The Validity of Item Tests of Scientific Method of Divergent Patterns in Biology Subject at Senior High Schools Viewed From The Number of Traits

(1) Fajar Nur Cahyani, (2) Bambang Subali
(1), (2) Yogyakarta State University

Corresponding Author: fajar0973pasca2015@student.uny.ac.id

Abstract. The initial stage of the research was developing a blue print based on learning continuum, formulating divergent test item models, assigning two experts on measurement and one expert on biological education assessment for reviewing, and performing a try out in the field. The population in this research is the hypothetical population that have characteristics such as selected samples by selecting five Public Senior High Schools in Bantul Regency ranging from schools which have gained highest national examination score to that of the lowest which consist of 954 testee. The data were analyzed using the Quest program referring to the Partial Credit Model (PCM). The analysis was performed to determine the validity of test items which have been developed viewed from single, double, and triple trait. The results of the try out show that all items are fitted with PCM for single and double trait based on INFITMNSQ acceptance limits. Meanwhile, one item is not fitted with PCM for triple trait precisely the trait to investigate. When INFIT-t acceptance limit is employed, some items are not fitted. They include 10 items of single trait, 3 items of double trait precisely the trait to design and perform an investigation and 2 items of the trait to report the results of the investigation. In addition, the non-fitted items for triple trait consist of 3 items of the trait to design an investigation, 2 items of the trait to perform an investigation, and 2 items of the trait to report the results of the investigation.

Keywords: Scientific Method, Divergent Pattern, Senior High School Biology Subject, Partial Credit Model.

1. Introduction
The implementation of a learning process at schools must not ignore the curriculum. In other words, every subject at schools must apply the curriculum. The curriculum of 2013 as a valid curriculum requires the learners to be able to apply a scientific approach. This is also stipulated in the Decree of the Culture and Education Minister No. 81 A of 2013 appendix IV stating that the learning process consists of five main learning experiences such as observing, questioning, gathering information, associating, and communicating which are parts of scientific approaches.

Biology is one of the subjects taught at high schools. Thus, Biology learning must employ a scientific approach to study the phenomena of living things and their environment. Natural phenomena or problems that occur in the surrounding environment can be overcome by conducting an investigation. This is in accordance with the view of [8] stating that biology as a science cannot be
separated from the process of investigation. Through the investigation, new biological knowledge will be obtained.

There are three stages in performing a scientific method namely observation, explanation, and testing [9]. The challenge of the scientific method is to discover new things about a spectrum and set a belief about the appropriate nature [14]. The first stage of a scientific investigation is to make a list of questions [6]. The key of the scientific method is using deductive logic to provide trusts on the real and true world [11].

The scientific method should have been taught to learners since elementary school level. Through Biology learning, all learners are expected to understand the basis of investigation and creative analysis based on data which have been gathered [7]. Empirical evidence generated from the experiments is needed in Biology. They are used to support a theory or to develop an explanation [17].

The techniques of assessment can be a written, oral and performance test. Assessment techniques in the form of test are included as a measurement technique because this technique will generate numerical data which describe the characteristics of learners’ ability [23], [24]. The mastery of scientific method is the ability or performance skill therefore the instruments used for assessment is a performance assessment.

Performance assessments can be performance records such as student works and documentations of concrete as well as rich teaching practices which can be directly related to the objectives and teaching process [19]. Based on Simson’s opinion (1972), a theoretical assessment addressed to a certain procedure is classified into the aspect of the psychomotor domain in which the stages consist of mastering the theory on techniques or procedures in the form of stages of activities. The theoretical mastery can be tested through a written test.

A test which can determine the critical thinking performance of the learners consists of multiple choice, open-ended, and multi-response format test [15], [12] have developed an instrument to measure analytical thinking skills and scientific process skills through an essay test. In addition, [5] have developed a closed and open test used to measure the writing skills of researchers in general communication in science. Also, [3] has developed a test to measure the nature of (whole) science (NOS) with Knowledge of the Nature of Whole Science (KNOWS). In addition, [13] use multiple-choice tests to determine high-order thinking skills in science and equipped with other instrument analysis.

