Development of Computer-Based Tests on Essay-type Questions: Application in Curriculum and Learning of Mathematics Courses

Syukrul Hamdi1, Badrun Kartowagiran2, Suyanta3, Amat Jaedun4, Siswantoyo5

1 Lecturer, Department of Mathematics Education, Yogyakarta State University

2,4 Lecturer, Undergraduate Program, Research and Evaluation of Education study Program, Yogyakarta State University

3Lecturer, Undergraduate Program, Chemistry Education Study Program, Yogyakarta State University

5 Lecturer, Undergraduate Program, Sports Science Study Program, Yogyakarta State University

syukrulhamdi@uny.ac.id

**Abstract.** This study aims to develop Computer-Based Tests (CBT) kits on essay type questions to prepare professional teachers in the era of the Industrial Revolution 4.0. This research is a type of development research which is carried out in 6 stages, namely (1) development of an essay type question instrument in the Curriculum and Learning of Mathematics course; (2) Focus Group Discussion; (3) development of computer-based test kits; (4) expert assessment on the developed computer-based test kit; (5) pilot study on computer-based test kits; (6) final revision of computer-based test kits. The focus group discussion was conducted involving 9 experts, then validation on the computer-based test kit was carried out by involving 4 media experts. The pilot study of the test kit was conducted on 65 students of mathematics pre-service teacher. Expert validation values are obtained through the Aiken formula; the level of readability and effectiveness of the test kits are described based on expert judgment. The results showed that the test kit (computer-based) had met the readability aspect and met the good standard with an average expert judgment of 0.78. Based on the results, it can be concluded that the computer-based test kits developed in this study can be used by lecturers or teachers as a means of online learning assessment, especially during the COVID-19 pandemic, which forces teachers and students to do online learning

1. Introduction

In the education 4.0 era, technological learning media designs and developments become the important factors to carry out the teaching and learning process. The use of learning media has even switched to not only as a supplementary component but even as a major component, especially during the pandemic, which requires all elements of teaching and pedagogy practice to carry out learning online and remotely. With the availability of useful media, the quality of the learning activities will be more optimal. One of the existing problems found in today’s education field is the availability of media that is suitable for the form or type of assessment being taught to students. Some of the learning media developed previously were still limited to the media used to help students understand easily in learning such as visual or audio-visual props. However, it is rare to find media that can be used by lecturers, especially to accommodate several forms of assessment commonly used. The assessment media which is currently easy to find by teachers are still frequently in the form of media such as Computer-Based Test on Multiple-choice Questions which is generally used during national exams.

Digital-based assessment for prospective teacher-students is massively encouraged by developing a computer-aided assessment model to support in-class time learning in the era of Education 4.0. Codreanu, et al [1] revealed that the skill of using computer-aided assessment by both pre-service teachers is a part of professional requisite and a fundamental requirement for better teaching and learning activities shortly. Therefore, this study attempts to offer pre-service teachers, particularly students in the math education department, a digital competence in assessment by developing a computer-based test on an essay-type question assessment model. It is a kind of digital competence support for the math education students to enhance their professional career prior to being in-service teachers in the practice-based training in the field [2][3].

Teachers’ digital literacy or teachers’ ICT skill is competence and understanding of information and communication technology had by teachers to support and enhance their professionalism pedagogy such as designing a lesson plan, preparing learning media integrated with technology, and attempt to use digital-based assessment or computer-aided. Nonetheless, Wang & Liu [4] emphasized that teachers’ educational technology competence referred to the teachers’ capability in using technology in daily use such as sending/receiving e-mail, online communication, social media integration to deliver material, and internet shopping. Besides, Claro et al [5] summarized several relevant kinds of research about pre-service teachers’ digital literacy concepts which included 1) operating ICT, which is teachers’ necessary skill to use, operate, handle, and solve technology related to educational issues such as spreadsheets, word processors, and navigation equipment, 2) working with information, which refers to cognitive skills in organizing, selecting, reviewing, and evaluating the information and data provided on the internet to be a piece of true information and also a capacity to communicate big data and information, 3) digital technology understanding, which refers to teachers’ ability and role to think critically, to adapt with the fast-changing development, and to adjust with technology demand in the real-word problems encountered.

In the case of the Covid-19 crisis, teachers or pre-service teacher students should have to utilize learning media with technology integration. Indeed, they should know to operate digital devices to deliver learning materials and perform the computer-aided test or digital-based test. Operating technology devices include an individual capacity to benefit technology in learning, teaching, and assessment [6]. However, teachers seemingly are inability to integrate their skills and knowledge into their teaching practice, particularly with assessment, in the class. It is commonly found that teachers are still dominant to apply conventional assessment model in the classroom [7][8]. Hence, CBT on essay-type question assessment model is developed to fulfill teachers’ needs and pre-service teacher students in implementing online and remote learning. In addition, the increasing number of research conducted is more concerned with how teachers integrated technology in the learning and teaching process than assessment integrated technology [9][10][11].

