**TÜRK FEN EĞİTİMİ DERGİSİ**
**Yıl 16, Sayı 4, Aralık 2019**

**Journal of**
**TURKISH SCIENCE EDUCATION**
**Volume 16, Issue 4, December 2019**

http://www.tused.org

# Measuring Junior High School Students' Science Learning and Science Process Skills through an Integrated Science Instructional Assessment

SUKARDIYONO [1], Dadan ROSANA [2], Wipsar Sunu Brams DWANDARU [3]

[1] Instructor Dr., Physics Education Department, Universitas Negeri Yogyakarta, Yogyakarta, 55281, INDONESIA
[2] Assoc. Prof., Dr., Science Education Department, Universitas Negeri Yogyakarta, Yogyakarta, 55281, INDONESIA
[3] Assist. Prof., Dr., Physics Education Department, Universitas Negeri Yogyakarta, Yogyakarta, 55281, INDONESIA, https://orcid.org/0000-0002-9692-4640

**Received:** 09.09.2017 **Revised:** 06.10.2018 **Accepted:** 09.10.2019

The original language of article is English (v.16, n.4, December 2019, pp.467-477, doi: 10.36681/tused.2020.1)

**Reference:** Sukardiyono, Rosana, D. & Dwandaru, W. S. B. (2019). Measuring junior high school students' science learning and science process skills through and integrated science instructional assessment. *Journal of Turkish Science Education, 16* (4), 467-477.

## ABSTRACT

This study aimed to produce a valid and reliable assessment tool for junior high school students' integrated science learning and determine their improvement of science process skills (SPS) using the integrated science instruction. The research procedure consisted of ten stages, i.e., preliminary study, planning, development of draft assessment tool, validation of draft assessment tool, first revision of draft version, pilot study, second revision (stage II), operational field tests, third revision (stage III), and dissemination. The Aiken V formula was utilized to obtain instrument's validity, whilst interclass correlation coefficients were calculated for its reliability. The results showed that the instructional assessment was valid with lowest (0.89) and highest (0.98) validity values. The reliability coefficients of the first, second and third observation sheets were found to be 0.93, 0.94 and 0.92, respectively, which fell into a very good category. This means that these observation sheets were feasible for measuring SPS. The results indicated that the students' SPS levels were very good for predicting, experimenting, observing, and measuring, while 'concluding and communicating' aspects were good.

**Keywords:** Integrated science instructional assessment, performance assessment, portfolio assessment, science learning, science process skills.

## INTRODUCTION

The study, which focused on the development of integrated science instructional assessment, purposes to overcome the assessment problem of science learning, especially in junior high school. Because previous assessments have tended only to measure cognitive abilities, teachers rarely pay attention in assessing students' science process skills (SPS)

Corresponding author e-mail: wipsarian@uny.ac.id

© ISSN:1304-6020

[Yildirim et al., 2016]. This shows an inconsistency with the nature of science learning, which includes products, processes, attitudes, and applications. Chiappetta (2010) and Flick and Lederman (2004) reveal that science learning should refer to three aspects of the nature of science, e.g., products, processes, and attitudes. This means that science learning and assessment call for an integrated manner of implementation. Integrated science learning intends to make science learning more meaningful, effective, and efficient, whereas its assessments enhance students' science learning (Alanazi, 2017).

Integrated science learning increases scientific literacy that enables students to understand and realize the needs of the community and fosters their engagement with technology-oriented economy (Sofowora and Adekomi, 2012). The science, technology, and society (STS) approach is a learning reform to meet community-related technological needs (Driver et al., 2000). Rosana (2017) shows that the STS approach gives a better impact on student's understanding of science and its relationship(s) with technology and society. Moreover, students benefit from the mastery of various concepts and demonstrate the ability to connect these concepts to everyday life (Wijayanti and Basyar, 2016).

Preliminary studies have shown that integrated science learning has not been implemented in several junior high schools in Yogyakarta Special Region (Retnawati et al., 2016). They have also depicted that very few science teachers tended to apply integrated science instructional assessment to measure cognitive, affective, and science process skills (Yildirim et al., 2016). Further, there has not been any example of science problem that can measure integrated science learning of student's cognitive abilities and science process skills. The aforementioned problem calls for an assessment instrument of science learning to measure the students' science process skills and cognitive abilities. That is, developing a science assessment instrument (called integrated science instructional assessment) needs to especially focus on performances and portfolio-based assessments of SPS. Portfolio assessment gives opportunities for students to show what the students have learned and mastered during the learning processes.

