The comparison between science virtual and paper based test in measuring grade 7 students’ critical thinking

P H Dhitareka*,
H Firman² and L Rusyati¹

¹ International Program on Science Education, Faculty of Mathematics and Science Education, Universitas Pendidikan Indonesia, Bandung, Indonesia
² Department of Chemistry Education, Faculty of Mathematics and Science Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

*Corresponding author’s e-mail: putra.habib.d@gmail.com

Abstract. This research is comparing science virtual and paper-based test in measuring grade 7 students’ critical thinking based on Multiple Intelligences and gender. Quasi experimental method with within-subjects design is conducted in this research in order to obtain the data. The population of this research was all seventh grade students in ten classes of one public secondary school in Bandung. There were 71 students within two classes taken randomly became the sample in this research. The data are obtained through 28 questions with a topic of living things and environmental sustainability constructed based on eight critical thinking elements proposed by Inch then the questions provided in science virtual and paper-based test. The data was analysed by using paired-samples t test when the data are parametric and Wilcoxon signed ranks test when the data are non-parametric. In general comparison, the p-value of the comparison between science virtual and paper-based tests’ score is 0.506, indicated that there are no significance difference between science virtual and paper-based test based on the tests’ score. The results are furthermore supported by the students’ attitude result which is 3.15 from the scale from 1 to 4, indicated that they have positive attitudes towards Science Virtual Test.

1. Introduction

Computer has bring our civilization into a state where information and technology have become the world’s biggest addiction [1]. The interest in development of computer-based test in educational assessment in schools and other educational institutions has increased in recent past years [2]. Executing assessments by using computer in the process is becoming more prevalent in educational assessment system that reflects the changes in pedagogical systems [3].

Paper-based Test is a conventional test type that commonly implemented in our educational system in recent years that require huge cost, plenty of time, and big efforts to assess the results done in a large scale [4]. Reducing paper consumption will indirectly reduce greenhouse gases and energy consumption that 40 reams of paper are equal to 1.5 acres of pine forest absorbing carbon for a year and each ream of paper is equal to roughly 12 pounds of carbon dioxide from the atmosphere [5]. Meanwhile, Computer-based testing is a green strategy used to reduce paper consumption that tons of paper would be saved if schools, universities and educational institutions were to replace paper-pencil testing with computer-based testing [6].
For instance, a research revealed that students more enjoy the computer-based test more than the paper-based test and they were more motivated to perform computer-based test repentance than another paper-based test [7]. Moreover, recent researches also show advantages of computer-based test for the students as well as for the lecturers [8]. Critical thinking is basic form of thinking used in problem solving scenarios and uses knowledge acquired over a lifetime in a manner that is logical. This form of thinking demonstrates the skills dimension of critical thinking [9].

With dynamic visuals, sound and user interactivity as well as flexibility to individual test-takers and real-time score reporting, computer-based testing expands testing feature beyond the limitations of traditional paper and pencil tests [10]. Thus, computer giving more eligible environment to assess critical thinking of the students. Students’ has their own capabilities with their own uniqueness and abilities. Schools often measure their capabilities by measuring their intelligenes the test that are equal for all the students regarding their content-related capabilities [11]. Students are gifted with multiple intelligences to support the teaching learning process and also to support their development as learner and also human in social life [12]. Some studies revealed that there is a significant difference between the two testing modes on test scores between computer-based testing and paper-based testing, while other studies reported opposite or inconsistent results [13]. This may lead into a sign that a further research is need to be done to make clear about computer and paper-based assessment that in this case the research used Science Virtual Test as an innovative computer-based assessment.

2. Methods
This research used quasi experimental method. Research that cannot fulfil the requirements as true experiment is categorized into non-experiment or quasi experiment [14]. It is in line with this research that still exposed to the threat of internal validity, as it only use one group. The design of the research is within-subjects design.

The location of this research was one of public junior high school in Bandung. The school used Kurikulum 2013 as the main curriculum. The population of this research was all 7th grade students. The sample was 71 students in two classes in seventh grade consist of 33 male and 38 female students. The sampling used is cluster sampling because the sample takes as a group rather than as an individual.

