator and the developed item, and also provided a column for expert criticism and suggestions to refine the integrated assessment instrument already developed. *Integrated assessment instruments* are also used as data collection tools to determine the results of the evaluation phase, which is the determination of quality.

Data Analysis Technique

Qualitative analysis is used to analyze chemical equilibrium indicators, analytical thinking skills indicators, and chemical literacy indicators. Qualitative analysis will be explained descriptively. The quality of integrated assessment instruments that have been developed, analyzed quantitatively. The quality of the instrument is seen from the validity, reliability, item difficulty.

The first quantitative calculation of content validity performed by expert judgement was analyzed by Aiken's V. Furthermore, the results of the implementation phase were analyzed using Rasch model approach. The results of this calculation are constructed the

validity, factor analysis, difficulty level, test information function, and reliability. Construct validity was analyzed using the result of anti-image value and factor analysis was analyzed from the eigenvalue of the correlation matrix. The difficulty level of the item was analyzed using the Winstep program on the output item entry. And the test reliability was calculated by looking at the intersection curve of the information function and the standard error of estimation.

FINDINGS AND DISCUSSION

ANALYSIS PHASE: In this phase, there are three analyzes: basic competence analysis, indicators of Analytical thinking skills analysis and chemical literacy indicators analysis.

Basic Competence Analysis

The core competencies are explained in more detail by the existence of basic competencies, where each core competency consists of two basic competencies. Basic competencies for the first competence are 1) Analyze the factors that influence the shift in equilibrium direction applied in the industry, and 2) Determine the quantitative relationship between the reactants with the reaction product of an equilibrium reaction. Basic competencies for the second competence are 1) Design, perform, and conclude and present experimental results of factors affecting the shift in equilibrium direction, and 2) Solve problems related to the quantitative relationship between reactants with the reaction results of an equilibrium reaction.

Analysis of analytical thinking skills indicators

Indicators of analytical thinking skills are differentiating, organizing and attributing (Ramirez & Ganaden, 2008). The product characteristics for each indicator of analytical thinking skills are presented in table 2:

Table 2

Indicators of analytical thinking skills

No	Indicator	Product Characteristic
1	Differentiating	Presented problems or phenomena and students are expected to determine a true element, which can be distinguished from other elements based on the theory of matter of chemical equilibrium and the relevance of it with the problem or phenomenon revealed.
2	Organizing	Presented problems or phenomena and students are expected to determine an element in accordance with the structure of the problem or phenomenon is based on the theory of chemical equilibrium material.
3	Attributing	Presented the problem or phenomena and students are expected to determine the value or purpose that became the basis of an element, which is in accordance with the theory of chemical equilibrium matter.

Analysis of chemical literacy indicators

Indicators of chemical literacy are explaining the phenomena using chemical concepts, using chemical understanding in problem-solving and analyzing strategies and benefits of chemical applications (Shwartz, Ben-Zvi, & Hofstein, 2005). The product characteristics for each chemical literacy indicator are presented in table 3:

Table 3
Indicators of chemical literacy

No	Indicator	Product Characteristic
1	Explain the phenomenon with the concept of chemistry	Presented phenomena that occur in everyday life related to chemical equilibrium material, so that students are expected to explain the phenomenon by using the concept of chemistry that is owned
2	Using a Chemical understanding to solve the problems	Presented an issue related to the chemical equilibrium constant, students are expected to use a chemical understanding of the chemical equilibrium constant in solving the problem
3	analyzing the strategies of Chemical application	Presented an issue related to the chemical equilibrium constant, students are expected to use a chemical understanding of the chemical equilibrium constant in solving the problem

The analysis phase is the foundation of all phases in the development process (Muruganatham, 2015). The analysis phase is in line with Obizoba (2015) that stated objective of identifying problems, objectives, content and context, so that the results of this phase of analysis will make it easier to identify the problems and processes needed to improve the input at the design stage.

DESIGN PHASE: This phase aims to develop a grid of questions based on core competence and basic competence of chemical equilibrium materials, as well as the indicators of analytical thinking and chemical literacy indicators. The integration of these three indicators namely integrated indicators. The integrated indicators in this study consist of 12 indicators, which are described in detail in table 4.