Performance-based assessment demonstrates or applies students' acquired knowledge and abilities. Performance assessment not only aims to determine students' abilities to apply their knowledge and skills in accordance with pre-determined learning objectives but also focuses on direct assessment (Hosnan, 2014). Stiggin and Chappuis (2012) argue that performance assessments engage students with activities that demonstrate their skills or produce their featured products of standardized mastery. Popham (2005) describes performance assessment as an approach to measure the students' performances of a particular task. Palm (2008) and Douglas (2009) state that performance assessment gives a better chance for measuring complex skills and communication. In a similar vein, Lai (2011) argues that performance assessment can assess deeper knowledge and skills as compared to traditional one. Performance appraisal, as an alternative assessment, requires students to actively demonstrate their performances. Namely, it assesses final results, processes and skills together (Meutia et al., 2013).

Using portfolio is another method to assess the students' performances. Hamp-Lyons and Condon (2000) argue that portfolio provides a broader measurement of what students can do. Johnson (2002) defines portfolio as a collection of work done by students. In other words, portfolio shows student's work from the beginning to the end of task/activity/learning. Yang (2003) defines portfolio as a compilation of student work, documentation of learning efforts, progress of learning, learning achievement, and reflective learning toward the material. Many studies emphasize the need of portfolio (e.g., Birgin, 2003; De Fina, 1992; Gussie, 1998; Micklo, 1997; Mumme, 1991; Norman, 1998). These studies indicate that portfolio provides more reliable and dynamic data for teachers, parents, and students. Portfolio assessment supplies clear information about how students overcome their weaknesses, and assists teachers in planning teaching progress. In view of Samad (2004) and Tabatataei (2012), portfolio assessment changes the teacher role of assessment. That is, portfolio interactively gives feedbacks for improving student learning progress, self-assessment, and self-reflection. The portfolio shows student progress, achievement, and self-reflection in one or more areas (Paulson and Paulson, 1991). Several studies refer to the importance of portfolio assessment (e.g., Defina, 1992; Hamp-Lyons and Condon, 2000; Harris and Sandra, 2001; Yancey, 1999). However, most of the related studies have been qualitative. Few studies have quantitatively adapted portfolio assessment.

This study used performance and portfolio assessments to measure the SPS, which is a broad set of capabilities, and can be transferred to many disciplines (Sheeba, 2013). The SPS, which reflects the behaviors of scientists, engage students in science-related activities (Alkan, 2016; Tawil and Liliasari, 2014). In line with the aforementioned definitions, Akinbobola and Afolabi (2010) describe the SPS as mental and physical abilities/competencies that are necessary for effectively studying Science and Technology in solving problems.

The SPS cannot be separated from conceptual understanding that involves science learning and applications (Karamustafaoglu, 2011). Students, who conduct a scientific inquiry to gain knowledge and skills through the SPS, directly concentrate on problem solving approaches within everyday life or practical knowledge as the aim of science (Abungu, Okere, and Wachanga, 2014). Aktamis (2008) states that the SPS includes skills that make students science-savvy in improving their living standards and facilitate their understanding of the nature of science (Kucuk and Cepni, 2015). The SPS affects their personal, social, and global lives. The SPS is necessary for generating and using scientific information, conducting scientific research, and solving problems. Also, the SPS is an important instructional approach to gain scientific knowledge, and scientific investigations through cognitive and investigative skills (Kruea-In and Fakcharoenphol, 2015). Hence, students can find facts, concepts, and theories with their gained science process skills and scientific attitudes (Nurhemi et al., 2011). The SPS affords students to be active, develop a sense of responsibility for self-study, enhance permanent learning, and understand research methods. Overall, they behave like scientists (Ergul et al., 2011).