3. Results and Discussion

3.1. General comparison between Science Virtual and Paper-Based Test
This section, the data of SVT and PBT score are normally distributed and has the same variance. Thus, paired-samples t test is used to analyze the score data. The result of the statistical test is shown on Table 1.

| Test Mode | Mean  | SD    | p-Value |
|-----------|-------|-------|---------|
| SVT       | 52.56 | 13.62 | 0.506   |
| PBT       | 51.46 | 12.18 |         |

Table 1 shows that the p-value is 0.506 which is higher than 0.05 as the significance level. Thus, it can be inferred that there is no significance difference between science virtual and paper-based test based on test score. Not as expected, SVT is expected to have higher result than PBT because the assumption that information given from SVT with dynamic content, interactivity, and other features will help the students more to answer the question rather than the information provided from the paper that only offer a static text or picture information. From the point of view to assume SVT as CBT, this research in line with some similar research which resulted that there are no significance difference between the test mode, and assumed that the tests measure the same construct, increasing the reliability and equality of both test modes [13,15].
3.1.1. Comparison based on test items’ type. Because the data is not normally distributed but has the same variance, thus, non-parametric test which is Wilcoxon signed ranks test is used to analyse the data. The results of the test are described on the Table 2.

Table 2. Comparison based on test items’ type

| Test Item’s Type                  | SVT | PBT | Asymp. Sig. |
|----------------------------------|-----|-----|-------------|
| Video (SVT) vs. Narration (PBT)  | 9.69| 2.46| 9.83        |
| Video (SVT) vs. Picture (PBT)    | 0.77| 0.65| 1.01        |
| Picture (SVT) vs. Picture (PBT)  | 2.92| 1.21| 2.87        |
| Article (SVT) vs. Article (PBT)  | 1.01| 0.72| 1.00        |

Table 2 shows that Asymp. Sig. of Video (SVT) vs. Picture (PBT) is 0.017 and it is smaller than the α-value which is 0.05. Thus, it can be inferred that there is significance difference between science virtual and paper-based test based on test’s item type which is on the video vs. picture. However, the finding is there is no significance difference and the video vs. narration instead showing higher result on narration test item’s type while video vs. picture result is as expected that SVT result higher and show significance differences.

The result of video vs. narration expected to be related with some assumptions. The assumption is that video content in SVT give less clear information rather than the narration. Perhaps, it is caused because the video content is too long and contain too many additional/distractor information, while students only get a limited chances to get access to the video because of the limited time. In the other side, narration text contain a steady information that students can read desired information rather than listen and watch the whole content in the fixed time frame.

Furthermore, the other result from the test item’s type (picture vs. picture and article vs. article) only result a slight different that it shows no significance differences. It is may inferred that indeed test modes are not influence students result whether on paper media or computer mediated, because both of the test has the same test construct and contain the exactly same questions and information.

3.1.2. Comparison based on critical thinking elements. In this section, the data is not normally distributed but has the same variance. Thus, non-parametric test is conducted to analyse the data. The statistical test used on this data is Wilcoxon signed ranks test. The result of data processing is shown on Table 3.

Table 3. Comparison based on critical thinking elements

| Critical Thinking Elements  | SVT | PBT | Asymp. Sig. |
|-----------------------------|-----|-----|-------------|
| Element 1: purpose          | 4.92| 1.52| 4.99        |
| Element 2: question at issue| 1.23| 0.93| 1.21        |
| Element 3: assumption       | 0.34| 0.51| 0.28        |
| Element 4: point of view    | 1.68| 0.84| 1.46        |
| Element 5: information      | 1.70| 0.99| 1.66        |
| Element 6: concepts         | 1.83| 1.10| 1.85        |
| Element 7: interpretation and inference | 1.41| 0.65| 1.37        |
| Element 8: implication and consequences | 1.92| 0.84| 2.03        |
Table 3 shows that the Asymp. Sig. values are all higher than α-value that is 0.05. Thus, there are no significance differences between Science Virtual and Paper-Based Test on critical thinking elements’ score. From table 4, the result seems to be not in pattern that the score are not different significantly that in the first of all the general score is not differ significantly. This result shows that SVT and PBT is equal in mediate the students to do the critical thinking. As expected from the previous result, with the same construct the result of both of the test will result no different.

