Rethinking standardized test of science education in Indonesian high school

Rethinking standardized test of science education in Indonesian high school

H Setiawan[1,*], K Garnier[2], and W Isnaeni [3]

[1]Biology Teacher in Global Mandiri Senior High School, Jakarta, Indonesia
[2]Education Faculty, Monash University, Australia.
[3]Biology Department, Mathematics and Natural Science Faculty, Universitas Negeri Semarang, Central Java, Indonesia.

[*]Corresponding author: heru.setiawan@teacher.globalmandiri.sch.id

Abstract. The purpose of this paper is to evaluate the pros and cons of standardized test assessment of Science in High School in Indonesia. A critical review of literature was used in this paper. The literature includes the current policy issues and peer-reviewed journal article from some educational databases. The data are analysed and synthesised qualitatively. Considering the literature portrays the relationship between students and assessment by focusing on how to improve assessment practices, the national educational system of Indonesia should reconsider multiple-choice assessment in standardized-test of Indonesian National examination. The assessment might have some advantages such as affordability of administration to a large number of students, objectivity, and reliability. However, the major problems about the standardized tests are that it less measure deep thinking skills and the test scores less reflect students' abilities or the quality of students' learning. Standardized testing means students would have more rote memorization, but they have less time to synthesize information or apply knowledge which is the challenge in the 21st century. Therefore, it seems not wise when the test is used as the only consideration to make important decisions about students for their graduation from high school.

1. Introduction

Standardized testing of national examination becomes a significant problem causing many high school students to be more stressed than ever in Indonesia. Furthermore, some of them commit suicide due to the pressures of excelling in the examination. Students' future depends on good test score in the exam. This means that the result of the test becomes an important decision not only for their graduation from school but also their future life. The score of the test also becomes an important consideration for enrolment into college. Although these issues happen, the official government believes that multiple-choice standardized testing is the objective and fair method of assessment to measure students' mastery of subject matter due to a large number of students. However, many experts disagree with this because it does not reflect students' learning thoroughly and it seems not suitable for what students' need in the digital era. Based on these endless debates, this paper is going to problematize standardized test in the Indonesian educational system. We will specifically focus on high school context.

This paper would explore the key issues and debates on the implementation of multiple choice assessment standardized test and we will also explain why the government needs to reconsider or rethinking the assessment practice. Then, in the next part, we will discuss why it is important to move

beyond standardized test to the alternative assessment or authentic assessment, which is claimed by many experts as assessment of the 21st century [1, 2]. We will also discuss the potential barriers. By exploring themes emerging from a debate around challenging the practice of assessment, it would have some significance. It has a greater purpose for students, it would reform the traditional assessment to meet the essential skills needed today and, in the future [3]. It would be also beneficial for school and systemic improvement. This is because in some cases, state governments are using the tests to make important decision about students' graduation. Furthermore, they deny students' diplomas based on unsatisfactory results of the final exam.

2. Methods

A critical review of literature is used in this paper. The kinds of literature include the current policy issues and debates related to assessment in education at a national level compared to theoretical papers published in peer-reviewed journal investigating science education assessment and assessment of the 21st century. The policy paper was collected from the official website of the Ministry of Education of Indonesia related to the standard of assessment system in education [4]. The issues investigated are collected from nationally published online newspapers. The Journal article were collected from published articles from some of the educational databases available from Monash University Library including ProQuest education journal, A+ education, ERIC (Educational Research Information Center), Scopus, and ISI Web of Science related to this topic. The key issues, related articles and policy papers are then analysed and synthesised qualitatively.

