Journal of Elementary Education
Vol.25, No. 1 pp. 91-105

Measuring the Indonesian Elementary Schools Student's Creativity in Science Processing Skills of Life Aspects on Natural Sciences Subject

In Yogyakarta Special Province (DIY)

Bambang Subali* & Siti Mariyam**

Abstract

This research aims at measuring the Indonesian Elementary Schools studentös creativity in science processing skills (SPS) of life aspects on Natural Sciences subject using tests of which items are fitted based on Partial Credit Model (PCM). This first stage of the research aims to get the anchor items which are used for tests in the second stage. In the first stage, the researchers developed a blue print of SPS and its items; performed expertös judgment, and tried out the test items to elementary students of grade IV and V. The findings show that one of all items is not fitted referring to PCM. In the second stage, the test items were divided into four test sets and each of test set is completed with *anchor items*. The testees were elementary school students of grade IV and V from five regencies/cities in Yogyakarta Special Province (DIY). The findings show that all items are fitted referring to PCM. In reference to log it scale, the items difficulty index of creativity for Basic Skills and Processing Skills is higher than the average ability of the studentös creativity skills. This indicates that their creativity is still low.

Keywords: Creativity, science processing skills, partial credit model

* Professor, Yogyakarta State University Indonesia Email: b_subali@yahoo.co.id
** Graduate, Yogyakarta State University Indonesia. Email: sitisudjoko@yahoo.com

Introduction

The essence of learning Natural Sciences (IPA) is to train learners to investigate natural phenomena to find scientific products through scientific process based on scientific attitude (Carin and Sund, 1989: 6). The scientific process is arranged in a particular order and is called a scientific method (Towle, 1989: 16-31). Science processing skills should be taught so that students master it. The science processing skills can be taught partially as basic skills and can also be taught in the form of integrative skills (Rezba et. All., 2007: 4). Meanwhile, Bryce, et. all. (1990: 3) divide it into basic skills, process skills, and investigative skills.

Solving the problemsto find a new product through a scientific method is a process of inquiry. All science sare inquiry. Biology belongs to sciences. Biologists try to answer questions about living things (Mayer, 1980: 3-4).

Finding a new product is a creative work. Thinkingcreatively belongs tothe highest cognitive domain in Bloom's taxonomy by Anderson & Krathwohl (2001: 68) and Dettmer (2006: 73). However, it does not mean that creativity can not be taught to elementary school students. This is supported by the ideas of Miller (2005: 65) that makes a simple definition of creative. According to him, creative is something that is not a duplication/imitation. The learners mastery on creativity as one of the aspects that should be developed need to be measured.

The main problem in measuring creativity is making sure that the aspects being measured is creativity and is not influenced by intelligence aspects (Cramond & Baer, 1994: 70). There are many researches which investigate the strategies of measuring creative thinking skills, as compiled by Kind & Kind (2007: 1-29). Moreover, the detail explanation about the existing creativity tests, including tests related to the process of divergent thinking is presented by Cropley (2000: 72-78). Added to this, Meeker (1969: 87-99) states that there are many ways and elements to be measured viewed from the aspects of how to measure the creativity. For example, one of the ways to measure divergent thinking skills can be specified based on the content and products.

According to Torrance (1974), measuring creativity is generally directed to measure divergent thinking skills. In line with Torrance, Gough (1979) states that measuring creativity means measuring creative personality. Added to this, Kelly (2004: 594-596) says that few researchers measure creativity as a multidimensional phenomenon using self-report scale.

Diakidoy & Constantinou (2000-2001) in Kind & Kind (2007: 1-29) have explored the context dependence of creativity among students by asking as many responses as possible to the three types of open-ended tasks and scored based on the divergent thinking ability in reference to Guilfordôs model which include: (a) fluency, i.e. the number of considerations within the given solutions, (b) flexibility, i.e. the number of different types of solutions, and (c) authenticity, calculated based on comparison scale: if the studentsô response is less than 5%, the score will be 3; if it is less than 15%, it will be scored 2; and if it is less than 50%, the score will be 1.

The creativity measurement of SPS at Senior high school students in Biology subject in DIY and Central Java used Item Response Theory (IRT)approach. This approach produces a calibration that places the ability of learners (person ability) and item difficulty on the same scale. Thus, they can be compared (calibration sets out to place the measurements of the person attainments and item difficulty on the same scale and uses the same units for both). The research used the IRT approach especially Rasch Model.