The American Association for the Advancement of Science (Aydin, 2013) divided SPS into two categories (basic and integrated SPS). Basic SPS consists of observing, classifying, data recording, measuring, using space and time relationship(s), using numbers, inferencing, and predicting. Moreover, integrated SPS comprises of transforming and controlling variables, interpreting data, making hypotheses, making operational definitions, using data and formulating models, and experimenting. Karamustafaoglu (2011) classifies basic science process skills as observing, classifying, measuring, and predicting, and integrated science process skills as identifying and defining variables, collecting and transforming data, creating tables and graphs based on data, describing relationship(s) between variables, interpreting data, manipulating material, recording data, formulating hypotheses, designing inquiries, inferencing, and making generalization. Goldston and Downey (2013) categorizes basic science process skills as observing, inferencing, classifying, measuring, estimating, predicting and communicating and integrated science process skills as experimenting, making hypotheses, identifying variables, making operational definitions, collecting data, reporting and interpreting data, and making model. Chiapetta and Koballa (2010) categorize and define various aspects of SPS. That is, basic science process skills include observing, classifying, making time-space relationship(s), measuring, inferencing, and predicting, whilst integrated science process skills contain making operational definition, formulating the model, controlling variables, interpreting data for producing explanations, inferencing or making hypotheses through graphed or located data, and experimenting.

Previous studies have stressed how to improve the SPS through the use of performance and portfolio assessments. Al-Rabaani (2014), who exploited a 14-item questionnaire for basic and integrated SPS, showed that there was no difference in the ability of moderate mastery of science based on gender. Feyzioglu et al. (2012), who examined the validity and reliability of the SPS assessment, administered to 222 vocational high school students. The reliability co-efficient of the test with 30 multiple-choice questions (i.e., observing, classifying, measuring, communicating, summarizing, predicting, formulating hypotheses, identifying variables, organizing and interpreting data, designing inquiry, and obtaining data) was found to be 0.83. The results of confirmatory factor analysis supported its validity and reliability. Moreover, Ozgelen (2012), who investigated sixth and seven grade students' SPS within the cognitive framework, found a low mean score for the integrated science process skills. He also reported that private school students had higher scores of the SPS than public school students.

As described above, producing a valid and reliable science instructional assessment is essential to measure students' SPS levels through performance and portfolio assessments. Hence, two research questions guided the current study; (i) is the integrated science instructional assessment tool valid and reliable for measuring students' SPS? and (ii) how does the integrated science instructional assessment determine the improvement of students' SPS?

## METHODS

### a) The Sample of the Study

The sample of the study consisted of 32 grade 7 students in one class. The SPS assessment tool was administered to them through 3 meetings with pre-determined practicum activities. A group of experts (one evaluation expert, one subject matter expert from Universitas Negeri Yogyakarta, and four science teachers from junior high school in Yogyakarta Special Region) validated the feasibility of the SPS test.

### b) Research Model

Through a quantitative research method, this study determined the validity of the integrated science instructional assessment based on performance and portfolio, and measured students' SPS levels via the assessment tool. This study followed Borg and Gall's (1983) development model, which incorporates ten developmental research steps: 1) preliminary research and gathering information, 2) planning, 3) development of draft assessment tool, 4) pilot study, 5) revising the assessment tool (first revision), 6) field testing, 7) revising the operational assessment tool (second revision), 8) operational field testing, 9) revising the final assessment tool (third revision), and 10) disseminating and implementing the final assessment tool. This study only performed two tests (pilot study and operational field tests with three revision steps and limited dissemination without implementation). The authors chose this model because of its detailed research procedures. The flowchart of research procedure is presented in Figure 1 in the Appendix.

### c) Data Collection Tools

The preliminary data were validated by the SPS observation sheet, worksheet-based SPS assessment, scientific attitude observation sheet, self-assessment of scientific attitudes, peer- assessment of scientific attitudes, lesson plans, and worksheets, which were collected from a group of experts (i.e., material and evaluation experts, teachers, and peers). The collected data were then tabulated for each type of the assessment tool over each validated item.

### d) Data Analysis

The content-validity coefficient is calculated using Aiken's formula (Aiken, 1985),

$$V = \sum \frac{s}{n(c-1)}$$

where $s$ is equal to ($r$ – Io), Io is the lowest validity assessment value, $c$ is the highest validity value, $r$ is the score given by the experts, and $n$ is the number of experts. The range of V is from 0 to 1. The limit of item validity depends on the number of value categories and the number of validators involved in the assessment. If six validators and 4 value categories are available, the item is valid in case of V ≥ 0.78.