3. Result and Discussion

3.1 Problematizing the use of multiple choice standardized test.

In this section, we present the key ideas that emerged from the analysis of the literature in the debate. Pros and cons of multiple choice assessment practice in Indonesia appear in the literature. Some researchers critique the implementation of multiple choice assessment due to some reasons. Multiple choice assessment in science education in high school only emphasizes the cognitive domain [5]. In other words, the test score seems not reflect students' abilities and it does not give much information about the quality of student learning and it does not measure students' performance thoroughly. Furthermore, based on investigation, the national examination question of science in Indonesia emphasizes low level of thinking assessing students' ability to describe and analyse [5] and it has little concern to assess their ability to apply, evaluate, synthesize and create [6]. This fact contradicts the standard of PISA where "PISA tests emphasize on students' ability in using scientific knowledge with the higher order cognitive processes of evaluation, critique and synthesis" [7]. Students can also more likely to guess the answer from a list through analysing patterns of alternative answers that seems not to need any deep thinking. In some cases, students use test-wise by looking for the one that best fits without any reason. Therefore, students may memorize facts or procedures well and obtain a good score, but less in the application. In this case, students are only accepting the fact and has less chance to challenge texts or to explore new ideas. Also, the standardized test creates inequality between students with different socio-economic status. Students who achieve good scores generally come from a wealthy family background. It does not mean that wealthier students are smarter than poorer students, but they have an advantage because the wealthier family can afford the best tutors to pass the examination. It also seems that through this fact, school kills creativity where schools are too much focuses on test preparation and testing, then they teach the students how to take the test rather than teach the students to become problem solvers.

3.2 Move beyond Standardized test to Authentic Assessment

Some experts argue that multiple choice question assessment contradicts the principle of authentic assessment [8], written in the 2013 curriculum of Indonesia. Authentic assessment is the measurement of intellectual accomplishments that are worthwhile, significant, and meaningful, which refers to the

context of the real-world situation to solve complex problems [9]. Some experts argue that it becomes assessment in the 21st Century [1, 2]. Evidence supports the idea that authentic assessment is beneficial for students. For example, students benefit from Project-based learning approach because it facilitates students' inquiry and works collaboratively to research and create projects that reflect their knowledge [10]. Furthermore, most of standardized test have now incorporated or sometimes converted entirely to performance assessment [11]. Google are launching product every day such as google classroom. This is not the result of testing, but project-based assignment that is supported by creativity that promotes thinking and learning.

Although multiple choice assessment matter, some researchers supported the use of the assessment because it has some strength compared to other types of assessment. First, Standardized tests with multiple choice are used to assess large groups of students because it can cover large amounts of material very efficiently and it is affordable for testing a large number of students. Scoring is easy, reliable and require less time and is an effective way to measure student knowledge on a large scale [12]. It is made equal to for students everywhere in objective and fair assessment to identify the achievement gaps between students. Such reasons are understandable, especially senior high school in Indonesia that yield high enrolment in Indonesia every year, so that we can compare large number of students. Furthermore, science education in high school has many subunits. Yet, the counter of multiple choice argues that multiple choice is assessment of learning, not assessment for learning. Teachers should see that process is important or even more important concern than result by implementing assessment for learning, not assessment of learning. When students think that result is the most important they might do anything to get the best score, such as cheating, because of their score-oriented mindset. They should have a paradigm that success is a journey as an ongoing process, not a destination to pass the exam that can be reflected in the practice of assessment. Some supporter of multiple choice (MC) question also argue that it considered an objective form of assessment with high reliability and affordability [13]. However, others criticized that with their objectivity, it does not provide students with the opportunity to explain their answers, thus potentially limiting the depth and scope of information gathered from students [14]. In science assessment, MC items tend to focus on facts rather than certain aspects of inquiry science such as complex arguments or coherent understanding [15, 16].

Considering the literature focusing on how to improve assessment practices, some literatures recommended that the assessment practice in high school should also measure 21st-century skills. The skills and knowledge that students learn in school and what they need in the 21st-century community should not be much different [17]. New assessment approach using 21st-century assessment might be the proper solution to close the gap between them. To address this issue, The Assessment and Teaching of Twenty-First Century Skills Project (ATC21S) was launched in 2009 [18]. The practice of multiple choice question is not suitable anymore with the challenges that students face in the real-world situation. For example, when they work as a doctor in a hospital, how can they implement multiple choice test in their workplace? What they need are abilities to collaborate and communicate with others to reach the goals of their organization. Also, teachers are responsible for not only encouraging their students to master all of the subject matter, but also giving them meaningful experiences to prepare students for their future careers and being ready to live in the society. In other words, assessment should be able to capture the more diverse skills which have any value in the workplace.