In this research, the results showed that the average score of creativity is far below the item difficulty index. Thus, the SPS creativity skills are low (Bambang Subali, 2010: 141). It has been reported that the development of SPS creativity related to the life aspects in Natural Sciences subjects at Elementary Schools has been done by the teachers (Bambang Vali & Siti Mariyam, 2013: 378). However, how the mastery of creativity has not been studied.

Statement of the Problem

Developing instruments for measuring creativity of Science processing skills on life aspects of Natural Sciences subjects at Elementary Schools which are fitted based on PCM and use the instruments to measure the creativity of Elementary Schools students grade IV and V.

Objectives of the Study

The objectives of this research are as follows:

1. Developing/constructing creativity tests for Science processing skills on life aspects of natural sciencs subjects at elementary school of which items are fitted with PCM in reference to IRT.
2. Presenting information about the test reliability indexs whichare used for measurementbased on the error of measurement scores.

3. Reporting the results of SPS creativity measurement of Elementary School students grade IV and V using tests which have been constructed and fitted based on PCM

Research Methods

The research was conducted in two stages. In the first stage, the researcherdevelopedthe blue print of Science Processing Skills (SPS). The formulation of SPS blue printreferred to SPS blue printsfrom the research conducted by Bambang Subali (2009) which were used for measuring creativity of SPS in Biology subjects for SeniorHigh School students. Also, the blue print refers to several sources such asRezba et.al. (2007), Bryce et al (1990), and Cox (1956). The SPSaspects consists of (a) basic skills, (b) processing skills, and (c)investigative skills. However, the researcher only used two aspects of SPS to measure elementary students' mastery on SPS namely (a) basic skills and (b) processing skills. The sub-aspects and indicatorsof basic skills and processing skills are enclosed in Appendix 1.

In reference to the blue print of SPS, the researcher developed creativity tests for SPS which consist of 63 items. After undergoing experts judgement, the items were tested to 637 students of grade V and VI using divergent scoring model. The testees were from four elementary schools in Regional Technical Implementation Unit (RTIU) Sleman Regency. The testee samples in the first stage comprise of 637 students from four elementary schools in RTIUsin Sleman regency. According toMuraki & Bock(1998: 35),the number of samples is eligible for the try outbecause ideally there must be at least 500 testees in order that a set of test can be used operationally. The objectiveof the try out is to obtain testitems which measure the creativity of SPS so that it can be used as the anchor for the testsin the second stage.

Each testee does the tests for threetimes. The tests items are essays in the form of divergent models and each testee is expected to give two correct answers for each item. The creative scoring refers to the models of Diakidoy & Constantinou (Kind & Kind, 2007) which focuses on the aspects of *fluency*and *flexibility.* Theaspect of *fluency*is based on the number of correct answers given.In this case, each item requests two correctanswers.With regard to the aspects of *flexibility* , two answers of each item given by the testees may be different. The tests are performed using a power testmodel. Testees are given time and they may stop themself in completing the tests.

In the second stage, the researcher divided the items into foursets of tests. Each consists of 20 items completed with five anchor items obtained from the first stage. With the use of anchor items, each studentonly completes one set of test but it can be mapped into a singlelogit scale. The difficulty index of anchor items from 1 to 5 in alogit scale is -0.42, -0.79, -0.82, -0.30, and -0.60.

The testee samples were taken from 10 RTIUs in five regencies/cities in DIY. TwoRTIUsfrom each regency / city was selected purposively. One of RTIUs is located in the center of government and another is located far from the center of government, except RTIUin the city of Yogyakarta because both are in the city center. Moreover, two privateElementary Schools and four public Elementary Schools from Each RTIU were selected. The test participants includedstudents of grade IV and V. There were 783 testees doing test I , 764 testees doing test II, 753 testees doing test III, and 760 testees doing test IV.

The creative scoring models of Diakidoy & Constantinou (Kind & Kind, 2007) include the aspects of*fluency*, *flexibility* and *authenticity.* The aspect of fluencyis based on the number of correct answers given and each item requests two correct answers. With regard to the aspect of flexibility, two answers of each item given by the testees may be different. Meanwhile, the aspect of *authenticity*is calculated based on thecomparison scale. In this case, if the responsesare less than 20% of all testess, the score will be 3; if it is less than 40% of all testees, it will be scored 2; and if it is less than 60% of all testees, it will be scored 1. Since each item has two correct answers, the maximum score is 6. Therefore, the items analysis uses a politomus of 7 categories scale.