3.2. Comparison of science virtual and paper-based test based on Multiple Intelligences

The scores of students on SVT and PBT are not normally distributed but has the same variance. Thus non-parametric test is applied to the data which is Wilcoxon signed rank test. The Result of the statistical test are presented on Table 4.

| Multiple Intelligences | SVT | PBT |
|------------------------|-----|-----|
| Naturalist             |     |     |
| Musical                |     |     |
| Logical                |     |     |
| Existential            |     |     |
| Interpersonal          |     |     |
| Kinaesthetic           |     |     |
| Verbal                 |     |     |
| Intrapersonal          |     |     |
| Visual                 |     |     |

Table 4 shows that the p-values are all higher than α-value that is 0.05. Thus, there are no significance differences between science virtual and paper-based tests’ score based on Multiple Intelligences. The expectation on this section, are some of the students with some certain Multiple Intelligences would have more advantageous in doing the test because of some features that the Science Virtual Test provided. The Multiple Intelligences that are most expected to have more advantageous on Science Virtual Test is the visual and logical because it is described that visual intelligence has ability to be more aware of something that seen by the eyes and visualize it into their mind, this is in line with the feature of Science Virtual Test that provides rich images and video contents [12]. While it is also described that logical students has the ability as it is need to be a scientists, which are good on science things and also mathematics. However, the findings in this research is that in general, there are no significance difference among the students’ intelligences group whether on Science Virtual or Paper-Based Test. This findings also support that the SVT and PBT has the equivalence of the test construct and do not differ between the samples of the Multiple Intelligences’ groups.

3.3. Comparison between science virtual and paper-based test based on gender

The scores data of science virtual and paper-based test based on gender is normally distributed and has the same variance. Paired-samples t-test as parametric test is used to analyse the data. The result of the statistical test was presented on Table 5.

Table 5 shows that the p-value are higher than α-value which is 0.05. Thus, there is no significance difference between science virtual and paper based tests’ score based on students’ gender. Regarding gender, this research result difference between males and females on SVT and PBT in general where females shows higher scores on both test modes. But, it has to be underlined that there are no significance difference statistically between science virtual and paper-based test for male or females on their result respectively. Some research resulted that female students reported have lower self-efficacy and high
anxiety in science class, but they earned higher grades than males, while males reported stronger self-efficacy and did not report anxiety but scores lower than females [16], so that this research is in line with that findings. It is assumed that gender may cause the difference results, but considering about the test modes, science virtual and paper-based test has no influence on gender, females is remain achieve the higher score and male remain achieve the lower score. This may be inferred that whether male or female not influenced by SVT nor PBT so that both test modes is not bias on gender.

Table 5. Comparison based on Gender

| Gender | SVT | PBT |
|--------|-----|-----|
|        | Mean | SD  | Mean | SD  | p-Value |
| Male   | 48.8 | 12.71 | 48.06 | 11.53 | 0.758 |
| Female | 55.8 | 13.71 | 54.42 | 12.10 | 0.540 |

3.4. Students’ attitude towards Science Virtual Test

The data about students’ attitude were gathered by Likert scale questionnaire consist of four attitude scale with a certain score, which are strongly agree (4), agree (3), disagree (2), and strongly disagree (1). The questionnaire consist of 15 positive statements and 3 negative statements. The range of students’ attitude are classified into three classification, which are from 1 to 2 as negative, 2.1 to 3 as moderate, and from 3.1 to 4 as positive. The result of the questionnaire was presented on Table 6.

Table 6. Students’ attitude questionnaire result

| Aspects | Score | Average |
|---------|-------|---------|
| Experience | 3.1 | |
| Preference  | 3.2 | 3.15 |
| Technical    | 3.1 | |
| Media       | 3.2 | |