The assessment techniques which are used to know the student’s mastery on scientific methods is an assessment technique which enhances learners thinking ability to develop and to be unlimited. Those techniques can be non-objective test or a test that contains open ended question which can provide the students an opportunity to give correct answers freely and broadly. Moreover, that technique may guide the students to perform divergent thinking in finding ways / steps / procedures to get solutions for solving problems. The divergent question is an open question which has many alternative answers and solutions [21]. [20] states that a divergent thinking test which has been developed since 1950 is a test that can measure a creativity. In addition, [28] say that divergent test model is a test used to measure the creativity of learners. The application of divergent tests model is influenced by experiences [22]. Divergent thinking is an open test which enables learners to generate different responses and ideas [1]. Divergent thinking is considered as an important component of creativity, can be classified as a broad type of thinking, and enhance the production of broad ideas [4].

In order to measure the divergent thinking ability related to the mastery of the scientific method, it is necessary to develop a measurement test of the scientific method of divergent model for biology subjects in Senior High Schools. This research intends to obtain information to investigate the validity of test items if the mastery of the scientific method is viewed as single trait, double trait, and triple trait.

2. Research Methods

The initial research activity was formulating a blue print based on the learning continuum of scientific method. Then, a divergent item test model was developed. The learning continuum of scientific method is designed by referring to the learning continuum of science process skill developed by Bambang Subali in 2009.
After being reviewed by 2 experts on measurement and one expert on biological education assessment, the tests were tried out in the field. The sampling technique used in this research is a purposive sampling technique. The sample consists of 954 students of grade X, XI, and XII from 5 Public Senior High Schools in Bantul Regency ranging from the highest to the lowest-achievement students. This sample as part of the hypothetic population that have characteristics as the characteristics of the research sample. The test has the three-categories and four-categories of politomus data were analysed referring to the Partial Credit Model (PCM) using the Quest software. The analysis was intended to know the mastery of scientific method viewed from single trait, double trait, and triple trait. If the mastery of the scientific method is viewed as two traits, they are divided into two traits namely a trait to design and perform investigations; and a trait to report the results of the investigations. If the mastery of the scientific method is viewed as three traits, they are divided into a trait to arrange investigations, to perform the investigations, and to report the results of the investigation.

The scoring was performed using a politomus and analysed using Partial Credit Model (PCM). The items which are fitted with the model in the Quest program (Adam & Kho, 1996) are determined based on mean score of INFIT Mean of Square (INFIT MNSQ) in which the range is 0.77 to 1.3, whereas the range for INFIT -t is -2 to +2.

3. Research Findings

The learning continuum of scientific method is designed by referring to the learning continuum of science process skill developed by Bambang Subali in 2009. The results of item analysis using the Quest program in which the mastery of the scientific method is viewed as a single trait shows that the mean and standard of deviation of INFIT MNSQ and INFIT -t are as follows.

Table 1. Mean and standard of deviation score of INFIT MNSQ and INFIT -t of scientific method mastery of divergent model viewed as a single trait

| Descriptions | Estimation for Items |
|--------------|----------------------|
| Mean and standard of deviation score of INFIT MNSQ | $1 \pm 0.10$ |
| Mean and standard of deviation score of INFIT -t | $-0.1 \pm 2.2$ |

Table 1 shows that the standard of deviation of small INFIT MNSQ is $\pm 0.10$ within the theoretical range from 0.77 to 1.33. Meanwhile, the standard of deviation of INFIT is greater ($\pm 2.2$) within the theoretical range from -2 to +2. Thus, there are possibly more items which are not fitted with PCM when using the INFIT -t criteria.

The following is presented the detail of items which are fitted with PCM based on INFIT MNSQ and INFIT -t score.

Table 2. The comparison of Item-PCM Conformity score when Scientific Method of Divergent Pattern for Biology subject at Senior High Schools is viewed as Single Trait Based on INFIT MNSQ and INFIT -t

| Items | Conformity of Items with PCM |
|-------|-----------------------------|
|       | INFIT MNSQ | INFIT -t |
|       | Score | Note | Score | Note |
| Item 1 | 1.22   | Fitted | 4.3   | Not fitted |
| Item 2 | 1.04   | Fitted | 1     | Fitted   |
Table 2 shows that all items are fitted with PCM models based on the fitted items acceptance limit of INFIT MNSQ 0.77 to 1.30 [2]. Moreover, 10 items are not fitted based on the fitted items acceptance limit of INFIT t -2 to +2 (Keeves & Masters, 1999; Bond & Fox, 2007). They are items number 1, 5, 6, 9, 11, 15, 16, 17, 20 and 21.

The next result of the analysis deals with the mastery of the scientific method viewed as two traits namely to design and perform investigations; and to report the results of the investigations.