An essay-type question test is a kind of assessment that requires students to answer questions in the form of analyzing, describing, explaining, discussing, comparing, and giving reasons [12]. The essay-type question test requires fairly high analytical skills for students since the questions are open-ended which obligated students to construct their answers are or responses [13] compared to Multiple Choice-Questions (MCQs). MCQs have several weaknesses in assessing students. MCQs are predisposed to form students' learning habits that are not deep and do not develop skills in writing [14][15][16]. Besides, an assessment with MCQs tends to let students answer pseudo-guessing and do not answer questions based on their knowledge and reasoning [17]. Therefore, it makes essay-type questions important to continue to be developed because not all assessments can be made through MCQs which lead to assessment bias [18]. The ETQ (Essay-Type Question) is to be developed as an open-ended assessment tool for assessing learning at the higher education level [19]

Besides the type of assessment measurement tool, selecting the appropriate media is also needed to boost the learning assessment system. The media needed by teachers needs to be initiated to be more complex and accommodate wider ranges of assessment models according to the demands of learning in the era of IR 4.0. Creativity is one of the keys needed in developing media in learning. Quoted from the World Economic Forum, by 2020mrobots will replace a lot of human positions at work so that many types of jobs will be lost and changed along with rapid developments due to the 4.0 industrial revolution. However, creativity cannot be owned by robots [20] and only humans who will have it. Creativity skills will be formed through a literacy culture. As a result of research by Appleton, Montero, & Jones [21] who find that a literacy approach needs to be developed through learning by using digital literacy and technology to increase creativity.

Developing a medium of Computer-Based Test (CBT) is one of the developments of complex learning assessment media which fit with the demand of the digital era. However, using CBT also has some technical challenges such as the compatibility of software and the user's mastery of technology. Claro et al [22] define several areas in the use of technology competence, comprising of technological application, pedagogical practice, academic administration, and evaluation as well as teacher development. CBT essay-type questions include a technology-enhanced test that is applied in Indonesia. The transition from PBT essay-type question to CBT essay-type question will have a major impact on test performance, especially time and accuracy response [23][24][25] in technology-assisted assessment learning. This is what is also needed in mathematics assessments which use a lot of essay-type questions including curriculum and mathematics learning courses in the Mathematics Department.

1. Method

The research method employed was development research with 6 phases, namely the developing curriculum and mathematics learning instruments, Focus Group Discussion (FGD), CBT development, CBT assessment by experts, CBT trial, and final revision of CBT. The FGD involved 9 experts, and the CBT validation involved 4 media experts. The trial involved 65 Mathematics pre-service teacher students who took the curriculum and mathematics learning courses. Expert judgment validation was analyzed using the Aiken formula, readability, and application effectiveness were described based on expert judgment.

The instrument for IT experts was used to obtain product quality data in terms of conformity with web-based software quality standards. The instrument used was developed from Pressman's theory [26] associating with web application testing which consisted of five aspects, namely Content, Interface, Navigation, Configuration, and Security. Data analysis from the FGD results was carried out by using quantitative and qualitative descriptive analysis. Furthermore, the result of the application trial for each class is described separately using a quantitative descriptive. Student learning outcomes of pre-service teachers were compared using the independent sample t-test. It was used to determine the effectiveness and functionality of applications developed in different classes.

1. Result

**The Instrument of Essay-Type Questions of Curriculum and Mathematics Learning Course**

The developed essay-type question instrument consists of 5 essay questions with different points. The five questions which had been arranged would be inserted into the developed application system. Before entering into the system, first prepare the rubric answers and the point of each step of the answers that will be entered into the system. The assessment rubric is required and inputted in the application providing that the system provides scoring and automatic assessment according to the rubric made by the lecturers. The essay-type test instrument blueprints of curriculum and mathematics learning course are presented in Table 1

**Table 1**. The essay-type test instrument blueprints of curriculum sand mathematics learning course

| **Competence** | **Indicator** | **Item number** |
| --- | --- | --- |
| Analyzing various basic concepts of education, and pedagogy-didactic mathematics | Describing curriculum components (passing competency standards, content standards, process standards, and assessment standards) | 1 and 2 |
| Describing how the 2013 Curriculum is | 3 |
| Describing the characteristics of learning mathematics | 4 |
| Developing sets of mathematics learning | 5 |