### e) Interpretation Method

The results of the initial draft feasibility assessment included suggestions on all 48 items in the performance and portfolio assessments. Six validators' scores of the assessment results were then analyzed using Aiken's formula to calculate V. The results of the instrument validation and item are divided into three categories: (1) valid without revision, (2) valid with revision, and (3) invalid. In this case, $n$ is six, the lowest validity score (lo) is one and the highest validity score ($c$) is four. The assessment of each item produces the V score and then compares it with the Aiken's V table of 0.78. The validity or feasibility of each item can then be determined. The validation results of the performance and portfolio assessments to measure SPS are reviewed based on substantial, construction, and language aspects.

## FINDINGS

**Table 1**. *The validity results of performance and portfolio assessment observation sheets*

| | Item | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Performance assessment observation sheet | Validity | 0.92 | 0.96 | 0.89 | 0.95 | 0.88 | 0.96 | 0.87 | 0.99 | 0.85 | 0.92 | 0.90 | 0.94 |
| Portfolio assessment observation sheet | | 0.97 | 0.88 | 0.98 | 0.82 | 0.95 | 0.87 | 0.92 | 0.98 | 0.95 | 0.98 | 0.95 | 0.88 |

As seen from Table 1, the validity results indicated that all items in the performance and portfolio assessment sheets were valid and higher than 0.78. The validators provided quantitative assessments, feedbacks and suggestions (see Table 2). The statements consisted of 48 items in 3 meetings representing the four indicators of the SPS. Validity assessment of the draft tool is shown in Table 3.

**Table 2**. *Validators' feedbacks and suggestions*

| Components | Feedbacks and Suggestions |
|---|---|
| *Science process skills* observation sheet | 1) Several indicators in the specifized table of the observation sheet need to be clearly elaborated.<br>2) Some words that are less suitable with the default language spelling need a writing improvement.<br>3) Effective language selection is needed to be improved.<br>4) Multiple negative statements, such as *except* and *not*, should be avoided.<br>5) Improvements on the items containing unnecessary statements.<br>6) Improvement on some improper spielers. |
| Syllabus | 1) Tidy up the writing error(s) of the question statement. |
| Lessons Plans | 1) The SPS aspects should be integrated in the learning objectives.<br>2) The objectives and subjects should be specifized on the focus of the sentences.<br>3) Images should be adjusted to the material and the needs of students.<br>4) As possible as the *guided* content should be reduced. |
| Worksheet | 1) The use of images should be more contextual and appropriate.<br>2) Using images from private collections is recommended instead of those from the internet.<br>3) The problem orientation ought to be re-checked to avoid "clue" in the question. |
| Formative test | 1) Problem 1 measuring the mass is incorrect.<br>2) Problem 15 does not need the related image because it can be solved without any image.<br>3) Problem 16 may obviously be answered because of a clue.<br>4) Problem 20 should avoid the use of the word "except".<br>5) Answer choices of c and d in problem 23 are incorrect.<br>6) In problem 29, the column entitled '*t*' should be replaced by 'final temperature'. |

**Table 3**. *Recapitulation of the experts' validation results*

| No | Validator | The number of item | | |
|---|---|---|---|---|
| | | Valid without revision | Valid with revision | Invalid |
| 1 | Experts | 38 items | 6 items | - |
| 2 | Teachers | 12 items | 32 items | - |

After feedbacks and suggestions were obtained, the next step covered to revise the initial draft. An initial assessment tool was apparent for further testing to determine the reliability of the item in the portfolio assessment instrument. Reliability of the assessment instrument was calculated via the final assessment results of five assessors in each activity. Eight students carried out each activity within 3 meetings. The reliability values of performance and portfolio assessments from each meeting are displayed in Table 4.

**Table 4**. *The reliability values of performance and portfolio assessments from each meeting*

| No | Observation Sheet | Reliability Value | Category |
|---|---|---|---|
| 1 | I | 0.93 | *Excelent* |
| 2 | III | 0.94 | *Excelent* |
| 3 | III | 0.92 | *Excelent* |

The measurement results of individual student's SPS were obtained from the average total score from each meeting. These averages had different scoring ranges because each aspect possessed different numbers of indicators. In view of by Djemari (2008), the classification was conducted with four categories (e.g., excellent, good, sufficient, and lacking). The assessment results of the adjusted scores are summarized in Table 5. Aspects such as predicting and drawing conclusions appeared in the measurement results from the student responses to the worksheets. Figure 2 shows the measurement results of the SPS in each meeting.