Table 3. Mean and standard of deviation score of INFITMNSQ and INFIT t of scientific method mastery of divergent pattern viewed as double traits (to design and perform investigations; and to report the results of the investigations)

| Description | Type of trait                        | Design and perform investigations | Report the results of the investigations |
|-------------|-------------------------------------|----------------------------------|----------------------------------------|
| Mean and standard of deviation score of INFIT MNSQ | Design and perform investigations | 1.01 ± 0.07                      | 0.99 ± 0.08                            |
| Mean and standard of deviation score of INFIT t  | Report the results of the investigations | 0.00 ± 1.6                      | -0.1 ± 1.5                             |

Table 3 shows that the standard of deviation of small INFIT MNSQ is 0.007 for first trait and 0.08 for second trait within the theoretical range 0.77 to 1.33. Meanwhile, the standard of deviation of INFIT is greater (1.6 for the first trait and 1.5 for the second trait) within the theoretical range -2 to +2. Thus, there are possibly more items which are not fitted with PCM when using the INFIT t criteria.

The following is presented the results of the test item analysis of scientific method mastery viewed as a test with two traits.
Table 4. The comparison of Item-PCM Conformity when the ability of Scientific Method of Divergent Pattern for Biology subjects at Senior High Schools is viewed as Double Trait (Design and perform investigations; and report the investigation results) based on INFIT MNSQ and INFIT t

| Items                      | Conformity of Items with PCM | INFIT MNSQ | INFIT t |
|----------------------------|------------------------------|------------|---------|
|                            | Score                        | Note       | Score   | Note    |
| a. Design and perform      |                              |            |         |         |
| investigations             |                              |            |         |         |
| Item 1                     | 1.18                         | Fitted     | 3.6     | Not fitted |
| Item 2                     | 1.02                         | Fitted     | 0.6     | Fitted |
| Item 3                     | 0.99                         | Fitted     | -0.4    | Fitted |
| Item 4                     | 0.95                         | Fitted     | -1.3    | Fitted |
| Item 5                     | 1.06                         | Fitted     | 1.2     | Fitted |
| Item 6                     | 1.11                         | Fitted     | 2.2     | Not fitted |
| Item 7                     | 1.02                         | Fitted     | 0.3     | Fitted |
| Item 8                     | 1.01                         | Fitted     | 0.2     | Fitted |
| Item 9                     | 1.02                         | Fitted     | 0.5     | Fitted |
| Item 10                    | 0.92                         | Fitted     | -1.8    | Fitted |
| Item 11                    | 0.9                          | Fitted     | -2.4    | Not fitted |
| Item 12                    | 1.01                         | Fitted     | 0.1     | Fitted |
| Item 13                    | 0.96                         | Fitted     | -1.1    | Fitted |
| Item 14                    | 1.01                         | Fitted     | 0.2     | Fitted |
| Item 15                    | 0.94                         | Fitted     | -1.8    | Fitted |
| b. Report the results of  |                              |            |         |         |
| investigation              |                              |            |         |         |
| Item 16                    | 1.12                         | Fitted     | 2.4     | Not fitted |
| Item 17                    | 0.99                         | Fitted     | -0.1    | Fitted |
| Item 18                    | 0.93                         | Fitted     | -1.1    | Fitted |
| Item 19                    | 0.93                         | Fitted     | -0.9    | Fitted |
| Item 20                    | 0.98                         | Fitted     | -0.6    | Fitted |
| Item 21                    | 0.89                         | Fitted     | -2      | Not fitted |
| Item 22                    | 1.07                         | Fitted     | 1.4     | Fitted |

Table 4 shows that all items of the first trait are fitted with PCM referring to the INFIT MNSQ score. Meanwhile, there are 3 out of 15 items that are not fitted with PCM referring to the INFIT t score. The non-fitted items consist of item number 1, 6, and 11. In addition, all items of the second trait are also fitted with PCM when they are based on MNSQ's INFIT score. Meanwhile, there are 2 out of 7 items that are not fitted with PCM namely item number 16 and 21 if they are based on the INFIT t score.

The results of the next analysis deal with the mastery of scientific methods viewed as three traits which include the trait of designing investigations, performing investigations, and reporting the results of the investigations. The comparison of mean and standard of deviation score of INFITMNSQ and INFIT-t is presented as follows.

Table 5. Mean and standard of deviation score of INFITMNSQ and INFIT-t of scientific method mastery of divergent patterns viewed as triple trait (designing investigations, performing Investigations, and reporting the results of investigation)
Considering the standard of deviation of INFIT-t which is greater than the standard of deviation of INFIT MNSQ for the three traits, the items that are possibly not fitted with PCM will be higher when referring to the INFIT-t acceptance limit.