The tests results were analyzed by using items analysis which utilize Quest program (Adam & Kho, 1996) and scaled with politomus 7 categories. The testing of the fitted items to the Partial Credit Model is based on the score of Infit Mean Square (Infit MNSQ) ranging from 0.77 to 1,30 (Wright & Masters, 1982). Because all correct answers for each item is independent, the nature tends to be unconditional with correct answerprobabilityof 0.5. The results of the analysis were presented in the form of item difficulty index and lowest to highest threshold score since each new occurrence increases the level of difficulty. The results of analysis also present the average ability ofthe testee (mean ability) and the ability of each thresholdsstep as well as its score. Threshold step is presented from 0 to 1, from 1 to 2 and so on until 5 to 6 as the highest score.

The instrument validity in this research referred to the validity principles of Wright & Master (1982). Regarding the use of IRT model, they state that:

> é ..the internal validity of a test can be analyzed inerms of the statistical fit of each it emto the model in a way that is independent of the sample distribution. To facilite our item fit analyses, we standardize these mean squares into fit statistics with expected means near zero and expected standard deviations near one. We use the term "valid" to refer to the success of this evaluation of fit. If the fit statistics of an item are acceptable .i.e., near zero, then we *say* the item calibration ñvalid". We also supervise the internal consistency of each personôs pattern of performance in the same way and, if the fit statistics for a person's performance acceptable, say that their measure is "valid".

The reliability in this research is calculated based on the error of measurement and internal consistency in reference to Adam & Kho (1996) which is calculated using a Quest Program.

Research Findings

After analyzing the three sets of tests given to 701 students from three elementary schools, there were only 637 students who completedthree times test based on developed tests sets. The findings obtained after the analysis using the Quest Program are presented as follows.

Table 1
The test of fitted items based on PCM model of three categories

```
--------------------------------------------------------------------------------
N = 637 L = 63 Probability Level= .50
--------------------------------------------------------------------------------
INFIT MNSQ              .56      .63       .71       .83       1.00      1.20      1.40
------------------+-----------+-----------+----------+-----------+-----------+-----------+----------
  1 item 1                             .           |     *      .
  2 item 2                             .           |      *     .
  3 item 3                             .           |       *    .
  4 item 4                             .           |        *   .
  5 item 5                             .           |      *     .
  6 item 6                             .           |   *        .
  7 item 7                             .           |           . *
  8 item 8                             .      *     |           .
  9 item 9                             .          *|           .
```

```
10 item 10    .     *      |           .
11 item 11    .            *           .
12 item 12    .           *|           .
13 item 13    .            *           .
14 item 14    .            |         * .
15 item 15    .         *  |           .
16 item 16    .      *     |           .
17 item 17    .          * |           .
18 item 18    .         *  |           .
19 item 19    .         *  |           .
20 item 20    .           *|           .
21 item 21    .         *  |           .
22 item 22    .            |   *       .
23 item 23    .            |  *        .
24 item 24    .            |  *        .
25 item 25    .           *|           .
26 item 26    .          * |           .
27 item 27    .            *           .
28 item 28    .           *|           .
29 item 29    .            *           .
30 item 30    .            |   *       .
31 item 31    .         *  |           .
32 item 32    .           *|           .
33 item 33    .           *|           .
34 item 34    .           *|           .
35 item 35    .          * |           .
36 item 36    .        *   |           .
37 item 37    .        *   |           .
38 item 38    .            |*          .
39 item 39    .           *|           .
40 item 40    .            *           .
41 item 41    .            *           .
42 item 42    .           *|           .
43 item 43    .           *|           .
44 item 44    .            |  *        .
45 item 45    .           *|           .
46 item 46    .            |   *       .
47 item 47    .            | *         .
48 item 48    .            |      *    .
49 item 49    .            |      *    .
50 item 50    .         *  |           .
51 item 51    .          * |           .
```

```
 52 item 52              .        *        .
 53 item 53              .        *        .
 54 item 54              .       |*        .
 55 item 55              .       |*        .
=====================================================================
---------------------------------------------------------------------------------
INFIT MNSQ    .56    .63    .71    .83   1.00   1.20   1.40
-----------------+---------+---------+---------+---------+---------+---------+--------
56 item 56              .        | *        .
 57 item 57              .      * |         .
 58 item 58              .       | *        .
 59 item 59              .       |    *     .
 60 item 60              .       |*         .
 61 item 61              .     *  |         .
 62 item 62              .       *|         .
 63 item 63              .     *  |         .
=====================================================================
```