**Table 5**. *The assessment results per meeting*

| No | SPS aspects | Average total scores | | | Averages | Categories |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | | |
| 1 | Predicting | 3.47 | 3.31 | 3.63 | 3.47 | Excellent |
| 2 | Experimenting | 3.6 | 3.19 | 3.66 | 3.48 | Excellent |
| 3 | Observing | 3.36 | 3.17 | 3.44 | 3.32 | Excellent |
| 4 | Measuring | 3.42 | 3.56 | 3.66 | 3.55 | Excellent |
| 5 | Concluding | 2.47 | 2.66 | 2.69 | 2.61 | Good |
| 6 | Communicating | 2.61 | 2.84 | 2.69 | 2.71 | Good |

## DISCUSSION and CONCLUSION

The results of the performance and portfolio assessments showed that the reliability values were larger than 0.70. This means that the observation-sheets based SPS assessment has an excellent reliability in accordance with Gliem and Gliem (2003). The performance and portfolio assessment sheets also indicated that the scoring rubrics of the SPS were clear and easily understood by the assessors. The performance and portfolio assessments, which fell into good category could be used for the operational field trial.

Because students were able to estimate events that would occur in accordance with their observations, the 'predicting' aspect had a very good category in regard to average scores of the three meetings. In addition, the fact those students were able to engage their senses in good learning showed the observing aspect as a very good category. Moreover, the fact that students could express the object or event accordingly revealed that the 'measuring' aspect was classified under 'very good' category.

In general, students seemed to be active during the experiments. Since the experiments were conducted within groups, students were able to share tasks and involve the practicum activities. This may result from 'very good' category for the experimenting aspect. The low average scores of the 'concluding and presenting or communicating' aspects pointed that students tended to be reluctant to give opinions and ask friends. Inability to draw conclusions may stem from a lack of the links amongst data, objectives, and the concepts they had already understood. Some students might have conducted the presentation by only reading their experimental results. Moreover, most students were unable to systematically explain and clarify their ideas using good and correct language.

The fact that the average score of the SPS aspects in the third meeting had a higher score than the first and second meetings may come from the third meeting in the practical activity 'Heat Transfer.' Thereby, students may have applied the SPS to the practical activity. This may also result from the feature of the aforementioned experiement in everyday life or students' familarity with the experiments. This means that students may have been familiar with the tools and materials used in the 'experimenting' aspect. Furthermore, the experimental procedure may relatively facilitate and guide students' performances. Students were more skilled in the 'measuring' aspect; e.g., using such instruments as a beaker glass and a stopwatch, holding a thermometer correctly, and reading the scale properly, such as observing the scale parallel to the eye. Furthermore, students could precisely determine temperature changes using a thermometer.

In light of the findings, this study concludes that the performance and portfolio assessment instruments measuring the SPS are feasible and valid according to the experts. Further, empirical test shows that the instruments have a good reliability and higher than 0.70. The measurement results of the SPS revealed a very good category for the 'predicting, experimenting, observing, and measuring' aspects, and a good category for the 'concluding and communicating' aspects.

**Suggestions**

Further studies should be conducted with a large sample size to test applicability of the integrated science instructional assessment tool. Moreover, future studies ought to test how the integrated science instructional assessment tool improves junior high school students' SPS performances.

## REFERENCES

Abungu, H. E., Okere, M. I. O., & Wachanga, S. M. (2014). The effect of science process skills teaching approach on secondary school student's achievement in chemistry in Nyando District, Kenya. *Journal of Educational and Social Research*, *4* (6), 359-372.

Aiken, L. R. (1985). Three coefficients for analyzing the reliability, and validity of ratings. *Educational and Psychological Measurement*, *45*, 131-142.

Akinbobola, A., & Afolabi, F. (2010). Analysis of science process skills in west african senior secondary school certificate physics practical examinations in Nigeria. *Bulgarian Journal of Science and Educatian Policy*, *4*, 32-47.

Aktamis, H., & Ergin, O. (2008). The effect of scientific process skills education on student's scientific creativity, science attitudes and academic achievements. *Asia-Pasific Forum on Science Learning and Teaching*, *9* (1), 1-20.