The results of analysis on the item conformity analysed in three dimensions based on INFIT MNSQ and INFIT t score are presented in table 3.

**Table 6.** The Comparison of Item-PCM Conformity when the ability of Scientific Method of Divergent Pattern for Biology subjects at High Schools is viewed as Triple Trait (Designing investigations, performing investigations, and reporting the results of investigations) based on INFIT MNSQ and INFIT t score.
Table 6 shows that all items (12 items) for the trait of designing investigations are fitted with PCM when referring to the acceptance limit of INFITMNSQ 0.77 to 1.30. In addition, there is one out of 3 items (item 15) for the trait of performing an investigation which is not fitted. Moreover, all items (7 items) for trait to report the results of investigation are fitted with PCM model. When using INFIT-t acceptance limit ranging from -2 to +2, there are 3 out of 12 items for the first trait that are not fitted with PCM (item number 1, 10, and 11), 2 of the 3 items for the second trait that are not fitted with PCM model (item number 13 and 15), and 2 out of 7 items for the third trait that are not fitted with PCM model (item number 16 and 21).

4. Research Findings
After performing item analysis, there two findings which need to be discussed. First, the sensitivity of using t distribution (via INVIT-t) and the distribution of X2 (INVITMNSQ) is different. The t distribution is tighter which results in greater chances for the items to be unfitted with PCM models. For the sake of the practical interests, the researchers prefer to use INFITMNSQ to investigate the validity of the item because this provides higher probability for items to be fitted with the model. [24],[25],[26],[27] report that all test items are fitted with the model because they use X2 distribution.

Second, the ability of scientific method which in this case is tested using written skill test is more appropriate if viewed as single trait or double trait because all items are fitted with PCM model when X2 distribution (INFIT MNSQ) is used. However, when it is viewed as a triple trait, it is inappropriate because the trait of performing an investigation is only measured using 3 items. Therefore, further research is required considering the length of the test will affect the level of the test validity. Tests that utilize too short or few items may produce low accuracy results since the indicators are not able to describe the ability being measured.

5. Conclusions and Suggestions
Based on the research findings, the conclusions are as follows. The test results show that all items for single trait and double trait are fitted with PCM model referring to acceptance limit of INFITMNSQ while one item of triple trait namely performing investigation is not fitted. When using the INFIT-t acceptance limit, 10 items of single trait, 3 items of double trait i.e. to design and perform investigations and 2 items on the trait of reporting the results of the investigation are not fitted. In addition, 3 items of triple trait i.e. designing investigations, 2 items on the trait of performing investigations, and 2 items on the trait of reporting the results of the investigation are not fitted. With regard to this, the researchers prefer to use the acceptance limit based on the X2 distribution rather than the t distribution.