Table 1 shows that one of 63 items is not fitted with PCM model of the three categories based on Infit MNSQ score. That item is number 7 with Infit MNSQ score of 1.33. However, the Infit t score is 1.4. This means that the item is still fitted with a model since the score limit of Infit t is ± 1.96 (alpha 5%) and even can be rounded to ± 1.96 (Bond & Fox, 2007: 43). After checking the item, there is a wrong construction of the sentence in the item.

Because each item is fitted with the model, it can be stated that the tests developed using a divergent answer model can be declared "valid" therefore the tests are categorized as reliable. This is supported by the data that the reliability of estimate from persons estimate is 0.90. Moreover, in reference to internal consistency calculation, it is 0.88. This score still meets the high level of reliability since it is far above the limit or 0.7.

In the second stage, after all items were divided into four test sets completed with five anchor items, test I was tested to 783 testees, test II was tested to 764 testees, test III was tested to 753 testees, and test IV was tested to 760 testees. The results of fit analysis based on PCM are presented in Table 2. Table 2 shows that all items are fitted based on PCM because the Infit MNSQ score is in the range of 0.77 - 1.30 based on the criteria determined by Adam & Kho, 1996).Because each item is fitted with the model, it can be stated that the tests developed using a divergent answer model can be declared "valid" (Wright & Master, 1982) therefore the testsare also reliable. This is supported by data that the reliability of estimate from persons estimate is0,62. Moreover, in reference to internal consistency calculation, it is 0.88. This score still meets the high level of reliability since it is far above the general requirement limit from 0,3 to 0,7.The results of the second test are presented completely. They includelearners ability index of the SPS aspects and the item difficulty index which is completed with threshold index.

Table 2

The testing of the fitteditems based on PCM using 7 categories completed with PersonAbility Index for Each Basic Skills Aspect of SPS and item difficulty indexwith their threshold index

A. Basic Skills

	Mean Ability	Mean difficulty	Threshold 0 up to 1	1 up to 2	2 up to 3	3 up to 4	4 up to5	5 up to 6	Infit MN SQ	Total Indicators	Items
1. Observing skills											
Mean	-0.71	-0.14	-0.25	-0.24	-0.23	-0.08	0.00	0.06		7	7
Minimum	-0.82	-0.67	-0.52	-0.52	-0.49	-0.39	-0.39	-0.39	0.97		
Maximum	-0.65	0.49	-0.03	-0.02	-0.02	0.20	0.30	0.30	1.13		
2.Data/information recording skills											
Mean	-0.74	0.06	-0.08	-0.05	-0.03	0.10	0.06	0.22			
Minimum	-0.80	-0.46	-0.38	-0.38	-0.38	-0.37	-0.37	-0.37	0.97	9	9
Maximum	-0.68	0.46	0.19	0.28	0.54	0.62	0.77	0.89	1.05		
3.Following instruction skills											
Mean	-0.72	-0.09	-0.16	-0.16	-0.15	0.04	0.04	0.09		4	4
Minimum	-0.76	-0.28	-0.24	-0.24	-0.21	-0.08	-0.08	-0.02	0.96		
Maximum	-0.62	0.09	-0.02	-0.02	-0.02	0.20	0.20	0.12	1.02		
4.Classifying skills											
Mean	-0.70	-0.08	-0.35	-0.35	-0.25	-0.05	0.11	1.19		2	2
Minimum	-0.71	-0.51	-0.56	-0.56	-0.45	-0.09	-0.09	0.11	0.93		
Maximum	-0.69	0.36	-0.13	-0.13	-0.05	0.01	0.31	2.27	0.96		
5.Measuring skills											
Mean	-0.71	-0.14	-0.20	-0.19	-0.17	0.04	0.06	0.19		8	
Minimum	-0.83	-0.73	-0.72	-0.67	-0.57	-0.31	-0.23	-0.13	0.97		
Maximum	-0.62	0.12	0.02	0.02	0.02	0.22	0.22	0.22	1.13		
6.Movement manipulation skills											
Mean	-0.64	0.43	0.36	0.36	0.36	0.49	0.49	0.49		4	4
Minimum	-0.70	0.22	0.13	0.13	0.13	0.26	0.26	0.26	0.96		
Maximum	-0.57	0.75	0.65	0.65	0.65	0.86	0.86	0.86	1.02		
7.Procedure/techniques/tools usage implementation skills											
Mean	-0.63	0.00	-0.10	-0.09	-0.05	0.08	0.11	0.26		14	14
Minimum	-0.83	-0.73	-0.27	-0.27	-0.27	-0.19	-0.19	-0.19	0.95		
Maximum	-0.62	0.12	0.25	0.25	0.25	0.32	0.4	0.98	1.05		