Table 4
Integrated indicators

No	Integrated Indicators
1	Be able to distinguish the meaning of equilibrium in the phenomenon of a chemical equilibrium reaction
2	Be able to relate the meaning of chemical equilibrium in the phenomenon of a chemical equilibrium reaction
3	Be able to distinguish the quantitative relationship of reactants and reaction products in the phenomenon of a chemical equilibrium reaction
4	Be able to organize the quantitative relationship of reagents and reaction products in solving the problem of a chemical equilibrium reaction
5	Be able to distinguish the quantitative relationship between reactants and reaction products in solving the problem of a chemical equilibrium reaction
6	Be able to relate the quantitative relationship between reactants and reaction products in solving the problem of a chemical equilibrium reaction
7	Be able to relate the influence of concentration, temperature, pressure, volume and catalyst in the phenomenon of a chemical reaction equilibrium concept
8	Be able to distinguish the influence of concentration, temperature, pressure, volume and catalyst in the phenomenon of a chemical reaction equilibrium concept
9	Be able to distinguish the influence of concentration, temperature, pressure, volume and catalyst in the phenomenon of a chemical reaction equilibrium concept
10	Be able to organize the effect of concentration, temperature, pressure, volume and catalyst in solving the problem of a chemical reaction equilibrium concept
11	Be able to relate the concept of equilibrium reaction in the determination of strategy and its utilization in the chemical industry
12	Be able to distinguish the concept of equilibrium reaction in the determination of strategy and its utilization in the chemical industry

The Integrated assessment instrument was in the form of open-ended multiple-choice questions. The instrument was designed with five options lettered A-E out of which only one was the correct as the accepted answer. The more options will increase reliability (Thorndike & Thorndike-Christ, 2010), and reduce success in guessing answers (Dehnad, Nasser, Hosseini, 2014).

The instruments developed based on integrated indicators and design scoring methods. The design of the scoring guidelines consisting of a score of 3, 2, 1, 0 (Bayrak, 2013) is shown in Table 5.

Table 5
Design of scoring guidelines

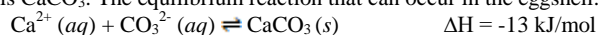
Score	Description
3	If multiple choice answers are true and the reasons are true for a complete reason
2	If multiple choice answers are true and the reasons are true for incomplete reasons
1	If the multiple choice answer is true and the reason is wrong
0	If the multiple choice answer is wrong and the reason is wrong

DEVELOPMENT PHASE: The product titled “Integrated Assessment-Analytical Thinking and Chemical Literacy (IA-ATCL)”. Initial products of IA-ATCL consists of 30 open-ended multiple-choice questions developed based on integrated indicators, where each indicator is represented by 1 or more questions. An example of IA-ATCL question is shown in table 6.

Table 6
An example of IA-ATCL items

TOPIC: EGGSHELL

1. In the eggshell contain nutrients such as calcium, iron, and other minerals that can be used as animal feed. The content of eggshells includes 95% calcium carbonate (CaCO_3), phosphorus, and high-protein chains close to 5%, less chitin, 0.05% chitosan. The chemical formula of the eggshell is CaCO_3 . The equilibrium reaction that can occur in the eggshell:



Based on the eggshell equilibrium reaction, the thinnest eggshell can be obtained at the time of ...

- | | | |
|------------|-----------|----------|
| A. Spring | C. Winter | E. Rainy |
| B. Snowing | D. Summer | |

The Reason:

On stem, items are presented reading materials of chemical literacy with certain topics related to everyday life. Each item consists of five answer options and a reason column. Each question comes with scoring guidelines, as an example shown in Table 7.

Table 7
An example of the scoring guideline

No	Answer	Description	Score
1	D	If multiple choice answers are wrong.	0
		If multiple choice answers are true and the reasons are wrong.	1
		If multiple choice answers are correct and reasons are incomplete. The reasons are incomplete: The temperature is raised, then the equilibrium shifts to the left	2

If multiple choice answers are true and the reasons are complete. 3

Complete Reason:

The equilibrium reaction is an exothermic reaction if the temperature is increased it will shift to the left, meaning CaCO_3 decomposes into Ca^{2+} and CO_3^{2-}

IMPLEMENTATION AND EVALUATION PHASE: The phase to determine how the quality of the Products.