6. References
[1] Acar, S. & Runco, M.A. 2015. Thinking in Multiple Directions: Hyperspace Categories in Divergent Thinking. Psychology of Aesthetics, Creativity, and the Arts, Vol. 9, No. 1, 41-53.
[2] Adams, R.J. and Khoo, Seik-Tom, Acer Quest Version 2.1. Camberwell, Victoria, The Australian Council for Educational Research, 1996.
[3] Allchin, Douglas, Evaluating Knowledge of the Nature of (Whole) Science, Science Studies and Science Education, 2010.
[4] Antink-Meyer, A., & Lederman, N.G. 2015. Creative Cognition in Secondary Science: An Exploration of Divergent Thinking in Science Among Adolescents. International Journal of Science Education, Vol. 37, No. 10, 1547–1563.
[5] Baram-Tsabari, A., & Lewenstein, B.V. 2013. An Instrument for Assessing Scientists’ Written Skills in Public Communication of Science. Science Communication, 35(1), 56-85.
[6] Benbow, A. & Mably, C. Master, The Scientific Method With Fun Life Science Projects. United State of America: Enslow Publisher, 2010.
[7] Campanile, M.F., Lederman, N.G., & Kampourakis, K, Mendelian Genetics as a Platform for Teaching About Nature of Science and Scientific Inquiry: The Value of Textbooks. Science and Education, 2013.
[8] Campbell, Niel A., Williamson, Brad, Heyden, Robin J, Biology: Exploring Life, New Jersey, Pearson Education, 2004.
[9] Carey, S.S, A Beginner’s Guide to Scientific Method, Fourth Edition, United State of America, WADSWORTH CENGAGE Learning, 2011.
[10] Dijk, E.M.V. 2014. Understanding the Heterogeneous Nature of Science: A Comprehensive Notion of PCK for Scientific Literacy. Science Education, Vol. 98, No. 3, pp. 397–411.
[11] Gimbel, S, Exploring The Scientific Method Cases And Question, London: The University of Chicago Press, 2011.
[12] Irwanto, Rohaeti, E., LFX, E.W., & Suyanta. 2016. The Development of an Integrated Assessment Instrument for Measuring Analytical Thinking and Science Process Skills. Proceedings of the International Conference on Education, Mathematics, and Science.
[13] Ketelhut, D.J., Nelson, B.C., Clarke, J., & Dede, C. 2010. A Multi-user Virtual Environment For Building And Assessing Higher Order Inquiry Skills In Science. British Journal of Education Technology, Vol. 41, No1, 56-68.
[14] Keeves, J.P. & Masters, G.N. (1999). Introduction. In: Masters, G.N. & Keeves, J.P. (1999). Advances In Measurement In Educational Research And Assessment. Amasterdam: Pergamon, An imprint of Elsevier Science.
Kosso, P, A Summary of Scientific Method. London New York, Springer Dordrecht Heidelberg, 15 2011.
Ku, K.Y.L, 2009, Assessing Students’ Critical Thinking Performance: Urging For Measurements [16] Using Multi-Response Format. Thinking Skills and Creativity, 4 (2009), 70-76.
[17] Mardapii, Djemari, Pengukuran, Penilaian, dan Evaluasi Pendidikan, Yogyakrta, Nuha Medika, 2012.
[18] McComas, W.F. & Kampourakis, K. 2015. Using The History Of Biology, Chemistry, Geology, And Physics To Illustrate General Aspects Of Nature Of Science. Review of Science, Mathematics and ICT Education. 9(1), 47-76.
[19] Miller, M.D., Linn, R.L., & Gronlund, N.E, Measurement and Assessment in Teaching. New Jersey, Pearson, 2009.
[20] Peck, C.A., Singer-Gabella, M., Sloan, T., & Lin, S. 2014. Driving Blind: Why We Need Standardized Performance Assessment in Teacher Education. Journal of Curriculum and Instruction (JoCI), Vol. 8, No. 1, Pp. 8-30.
[21] Robinson, A, Genius: A Very Short Introduction, New York, Oxford University Press, 2011.
[22] Runco, M., Creativity Theories and Themes: Research, Development, and Practice, USA, Torrance Center for Creativity and Talent Development, 2014.
[23] Runco, M.A., & Acar, S. 2012. Divergent Thinking as an Indicator of Creative Potential. Creativity Research Journal, 24(1), 66-75.
[24] Simpson, E.J. (1972) The Classification of Educational Objectives in the Psychomotor Domain. Gryphon House, Washington DC.
[25] Subali Bambang, Prinsip Asesmen dan Evaluasi Pembelajaran Edisi Kedua, Yogyakarta, UNY Press, 2016.
[26] Subali, Bambang; Paidi, & Mariyam, Siti. 2016. The divergent thinking of basic skills of sciences process skills of life aspects on natural sciences subject in Indonesian elementary school students. Asia-Pacific Forum on Science Learning and Teaching, Volume 17, Issue 1, Article 2.
[27] Subali, Bambang & Mariyam, Siti. 2015. Measuring the Indonesian Elementary Schools Student’s Creativity in Science Processing Skills of Life Aspects on Natural Sciences Subject in Yogyakarta Special Province (DIY). Journal of Elementary Education Vol.25, No. 1 pp. 91-105.
[28] Subali, Bambang & Mariyam, Siti. 2013. Pengembangan Kreativitas Keterampilan Proses Sains dalam Aspek Kehidupan Organisme pada Mata Pelajaran IPA SD. Cakrawala Pendidikan. November 2013, Th XXXII, No. 3.
[29] Subali, Bambang. 2009. Pengukuran keterampilan proses sains pola divergen mata pelajaran biologi sma di Provinsi DIY dan Jawa Tengah. Doctorate Dissertation. Program Pascasarjana Universitas Negeri Yogyakarta. Unpublished.

[30] Zheng, L., Proctor, R.W., & Salvendy,G. 2011. Can Traditional Divergent Thinking Tests Be Trusted in Measuring and Predicting Real-World Creativity?. Creativity Research Journal, 23(1), 24-37.