Table 2 shows that all items which measure the creativity of basic skills sub aspects of SPS are fitted based on the PCM model. Also, itshows that all testees ability index (person ability index) for all sub-aspects of SPS on the basic skills is lower than items difficulty index. This indicates that the studentøs creativity mastery related to the basic skills of SPS is still low.

Table 3

The testing of Fitted Items in reference to PCM with 7 categories completed with the PersonAbility Index for Each Processing Skills Aspect of SPS and Item Difficulty Index) with their threshold index

II. Processing skills											
	Mean Ability	Mean difficulty	Threshold 0 up to 1	1 up to 2	2 up to 3	3 up to 4	4 up to5	5 up to 6	Infit MNSQ	Total Indicators	Items
1. Inferencing skills											
Mean	-0.70	0.07	-0.07	-0.04	-0.03	0.08	0.26	0.33		7	9
Minimum	-0.82	-0.67	-0.31	-0.31	-0.31	-0.26	-0.26	-0.26	0.97		
Maximum	-0.65	0.49	0.35	0.35	0.35	0.44	0.77	0.77	1.09		
2. Predicting skills											
Mean	-0.71	0.02	-0.29	-0.22	-0.15	0.08	0.17	0.44		4	4
Minimum	-0.70	0.22	-0.98	-0.74	-0.58	-0.35	-0.18	0.08	0.97		
Maximum	-0.57	0.75	0.23	0.23	0.23	0.54	0.54	0.57	1.09		
3. Selecting procedures skills											
Mean	-0.66	-0.07	-0.19	-0.17	-0.15	-0.03	0.09	0.09		4	4
Minimum	-0.70	0.22	-0.49	-0.49	-0.49	-0.33	-0.33	-0.33	0.99		
Maximum	-0.57	0.75	0.03	0.03	0.05	0.12	0.59	0.59	1.03		

Table 3 shows that all items which measure the creativity of processing skills sub aspects of SPS are fitted based on the PCM model. Also, it shows that all testees ability index (person ability index) for all sub-aspects of SPS on the processing skills is lower than items difficulty index. This indicates that the studentsô creativity mastery related to the processing skills of SPS is still low.

The mean score and standard deviation creativity of SPS on life aspects of the students grade IV and V in science subjects in Elementary Schools in 10 RTIUs DIY are presented in Table 5.

Table 4

The mean scores and standard deviation of KPSAK creativity grade IV, natural sciences subjects at Elementary Schools in 10 UPTDs in DIY and various test sets being tested.

Elementary School	N	Raw Scores				Maximum score		Estimation Skor			
		Min	Max	$\bar{x}$	S	Raw	Estimation	Min	Max	$\bar{x}$	S
Grade IV	1548	0	84	35.31	16.68	120	>3.20	<-1.95	-0.34	-0.66	0.21
Grade V	1512	0	84	41.49	15.01	120	>3.20	<-1.95	-0.34	-0.59	0.14

Notes: 1) Students of grade IV who obtained a 0 score consist of 11students while grade V consists of 2 students

2) The estimation of the log it scores toward the raw score of 1 is -1.95andtoward the maximum raw score of +3.20 is 383. Therefore, 0 < -1.95 and 384 > + 3.20.

Table 4 shows that the average raw scores and SPS creativity logit scale scores of grade IV are below grade V. In reference to the average scores, it shows that the results are still far lower than the maximum score.