Validation

In content validation it would be useful to involve at least 5 or 7 experts in its field for Judge-scale content domains using scales rating (Yaghmale, 2003). In this Research, validation was performed by 7 experts, consisting of 1 chemical equilibrium material expert, 1 evaluation expert, and 5 school practitioners. Content validity of each items shows in figure 1.

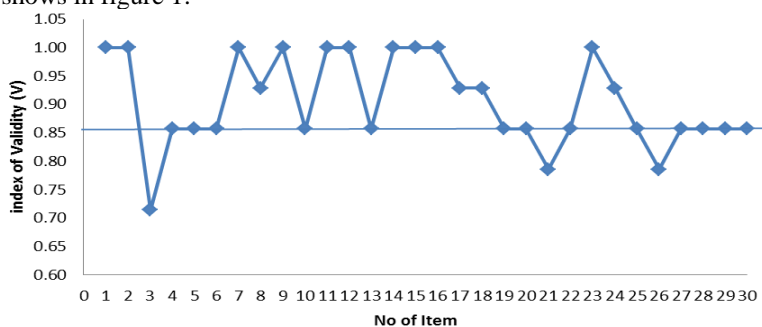


Figure 1
The Content Validity

The result of the content validity calculation of IA-ATCL show that 27 item were developed valid because the value of Aiken's V was greater than 0.86 (Aiken, 1985), but they needed improvements for some numbers based on expert revision (Medeiros, Junior, Torres, Vitor, Santos & Barichello, 2015). The items with index of validation below 0.86 were removed (Kovacic, 2017). The results of the content validity will support the results of construct validity (Yaghmale, 2003).

Table 8

The value of KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.804
Bartlett's Test of Sphericity	Approx. Chi-Square	1.090E3
	df	190
	Sig.	.000

Table 8 shows the requirement of construct validity that the value of KMO-MSA is $0.804 > 0.5$, which indicates that the samples taken are categorized as satisfactory, and Bartlett test significance value is $0.000 < 0.05$ indicating that research variable can be predicted and analyzed further (Beavers, Lunsbury, & Richard, 2013).

The construct validity is calculated based on factor analysis, by interpreting the anti-image matrices data. 25 items of IA-ATCL are considered valid because the value of anti-images is greater than 0.5 (Chan & Idris, 2017). The value of anti-images for each item is shown in figure 2.

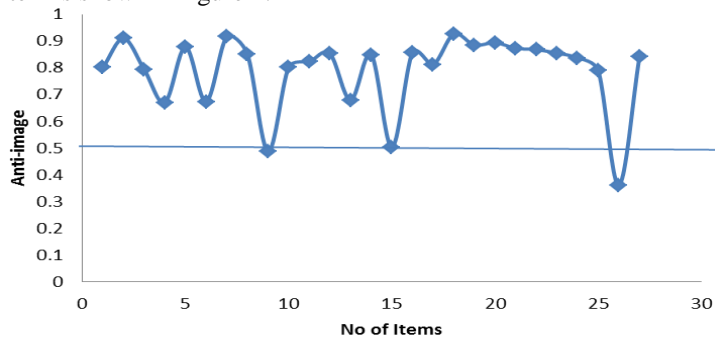


Figure 2
The Construct Validity

Item Fit

Item fit is used to determine if an item is functioning optimally in measuring. Adam & Khoo (1990) explained that a good item would have a fit item value between 0.77-1.30. The result of the measurement of the fit items is presented in figure3.

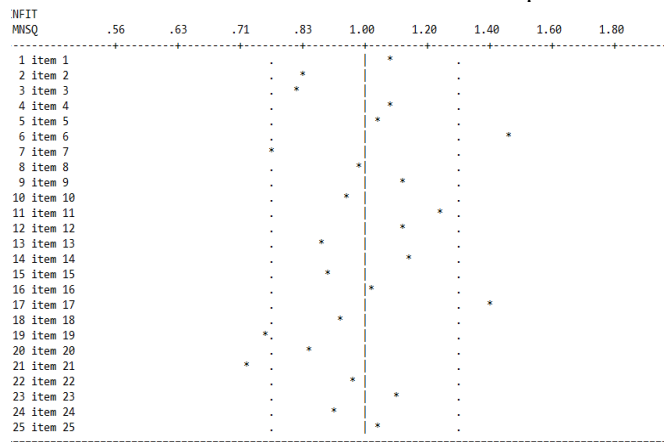


Figure 3
Item Fit

It shows that the 4 items are not fit with the model, so it needs to be fixed or may be eliminated (Yusoff, M.S.B, 2015). The next step taken is to eliminate the 4 problems, and to calculate the item fit analysis again. In addition to these four items, there is one problem that is eliminated as well, that is, the problem is in the range of 0.77. So there are 20 questions that are stated fit.