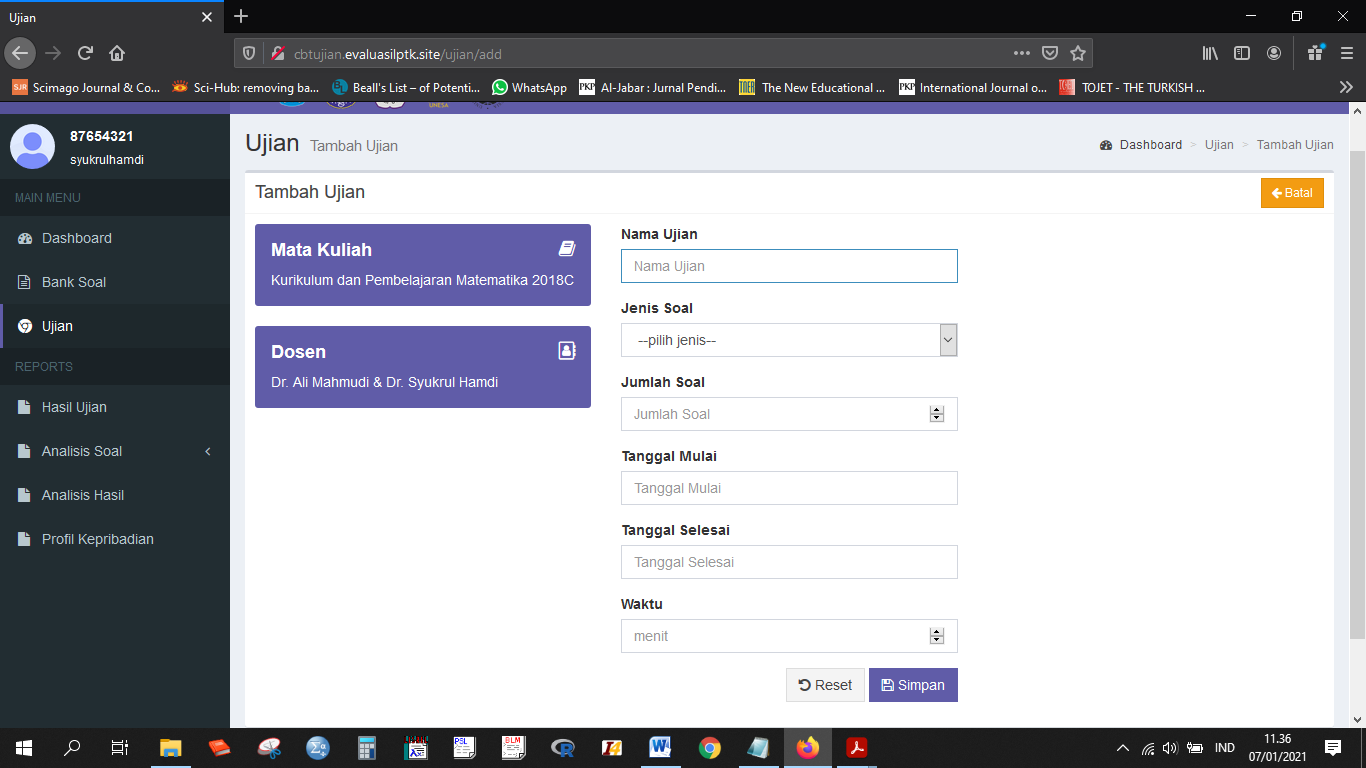
Based on the result of expert judgment validation and the result of analysis by using the Aiken formula, the instrument of essay-type question test of curriculum sand mathematics learning course was categorized as valid with each Aiken index greater than 0.7.

**The Result of Computer-Based Test on Essay-Type Question**

The developed application comprises of applications for lecturers, admins, and prospective teacher students. The introductory display of the developed media is presented in Figure 1. The display of the dashboard appeared after login by inputting your email and password. The dashboard display contains account information from the user. From this page, users can create a question item bank, create exams, and view test results from respondents who have taken the exam. The display in Figure 2

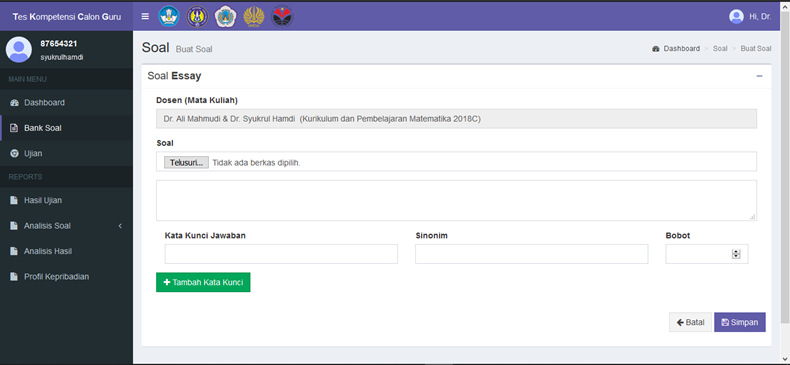
|  |  |
| --- | --- |
| **Figure 1**. CBT Introductory Display | **Figure 2**. The display after entering or login into the Application |

If the users want to do an examination, a sub-menu has been provided to create an exam where this sub-menu consists of the name of the exam, type of question, number of questions, start date, end date, and time. The test menu is presented in Figure 3.