Discussions

The research findings show that the average creativity ability of SPSon life aspects of the Elementary Schoolsstudents grade IV and V is low. This is opposed to the statement of teachers based on research conducted byBambang Subali& Mariyam (2013). Most of the teachersstated that they had taught creativity to their students like the model.This conditionmay be caused by the teachers who do not know well how to develop studentǿs creativity.Ideally, creativity learning must usean applied learning model and an ideationallearning model (Dettmer, 2006: 70-78). Teachers can also apply the principle of SCAMPER to stimulate children to be creative. They include (a) substituting / replacing, (b) combining, (c) adapting, (d) modifying, adding, (e) putting something for another use, (f) eliminatinng or reducing and (g) reconstructing or reversing (ichalko, 2000: 18-21)

The second possibility is that the teachers olny focus on understanding concepts as a target in teaching.Ideally,teachers mustbe ableto balance between standard-based learning and creativity-based learning because both are contrast. Creative thinking isclearly separated from the sequential thinking. Meanwhile, analytical thinking abilityis associated with the standard and traditional education. If one of those components is apllied, the balance in the style of thinking may be affected. Determining how to integrate creativity into a standards-based system needs to consider the learning needs of talented students (Burke, 2007: 58-63).

The third possibility is that teachers do not realize that the development of creativity in natural sciences teaching aims at directing learners to perform discovery or inquiry openly or do the tasks related to the investigation. Therefore, learners undertake creative activities as performed by scientists when doing a scientific research. Cognitive approach states that learning can adjust to develop creative thinking patterns. The students of Natural Sciences subjects area 'simple thinker'. Therefore, they may use any scientific process in a way that is different from the scientists (Kind & Kind, 2007: 1-37). However, its application in elementary schools is not easy because elementary schools students are mostly less potential. Therefore, the teachers may focus on teaching stdudents in order to understand the concepts. In this case, teachers develop studentsô convergent thinking skills. Teachers will rarely give questions which require divergent answers. Croom & Stair (2005: 12-14) state that the divergent questions are questionsthat would not require yes or no answer. Questions that begin with words like: "why", "how", "what do you think", etc which will provide many possible answers. Thus, students will answer these questions with many possible answers as the characteristic divergent thinking. However, there are still many elementary school children who perform convergent thinking skills based on their mental development which are in the concrete phase.

The teacherôs worry not to teach creativity to their students who have low academic potential may not be an excuse. Added to this, smart students are not always creative. In reference to the research findings of Ferrando et. al (2005: 21-50), there is a low correlation between creativity and intelligence. In other words,learners who have high IQ are not always more creative. According to Cromie (2007: 1), not all studies show a correlation between studentsô IQ and creativity. Some studies show that the improvement of creativity in line with the increase of IQ up to 120. Kyung Hee Kim (2005: 1) states that the results of a meta-analysis of 447 correlationcoefficient show that many creativity test scores have nothing to do with IQ scores, but many others do. Moreover, Rawat, et.al. (2012: 264-275) argues that the development of creativity is closely related to the development of skills to form a corresponding consideration in different situations. Therefore, the development of creativity should be taught as early as possible to students.

Closing

In reference to the research findings, it can be concluded that the instruments for measuring SPS creativity of life aspects which had been developed and tested to 2030 testees in 10 RTIUs in the DIY province show that all items are fitted with the model of PCM. The average ability scores of the students were still below the items difficulty levels. The recommendation given by the researchers is that it requires efforts to improve teachers ability in teaching SPS creativity of life aspects to students.

References

Adams, R.J. & Kho, Seik-Tom. (1996). *Acer quest version 2.1.* Camberwell, Victoria: The Australian Council for Educational Research.

Anderson & Krathwohl. Ed. (2001). *A taxonomy for learning, teaching, and assessing*. New York: Addison Wesley Longman, Inc.

Bambang Subali (2011). *Pengukuran Kreativitas Keterampilan Proses Sains dalam Konteks Assessment For Learning.* Cakrawala Pendidikan, Ed. XXX, No. 1, February 2011. p. 130-144.

Bond, T.G. & Fox, Ch.M. (2007). *Applying the rasch model*: *Fundamental measurement in the human sciences*. 2-nd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Bryce, T.G.K., McCall, J., MacGregor, J., Robertson, I.J., & dan Weston, R.A.J. (1990). *Techniques for assessing process skills in practical science: Teacher's guide*. Oxford: Heinemann Educational Books.

Burke, A.A. (2007). *The benefits of equalizing standards and creativity: discovering a balance in instruction*inGifted Child Today, 30, 1, p. 58-63 (retrieved in 29 October 2007).