Items Difficulty

Based on the data, the difficulty of the items at IA-ATCL has spread evenly at each level. Complete data items difficulty is presented in figure 4.

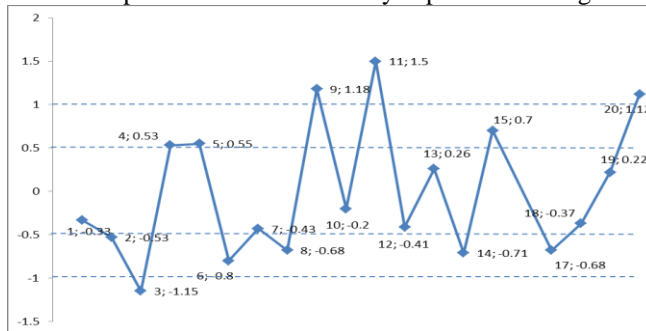


Figure 4
Items difficulty of IA-ATCL

According to Adedoyin & Mokobi (2013), if the calculation of the difficulty of the items has a δ (Measures) ≥ 1 then it is categorized Very Difficult, $0.5 \leq \delta < 1$ categorized Difficult, $-0.5 \leq \delta < 0.5$ categorized Moderate, $-1 \leq \delta < -0.5$ categorized easy and $\delta \leq -1$ categorized very easy. The results of the difficulty calculation on IA-ATCL using Winstep, indicating that there are 3 items (15%) included in very difficult category, 3 items (15%) included in difficult category, 8 items (40%) included in moderate category, 5 items (25%) included in the easy category, and 1 question (5%) included in the category very easy. The level of difficulty of this study is already good, because an item is good if it is at -2.0 logit and 2.0 logit (Hambleton & Swaminathan, 1985).

Item Function Information

The maximum value of the test information function is at the value of 20.774 logit which is in the ability (θ) of the participant in the value of -0.20 logit with an estimated standard error of 0.2194. It shows that the tests are good and include high category so it can be concluded that the integrated assessment instrument has a high degree of accuracy to estimate the student's ability (Wiberg, 2004). Figure 4 shows that the intersection curve between the test information function and the standard error estimation are at $\theta = -3.05$ logit and $+3.00$ logit. This indicates that the integrated assessment instrument developed is reliable if it is tested to participants who have capabilities in the range of $-3.05 < \theta < +3.00$ logit. Thus, it can be concluded that the integrated assessment instrument developed is reliable to be tested to test participants with low, moderate or high ability (Sumintono & Widhiarso, 2005). Furthermore, Hambleton, Swaminathan & Rogers (1991) suggests that the use of the test information function is more reliable than the use of reliability in classical test theory, due to several factors: the shape depends only on the items in the test, and has an estimated measurement error at each level of ability.

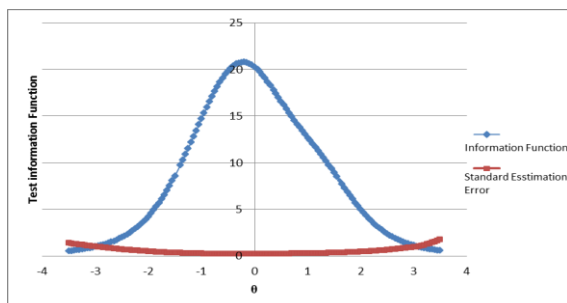


Figure 5

The intersection curve between the test information function and the standard error estimation

This study is in line with the research conducted by El-yassin (2015) that develops integrated assessment in medical education. The results of this study also shows that integrated assessment is one of the ways to enhance student outcomes. Another study was also conducted by Bourke & Mentis (2014) who examined the integrating assessment approaches for inclusive education. The results of this study also show that integrated assessment is the best way to assess high-need students, because they have to assess many aspects at a time. The interrelation between the research results and analyses and this present research is that the integrated research is one of the best ways to measure several aspects at a time, so expect the assessment of student learning outcomes can be maximized.