Table 6 shows that the values are 3.1 on experience and technical and 3.2 on preference and media then the average is 3.15 which are make all the values fell into 3.1 to 4 classification. Thus, the students have positive attitudes towards Science Virtual Test in general and also on all aspects. The two most positive aspects are preference and media aspects and the lower are experience and technical aspects. The implementation of Science Virtual Test on the date is using a single projector and speaker connected to one laptop, due to the circumstances of the computer lab at the school which is still under maintenance. Some of the features of Science Virtual Test is demonstrated on the front of the class. The features mentioned are technical use, administration (filling name and class), and getting the automate result after the finish the test. On science attitude towards SVT, it is showed that students show positive attitude towards SVT with the average score of 3.15. Nevertheless, if we investigate it in detail per aspects, it is also shows that students have positive attitude towards Science Virtual Test. On experience aspects, the score is 3.1 means that the students agreed that they have positive experience on Science Virtual Test. Next is preference aspects shows that the score is 3.2 means that the students also reported that they positive attitude towards implementation of Science Virtual Test and more prefer Science Virtual Test rather than paper-based test, this is in line with the positive experience that they have. Technical aspect also shows a very good respond from the students that the score show 3.1, indicate that the students shows positive attitude towards Science Virtual Test on the technical aspects. This is indicated they were enjoyed using Science Virtual Test in the class. From the media aspects, the score is 3.2 means that students agree that the media is already good and acceptable, this means that the Science Virtual Test is well accepted by the students in the term of the contents and visual appearance as well as the user interface which is also makes it more reasonable to implement SVT on junior high school students.
4. Conclusion
This research about comparison of Science Virtual and Paper-based Test towards students’ critical thinking on seventh grade secondary school has been conducted accordingly as the theoretical foundation on research in education. In general, there is no significance difference between science virtual and paper-based tests’ score and also the critical thinking elements. However, Science Virtual Test shows higher scores than paper-based test but very in small number differences, but there is a difference on the picture test item type rather than the video test item’s type, which is still need to be investigated in the next research. About the Multiple Intelligences, There are no significance differences between science virtual and paper-based tests’ score based on students’ Multiple Intelligences. There is also no significance difference between science virtual and paper-based tests’ score based on students’ gender. Lastly, students have positive attitudes towards Science Virtual Test in all aspects with the average of 3.15 from the scale from 1-4, and teachers also have a good perspective on computer-based test (including Science Virtual Test) to be implemented on school as a common assessment tool. Though Science Virtual Test resulted no different score to compare with PBT, but students have more positive attitude toward Science Virtual Test.

5. References
[1] Chand P, Murthy P, Gupta V, Kandasamy A, Jayarajan D, Sethu L and Arora S 2014 Technology Enhanced Learning in Addiction Mental Health: Developing a Virtual Knowledge Network: NIMHANS ECHO In Technology for Education (T4E) IEEE Sixth International Conference 229-232
[2] Valenti S, Cucchiarelli A and Panti M 2002 Computer based assessment systems evaluation via The ISO9126 quality model Journal of Information Technology Education 3 157-175
[3] Tzu K C C 2012 Elementary EFL Teachers’ computer phobia and computer self-efficacy in Taiwan TOJET: The Turkish Online Journal of Educational Technology 2 1-8
[4] Csapó B, Molnár G and Nagy J 2014 Computer-based assessment of school readiness and early reasoning Journal of Educational Psychology 106 639
[5] DeRosa J 2007 The Green PDF: Reducing Greenhouse Gas Emissions One Ream at A Time http://www.scribd.com
[6] Piaw C Y 2012 Replacing paper-based testing with computer-based testing in assessment: are we doing wrong? Procedia-Social and Behavioral Sciences 64 655-664
[7] Haahr J H and Hansen M E 2006 Adult Skills Assessment in Europe: Feasibility Study http://www.voced.edu.au
[8] Arachchi S M, Dias K and Madanayake R S 2014 A Comparison between evaluation of computer based testing and paper based testing for subjects in computer programming International Journal of Software Engineering & Applications 5 57
[9] Facione P A 2000 The disposition toward critical thinking: its character, measurement, and relation to critical thinking skill Informal Logic 20 61–84
[10] Scalise K and Gifford B 2006 Computer-based assessment in e-learning: a framework for constructing “intermediate constraint” questions and tasks for technology platforms The Journal of Technology, Learning, and Assessment 4 1-45
[11] Hajhashemi K and Eng W B 2010 A Validation Study of the Persian Version of McKenzie's Multiple Intelligences Inventory to Measure Profiles of Pre-University Students http://www.ssoar.info
[12] Armstrong T 2009 Multiple Intelligences in The Classroom (Alexandria: ASCD)
[13] Mojjarrad H, Hemmati F, Jafari Gohar M and Sadeghi A 2013 Computer-based assessment (CBA) vs. Paper/pencil-based assessment (PPBA): An investigation into the performance and attitude of Iranian EFL learners' reading comprehension International journal of Language Learning and Applied Linguistic World 4 418
[14] Gravetter F J and Forzano L B Research Methods for The Behavioral Sciences 4th ed (Belmont: Wadsworth Cengage Learning)
[15] Mohammadi M and Barzgaran M 2010 Comparability of computer-based and paper-based versions of writing section of PET in Iranian EFL Context *The Journal of Applied Linguistics.* 3 141-162

[16] Britner S L 2008 Motivation in high school science students: a comparison of gender differences in life, physical, and earth science classes. *Journal of Research in Science Teaching* 45 955-970