Alanazi, F. H. (2017). Effectiveness of the proposed training formative assessment programmeand its impact on teaching style improvements of Saudi science teachers in Saudi Arabia. *Journal of Turkish Science Education*, *14* (1), 35-56.

Alkan, F. (2016). Experiental learning: Its effects on achievement and scientific process skills. *Journal of Turkish Science Eduction*, *13* (2), 15-26.

Al-rabaani, A. (2014). The acquisition of science process skills by Omani's pre-service social studies' teachers. *European Journal of Educational Studies*, *6* (1), 13-19.

Aydin, A. (2013). Representation of science process skills in the chemistry curricula for grades 10,11 and 12. *International Journal of Education and Practice*, *1*, 51-63.

Birgin, O. (2003). *Investigation of the application level of a computer based portfolios*. Master's Thesis, Karadeniz Technical University, Trabzon.

Borg W., & Gall, M. (1983). *Educational research: An introduction.* New York: Longman Inc.

Chiappetta, E. L., & Koballa, T. R. (2010). *Science instruction in the middle and secondary schools*. 7th edition. New Jersey: Merrill/Prentice Hall.

De Fina, A. (1992). *Portfolio assessment: getting started*. New York: Scholastic Professional Books.

Djemari, M. (2008). *Instrument and non-test drafting technique.* Yogyakarta: Mitra Cendikia Offset.

Douglas, G. (2009). *Performance assessment: a key component of a balanced assessment system*. Department of Research, Evaluation, and Assessment. Virginia Beach City Public Schools.

Driver, R., Leach. J., Miller, R., & Scott, P. (2000). *Young people's images of science*. Buckingham: Open University Press.

Ergul, R., Simsekli, Y., & Calis, S. (2011). The effects of inquiry-based science teaching of elementary school student's science process skills and science attitudes. *Bulgarian Journal of Science and Educatiaon Policy*, *5*, 48-68.

Feyzioglu, B., Demirdag, B., Akyildiz, M., & Altun, E. (2012). Developing a of a science process skills test for secondary students: validity and reliability study. *Educational sciences: Theory and Practice*, *12* (3), 1899-1906.

Flick, L., & Lederman, M. (2004). *Scientific inquiry and nature of science*. Boston: Kluwer Academic Publishers.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting cronbach alpha reliability coefficient for Likert-type scales, *The Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education* (pp. 82-88), Ohio State University, Ohio.

Goldston, M., & Downey, L. (2013). *Your Science Classroom. Becoming an Elementary/ middle school science teacher*. USA: SAGE Publications.

Gussie, W.F. (1998). *Assessment of the implementation of portfolio assessment in the K-8 school districts in New Jersey*. Doctoral Dissertation, Widener University.

Hamp-Lyons, L., & Condon, W. (2000). *Assessing the portfolio: principles for practice, theory, and research*. New Jersey: Hampton Press.

Harris, S., & Sandra, M. (2001). Portfolio use in two principal programs. *Rural Educator*, *23*, 19-23.

Hosnan (2014). *The scientific and contextual approach in 21st century learning*. Bogor: Ghalia Indonesia.

Lai, E. (2011). *Performance-based assessment: some new thoughts on an old idea*. Pearson Education.

Karamustafaoglu, S. (2011). Improving the science process skills ability of science student teachers using I diagram. *Eurasian Journal of Physics and Chemistry Education*, *3*, 26-38.

Kruea-In, C., Kruea-In, N. & Fakcharoenphol, W. (2015). A study of Thai in-service and pre-service science teacher's understanding of science process skills. *Procedia-Social and Behavioral Sciences*, *197*, 993-997.

Kucuk, M., & Cepni, S. (2015). A qualitative study to explain middle school student's understanding of nature of science. *Journal of Turkish Science Education*, *12*(3), 3-20.

Meutia, H., Johar, R. & Ahmad, A. (2013). The ability of University students as teacher candidates in implementing perforamance assessment to assess students learning outcomes in Mathematics learning. *Jurnal Peluang*, *1*, 63-70.

Micklo, S. J. (1997). Math portfolio in the primary grades. *Childhood Education*, *97*, 194-199.