CONCLUSION

This research has successfully developed 20 items of questions that can measure students' analytical thinking and chemical literacy skills. The instrument consists of 20 questions scattered into the level of difficulty. 3 items were included in very difficult category, 3 items were included in the difficult category, 8 items were included in the moderat category, 5 items were included in the easy category, and 1 question was included in the very easy category. Morckish (1995) explains that integrated assessments are easy to use and support for the advancement of learning theories, teachers, and learning styles of learners. Along with this, integrated assessment can be used to measure learners' involvement in learning (Dorman, Waldrip, Fisher, 2006), so that the instrument is expected to be used as an appraisal instrument on applying a learning model or approach aimed at improving analytical and literacy thinking skills chemistry of students. This instrument can also serve as a guideline in the development of integrated assessment instruments on other chemical materials.

REFERENCES

Adam, R.J., & Khoo, S. (1996). *Quest: The Interactive Test Analysis System Version 2.1*. Camberwell, Victoria: The Australian Council for Educational Research.

- Adedoyin, O.O., & Mokobi, T. (2013). Using IRT psychometric in examining the quality of junior certificate mathematic multiple choice examination test items. *International of Asian Social Science*, 3(4), 992-1011.
- Aiken, L.R. (1985). Three coefficient for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45, 131-142.
- Anwar, B. & Mumthas. (2014). Taking triarchic teaching to classrooms: giving everybody a fair chance. *International Journal of Advanced Research*, 2(5), 455-458.
- Art-in & Sitthipon. (2012). Development of teachers' learning management emphasizing on analytical thinking in Thailand. *Procedia - Social and Behavioral Sciences*, 46, 3339 – 3344.
- Bayrak, B.K. (2013). Using two-tier test to identify primary students' conceptual understanding and alternative conceptions in acid base. *Mevlana International Journal of Education*, 3(2), 19-26.
- Beavers, A.S., Lounsbury, J.W., Richard, J.K., et al. (2013). Practical considerations for using factorial analysis in Educational Research. *Practical Assessment, Research & Evaluation*, 18(6), 1-13.
- Bourke, R., & Mentis, M. (2014). An assessment framework for inclusive education : integrating assessment approaches. *Assessment in Education : Principles, Policy & Practice.*, 37-41.
- Celik, S. (2014). Chemical literacy levels of science and mathematics teacher candidates. *Australian Journal of Teacher Education*, 39 (1), 1-15.
- Chan, L.L., & Idris, N. (2017). Validity and reliability of the instrument using exploratory factor analysis and cronbach's alpha. *International Journal of Academic Research in Business and Social Sciences*, 7(10), 400-410.
- Chonkaew, P., Sukhummek, B., & Faikhamta, C. (2016). Development of analytical thinking ability and attitudes towards science learning of grade-11 students through science technology engineering and mathematics (STEM education) in the study of stoichiometry. *Chemistry Education Research and Practice*.
- Cigdemoglu, C., & Geban, O. (2015). Improving students' chemical literacy level on thermochemical and thermodynamics concepts through context-based approach. *Chemistry Education Research and Practice*, 1-15.
- Dehnad, A., Nasser, H., & Hosseini, A.F. (2014). A Comparison between three-and four-option multiple choice questions. *Procedia-Social and Behavioral Sciences*, 98, 398-403.
- Dorman, J. P., Fisher, D. L., & Waldrip, B. G. (2009). Using the student perceptions of assessment questionnaire (SPAQ) to develop an assessment typology for science classes. *Journal of Science Education*, 9(1), 13-17.