Carin, A.A. dan Sund, R.B. (1989). *Teaching Science Through Discovery*. Columbus: Merrill Publishing Company.

Chiapetta, E. L. (1997). Inquiry-Based Science, in *The Science Teacher*. Oct 1997; 64, 7. ProQuest Education Journals, 22-26.

Cochran, S.M. & Lytle, S.L. (2006). *Troubling images of teaching in no child left behind* in *Harvard Educational Review*. Cambridge: Winter 2006. Vol. 76, Iss. 4; pp. 668-700 (Retrieved in 19 August 2007)

Cox, D.R. (1958). *Planning of experiments*. New York: John Wiley & Sons, Inc.

Cromie, W.J. 2007. *Creativity Tied to Mental Illness*: *Irrelevance Can Make You Mad,* (Online), (http://www.news.harvard.edu/gazette/...reativity.html, Retrieved in 29 January 2009).

Croom, B. & Stair, K. (2005). Getting from Q to A: Effective questioning for effective learning dalam The Agricultural Education Magazine, 78, 1, 12-14 (Retrieved in19 August 2007)

Dettmer, P. (2006). New Blooms in Established Fields: Four Domains of Learning and Doing [Versi elektronik]. *Roeper Review*, 28, 2, 70-78.

Ferrando, M., Prieto, M.D., Ferrandiz, C. & Sanchesz, C.(2005). Intelligence and Creativity. *Electronic Journal of Research in Education*, ISSN: 1696-2095, 7, 3(3): 21-50.

Kim, K. H. (2005). Can only intelligent people be creative? A meta-analysis. *The Journal of Secondary Gifted Education,* (16),(2-3): 57-66.

Kind, P. M. & Kind, V. (2007). Creativity in science education: Perspectives and challenges for developing school science [Versi elektronik]. *Studies in Science Education*, 43, 1-37.

Michalko, M. (2000). *Four steps toward creative thinking*in The Futurist; May/Jun 2000; 34, 18-21; ProQuest Education Journals.

Miller, J.L. (2005). *Mind magic: How to develop the 3 components of intelligence that matter most in today's world*. New York: McGraw-Hill.

Muraki, E. & Bock, R.D. (1998) *Parscale: IRT item analysis and test scoring for rating scale data.* Chicago: Scientific Software International, Inc.

Pollman, J., Uprichard, E., Malone, U., & Coop, R. (1973). *Multivariate Analysis of The Relationship Between Creativity and Intelllegence*. A Paper presented at annual meeting of American Educational Reserach Association, New Orleans, Lousiana, February 25-March 1, 1973.

Rawat, K. J.; Qazi, W. H. S. (2012). Creativity and education dalam. *Academic Research International 2*.2 (Mar/Apr 2012): 264-275.

Rezba, R.J., Sparague, C.S., Fiel, R.L., Funk, H.J., Okey, J.R., & Jaus, H.H. (1995). *Learning and assessing science process skills*. 3rd ed. Iowa: Kendall/Hunt Publishing Company.

Richert, A.E. (2002). Narratives that teach: Learning about teaching from the stories teachers tell. In: Lyons, N. & LaBoskey, V.K. (2002). Narrative inquiry in practice advancing the knowledge of teaching. New York: Teachers College Press.

Sheppard, B; Canning, M., Tuchinsky, M, & Campbell, C. (2006). Discovering creative solutions to everyday challenges. Chicago: Dearborn Trade Publishing. A Kaplan Professional Company.

The Decree of National Education and Culture Minister of the Republic of Indonesia No 64 (2013) about standard of content of basic and secondary education.

The Decree of National Education and Culture Minister of the Republic of Indonesia No 22 year 2006 about standard of content of basic and secondary education.

Torrance, E.P. (1979). Three stage model for teaching for creative thinking. Dalam: Lawson, A.E. *The psychology of teaching for thinking and creativity*. Columbus: ERIC.

Towle, A. 1989. *Modern biology*. Austin: Holt, Rinehart and Winston.

Williams, J. (2013). Science-Creativity is all in the mind: Resources dalam *The Times Educational Supplement* 5030 (Feb 8, 2013): 43

Wright, B.D. & Masters, G.N. (1982). *Rating scale analysis*. Chicago: Mesa Press.

Zahida, A. S. & Khalid, M. .K. (2014).Development of an achievement test in the subject of health and physical education at intermediate level in Pakistan. *Journal of Elementary Education,*24(I) 61-70.