Mumme, J. (1991). *Portfolio assessment in mathematics. California Mathematics Project*, Santa Barbara: University of California.

Norman, K. M. (1998). *Investigation of the portfolios as an alternative assessment procedure*. Doctoral Dissertation, The University of Memphis.

Nurhemy, T., Santosa, S. & Probosari, R. (2011). The application of active learning with silent demonstration to increase the science process skills of grade VIII Students of SMP 14 Surabaya. *Pendidikan Biologi*, *3*, 61-71.

Ozgelen, S. (2012). Students science process skills within a cognitive domain framework. *EURASIA Journal of Mathematics. Science & Technology Education*, *8* (4), 283-292.

Paulson, F. L., & Paulson, P. R. (1991). Portfolios: stories of knowing. *Annual Meeting of the Claremont Reading Conference*.

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research, and Evaluation, 13*, 1-12.

Popham, W. (2005). *Classroom assessment, What teachers need to know*. USA: Allyn and Bacon.

Retnawati, H., Samsul, H., & Ariadi Chandra, N. (2016). Vocational high school teachers' difficulties in implementing the assessment in curriculum 2013 in Yogyakarta Province of Indonesia, *9*(1), 33-48.

Rosana, D., Kadarisman, N., Maryanto, A., & Sugiharsono. (2017). The evaluation of science learning program, technology, and society application of audio bio harmonic system with solar energy to improve crop productivity. *Jurnal Pendidikan IPA Indonesia*, *6*(1).

Samad, A. A. (2004). *Essential language testing for Malaysian teachers*. Kuala Lumpur: Malindo Printers Sdn Bhd.

Sheeba, M. (2013). An anatomy of science process skills in the light of the challanges to realize science instruction leading to global excellence in education. *Education Confab, 2*, 108-123.

Sofowora, O. A., & Adekomi, B. (2012). Improving science, technology, and mathematics education in Nigeria: A case study of Obafemi Awolowo University Ile-Ife. *African Journal of Educational Studies in Mathematics and Sciences*, *10*, 1-8.

Stiggins, R., & Chappuis, J. (2012). *An introduction to student-involved assessment for learning*. USA: Pearso Education.Inc.

Tabatataei (2012). The effect of portfolio assessment technique on writing performance of EFL learners. *English Language Teaching*, *5* (5).

Tawil, M., & Liliasari (2014). *Keterampilan-keterampilan sains dan implementasinya dalam pembelajaran IPA*. Makassar: Badan Penerbit UNM.

Wijayanti, A., & Basyar, M. A. K. (2016). The development of thematic–integrated e-portfolio media web blog based to increase the scientific literacy of elementary teacher education program students. *Jurnal Pendidikan IPA Indonesia*, *5* (2), 284-290.

Yang, N. (2003). Integrating portfolios into learning strategy-based instruction for EFL college students.International. *Review of Applied Linguistics in Language Teaching*, *41* (4), 293-317.

Yancey, K. B. (1999). Looking back as we look forward: Historicizing writing assessment. *College Composition and Communication*, *50*, 483-503.

Yildirim, M., Calik, M., & Ozmen, H. (2016). A meta-synthesis of Turkish studies in science process skills. *International Journal of Environmental & Science Edcation*, *11* (14), 6518-6539.