- El-Yassin, H.D. (2015). Integrated assessment in medical education. *J Contemp Med Sci, 1*(4), 36–38.
- Fensham, P.J., & Bellocchi, A. (2013). Higher order thinking in chemistry curriculum and its assessment. *Thinking Skills and Creativity, 10*, 250–264.
- Fives, H., Huebner, W., Birnbaum, A.S., & Nicolich, M. (2014). Developing a measure of scientific literacy for middle school students. *Science Education, 98*(4), 549–580.
- Hambleton, R.K., & Swaminathan, H. (1985). *Item Response Theory Principles and Applications*. Boston: Kluwer Nijhoff Publishing.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of items response theory*. Newbur Park: Sage Publication.
- Jurecki, K., & Wander, M.C.F. (2012). Science literacy, critical thinking, and scientific literature: guidelines for evaluating scientific literature in the classroom. *Journal Of Geoscience Education, 60*, 100–105.
- Lateef, A., Dahar, M.A., & Latif, K. (2016). Impact Of Higher Order Thinking Skills Of University Students On Their Academic Performance. *Sci.Int.(Lahore), 28*(2), 2031-2035.
- Medeiros, R.K.S., Junior, M.A.F., Torres, G.V, Vitor, A.F., Santos, V.E.P., & Barichello, E. (2015). Content validity of an instrument about knowledge on nasogastric intubation. *Bioscience Journal, 31*(6), 1862-1870.
- McPhun, Helen. (2010). *Integrated assessments – engaging ways to enhance learner outcomes*. Good Practice Publication Grant.
- Montaku, S., Kaittikomol, P., & Tiranathanakul, P. (2012). The model of analytical thinking skill training process. *Research Journal of Applied Science, 7*(1), 17-20.
- Mockrish, R.M.Jr. (1995). *Integrated assessment for an integrated curriculum: A field guide to middle school education*. Ann Arbor, Michigan: UMI Company.
- Muruganantham, G. (2015). Developing of E-content package by using ADDIE Model. *International Journal of Applied Research, 1*(3), 52-54.
- OECD. (2013). *PISA 2015 Draft science framework*. Paris: OECD.
- Obizoba, C. (2015). Instructional design models—framework for innovative teaching and learning methodologies. *International Journal of Higher Education Management, 2*(1), 40-51.
- Petrovska, S., & Veselinovska, S.S. (2013). Contemporary pedagogical approaches for developing higher level thinking on science classes. *Procedia - Social and Behavioral Sciences, 92*, 702 – 710.
- Ramirez, R.P.B., & Ganaden. (2008). Creative activities and students' higher order thinking skills. *Education Quarterly, 66* (1), 22-33.

Santhitiwanich, A., Pasiphol, S., Tangdhanakanond, K. (2013). The integration of indicators of reading, analytical thinking and writing abilities with indicators of subject content. *Procedia - Social and Behavioral Sciences*, 116, 4854 – 4858.

SAQA. (2005). *Guidelines for integrated assessment*. South African: The South African Qualifications Authority.

Shwartz, Y., Ben-Zvi, R., & Hofstein, A. (2005). The importance of involving high-school chemistry teachers in the process of defining the operational meaning of 'chemical literacy'. *International Journal of Science Education*, 27(3), 323-344.

Schwartz, Y., Ben-Zvi, R., & Hofstein, A. (2006). Chemical literacy: what does this mean to scientists and school teachers?. *Journal of Chemical Education*, 83(10), 1557-1561.

Sumintono, B. & Widhiarso, W. (2005). *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial*. Bandung: Trimkom Publishing House.

Temel, S. (2014). The Effects of problem based learning on preservice teachers' critical thinking dispositions and perceptions of problem-solving ability. *South African Journal of Education*, 34(1), 1-20.

Thorndike, R. M., & Thorndike-Christ, T. (2010). *Measurement and Evaluation in Psychology and Education* (8th Ed). Upper Saddle River, NJ: Pearson/Merril Prentice Hall.

Turiman, P., Omar, J., Daud, A.M., & Osman, K. (2012). Fostering the 21st century skills through scientific literacy and science process skills. *Procedia - Social and Behavioral Sciences*, 59, 110 – 116.

Vanada, D.I. (2010). Practically creative: the role of design thinking as an improved paradigm for 21st century art education. *Techné Series A*, 21(2), 21-33.

Wiberg, M. (2004). *Classical Test Theory vs. Item Response Theory*. Sweden: Umea University.

Yaghmale, F. (2003). Content validity and its estimation. *Journal of medical education*, 3(1), 25-27.

Yusoff, M.S.B. (2015). Psychometric properties of the secondary school stressor questionnaire among adolescents at five secondary schools. *Journal of Taibah University Medical Sciences*, 10(2), 159-168.