Assessment should measure what is the most important for students for long life learning where they not only memorize concept, but they can apply the knowledge. Teachers should implement any assessment practice that will benefit their students. It needs reform the current practice of assessment in science in Indonesia from low level of thinking to the standard of assessment in the 21st century learning [19]. The framework of assessment includes (1) learning and innovation skills, (2) information, media, and technology skills, (3) life and career skills. The learning and innovation skills include creativity, innovation, critical thinking, problem-solving, collaboration, and communication. Then, the information, media, and technology skills include information literacy, media literacy, and ICT literacy. These skills are very fundamental to be integrated [20]. However, the implementation of authentic assessment in Indonesia might face some obstacles. This forms the basis of our next discussion.

3.3 Some potential barriers or obstacles to the implementation of authentic assessment

According to the fact drawn from literature, there are two major obstacles for implementing authentic assessment into standardized test in Indonesia including internal and external barriers. The first, maintaining the practice of standardized test is also influenced by a political factor. To change the current practice is not easy because the proposal must be approved by the House of People's Representatives and president. As an essential part of the democratic process, any change to the law must also be debated in Parliament. Furthermore, many politicians would object to or resist to changing the system. Ministry of Education Indonesia argued that standardized test using multiple choice is the most proper instrument to test cognitive aspect as an indicator of the quality of education in Indonesia [21].

Also, the institutional policies such as time constraint, the number of teachers and students might make authentic assessment difficult to be implemented. Teachers often do not have much time to do the assessment in the classroom by observing students' activity one by one in detail. This is supported by some researchers that the implementation of authentic assessment is time-consuming [9]. Teachers that has little time will face big obstacles to implement this type of assessment. A quote from teachers said that "the time is limited…there's a lot of things to cover and you are pressured at all times" [20]. It seems difficult to assess the whole competence of all students thoroughly. Meanwhile, many decisions should be taken as soon as possible by teachers in the classroom. Furthermore, the number of students in each class and the number of teachers is not proportional to school hour in Indonesia. The allocated time for science in 2013 curriculum in senior high school is only 5 hours within a week with most of their time for class theory [4]. The number of students is approximately 30 students and only 2 teachers each school on average. This is clearly seen in Government Policy No. 23 of 2013 released by the Ministry of Education Indonesia which states "the number of learners in the class for secondary school is no exceed than 32 ... "(p.55).

Second, there are three main aspects of external obstacles including the key stakeholder that implement the curriculum cultural influences, and infrastructure. The key stakeholder of assessment is teachers. Teachers have a crucial role such preparing the learning program, implementing the learning program, carry out the assessment. However, some research found that Indonesian teachers' qualification and professionalism is still low. Most sciences teachers in Indonesia do not have relevant study discipline [5]. They are not studying science in their bachelor or master's degree program. Furthermore, some of them come from Industrial Engineering and building Engineering which may lack professionalism in teaching science. Teachers professionalism gives a significant impact to implement authentic assessments [20]. It would create poor understanding of authentic assessment in science and poor of creativity. Teachers with low creativity tend to be unable to overcome some problem in authentic assessment, analysing learning outcomes, and conducting follow-up programs. The Indonesian teachers lack understanding in assessment system are developing the instrument, designing the assessment rubric, and gathering data from multiple measurement techniques [21].

Some evidences drawn from literature also depict that teachers' difficulties in implementing authentic assessment are also influenced by cultural factors such as characteristics of the students. Bad characteristics of some Indonesian students found in the class are less responsible, less independent in doing the task, and have low motivation [5] which is influenced by family background, economic factor, ethnicity, gender, and race [19]. For example, students are difficult to be asked to express opinions or answer questions verbally [19]. Limited facilities and infrastructure in the school are also the main obstacle to implement authentic assessment. Complete and convenience infrastructure will support teachers and students to do the tasks [19]. For example, when there is no tools and material in the Secondary school in Indonesia, then students cannot conduct experiment in the laboratory. Teachers cannot assess students' ability in operating laboratory tools and conducting experiment properly. These issues are the challenges need to be considered not only by Indonesian government but also by educators, students, parents, state, and country because they use much of their time, effort, and money to succeed in the examination from school.

4. Conclusion

We argue that the national education system should reconsider multiple-choice assessment in standardized-test of Indonesian National examination. The assessment might have some advantages such as affordable for large number of students, objective, reliable, and cheaper than other forms of tests. However, one of the biggest problem about the standardized tests is that it might not measure deep thinking skills. It also reflects assessment of learning not assessment for learning. In other words, the test scores cannot reflect students' abilities and it not gives much information about the quality of student learning. Multiple-choice items are poor tools for measuring the ability to apply knowledge to complex problems which the challenge of 21st-century. In this era, it is more important to teach the students that stimulate their creativity and how to solve the real-world problem rather than teaching students how to face the test. Hence, multiple choice should only be used as the part of assessment such as to check on factual and procedural knowledge. It should not substantially control the curriculum or instruction to make important decisions about students.