## APPENDIX

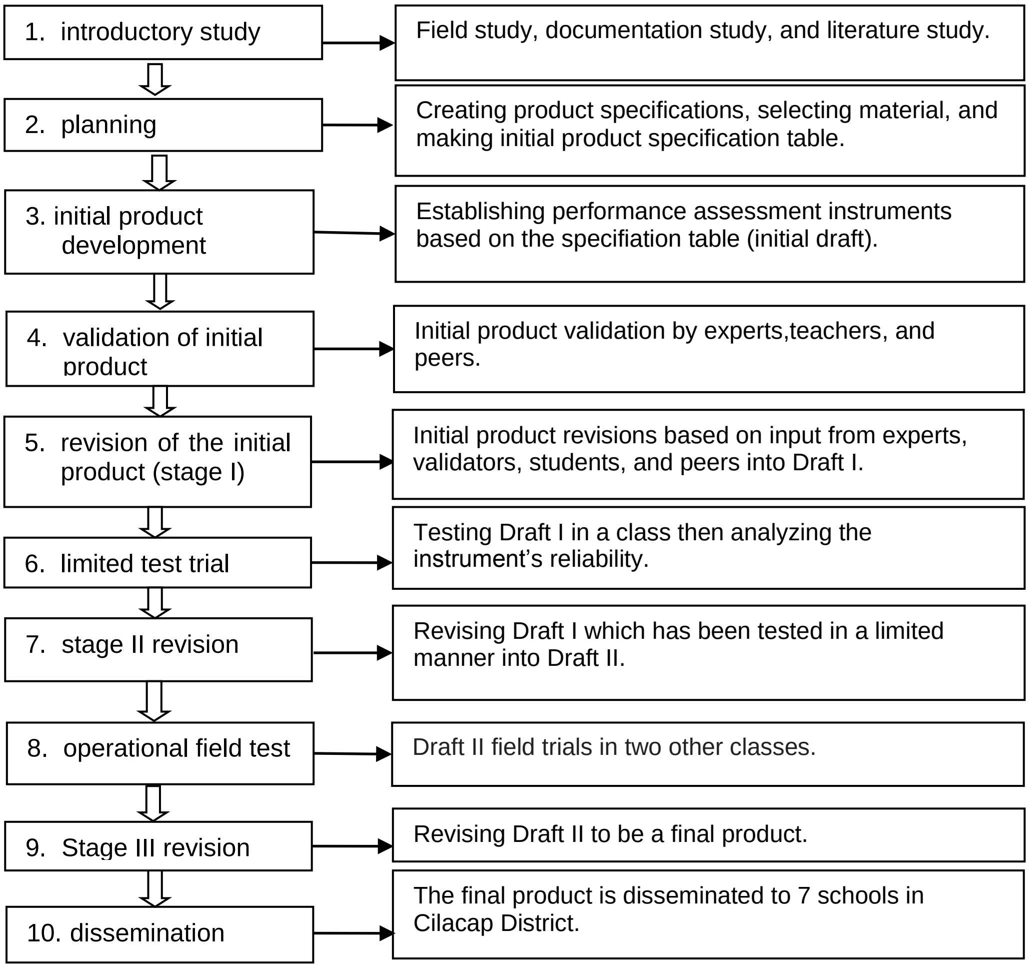


***Figure 1***. Flowchart of performance and portfolio assessments in regard to developmental procedures of the instrument(s)

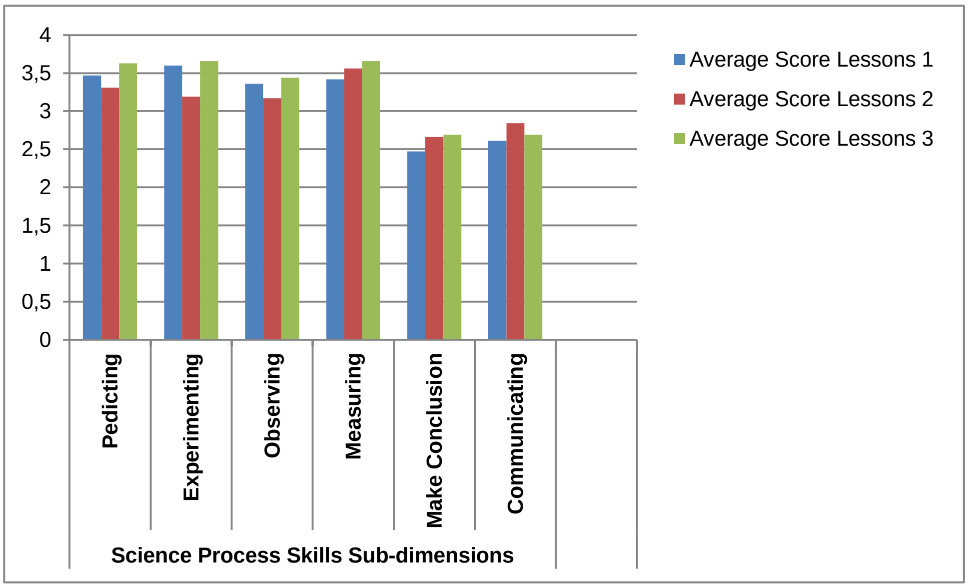


***Figure 2***. The measurement results of the SPS through the observation sheet