As the alternative of standardized test, authentic assessment can reflect the actual performance of students. It should also assess 21st-century skills such as high order thinking skills, collaboration, communication, problem-solving skills to prepare students in the real world. Authentic assessment is implemented in Indonesian high school especially in the 2013 curriculum. However, the implementation faced internal and external barriers. The former, maintaining the practice of standardized test is influenced by a political factor. It includes some institutional policy in educational system of Indonesia itself such as time constraint, the ratio between teachers and students in Indonesia makes authentic assessment challenging to be implemented. The external obstacles include the expertise of key stakeholder, cultural factor of students in Indonesia, and infrastructure available in the school.

This paper has an implication for research. More investigation needs to be conducted to assess the use of authentic assessment in Indonesia to provides more empirical evidence of the benefits for students. Future researchers can also develop an approach, model, module, guidance, or instrument to implement the authentic assessment based on high order thinking skills to help secondary science teachers in the classroom.

Acknowledgement

This paper is funded by *Lembaga Pengelola Dana Pendidikan*/LPDP (Indonesia Endowment Fund for Education), Ministry of Finance of Republic of Indonesia as the parts of Scholarship Contract No. PRJ-6893/LPDP.3/2016 for International Conference.

References

[1] Greenstein L M 2012 *Assessing 21st century skills: A guide to evaluating mastery and authentic learning* (Melbourne: Corwin Press)
[2] Voogt J, Erstad O, Dede C & Mishra P 2013 *J. comput. assist. learn.* **29(5)** 403
[3] Timmis S, Broadfoot P, Sutherland R & Oldfield A 2016 *Br. Educ. Res. J.* **42(3)** 454
[4] The Ministry of Education and Culture Indonesia 2016 *The Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 23 of 2016 Regarding Educational Assessment Standards*. Retrieved from: http://bsnp-indonesia.org/wp-content/uploads/2009/09/Permendikbud_Tahun2016_Nomor023.pdf
[5] Pantiwati Y,Wahyuni S and Permana F H 2017 *J. Educ. Pract.* **8(8)** 117
[6] Ardan A S, Ardi M, Hala Y, Supu A and Dirawan G D 2015 *Int. Educ. Stud.* **8(4)** 52
[7] Osborne J 2014 *Sch. Sci. Rev.* **95(352)** 53
[8] Abosalem Y 2016 *J. Second. Educ.* **4(1)** 1
[9] Wiewiora A & Kowalkiewicz A 2019 *Assess. Eval. High. Educ.* **44(3)** 415
[10] Chen C H & Yang Y C 2019 *Educ. Res. Rev.* **26** 71
[11] Butler A C 2018 *J. Appl. Res. Mem. Cogn.* **7(3)** 323

[12] Shin J, Guo Q & Gierl MJ 2019 *Front. Psychol.* **10** 825.
[13] Broadbent J, Panadero E & Boud D 2018 *Assess. Eval. High. Educ*. **43(2)** 307
[14] Wind S A & Peterson M E 2018 *Lang. Test.* **35(2)** 161
[15] Liu O L, Lee H and Linn M C 2011 *Educ. Assess.* **16(3)** 164
[16] Wooten M M, Cool A M, Prather E E and Tanner K D *2014 Phys. Rev. Special Top.-Phys. Educ. Res.* **10(2)** 020103
[17] Black P 2015 *Assess. Educ.: Princ. Policy Pract.* **22(1)** 161
[18] Griffin P and Care 2014 *Assessment and teaching of 21st century skills: Methods and approach* (Netherland: Springer)
[19] Retnawati H, Hadi S and Nugraha A C 2016 *Int. J. Instr*. **9(1)** 33
[20] Bramwell-Lalor S & Rainford M 2015 *Eur. J. Sci. Math. Educ.* **4(3)** 380
[21] Fauzi A, Bundu, P and Tahmir 2016 *Int. J. Environ. Sci. Educ*. **11(10)** 3231