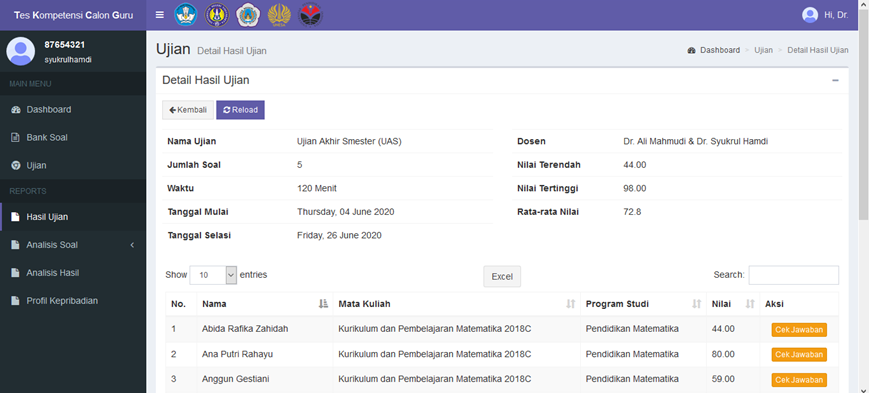
**Figure 3**. The Display in Making an Exam

Screen display for inputting questions and answer keys, synonyms, and points. Columns for inputting questions, answer keys, synonyms, and points have been provided and separated respectively in this application. The display of inputting questions is presented in Figure 4



**Figure 4**. The Display to Input Questions

If the exam has been completed, the respondent's score or score is recorded and stored in the system so that it can be accessed by the lecturer to detect the result of the respondent's score. The answers of each respondent are recorded and stored separately with other respondents to make it easier for lecturers to document them. An example of the test results is presented in Figure 5.



**Figure 5**. The Display of CBT Implementation Result

The result of the CBT development was assessed by experts. The assessment by media experts employed an assessment instrument consisting of content, interface, navigation, configuration, and security assessment components. The assessment was conducted by 4 experts. The following are the converted data testing results by experts with the assessment categories shown in Table 2..

**Table 2**. The Assessment Result of Media Experts

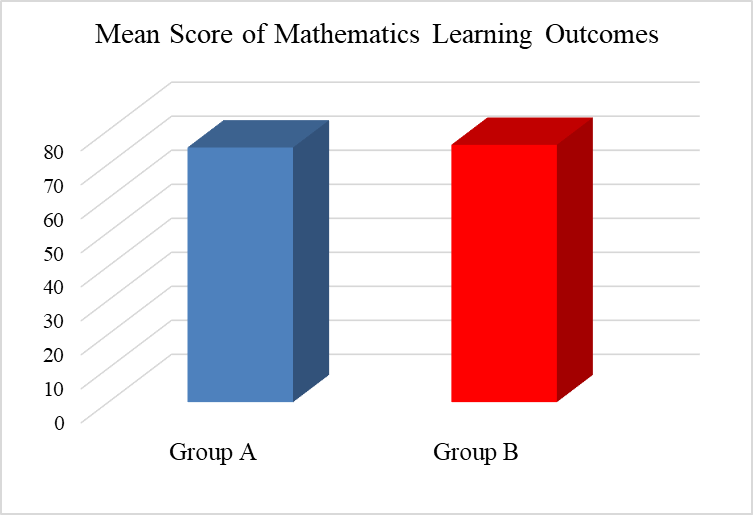
|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Aspect** | **Assessment Result** | **Category** |
| 1 | Content | 0.77 | Good |
| 2 | Interface | 0.77 | Good |
| 3 | Navigation | 0.75 | Good |
| 4 | Configuration | 0.75 | Good |
| 5 | Security | 0.85 | Very Good |
| Average Score | | 0.78 | Good |

The application which is in the good category based on experts’ judgment is employed for trial examination or implementation of the application which has been developed.

**The result of Computer-Based Test Implementation**

Student learning outcomes using the Computer-Based Tests on Essay-type Questions application in group A with a total sample of 35 students obtained the highest score of 93 and the lowest score of 37. The standard deviation gained in group A's mathematics learning results was 13.372 and the score variance of 178.81. The learning results of group B with 31 students obtained the highest score of 98 and the lowest score of 57. The standard deviation was 11.32 and the score variance was 128.1

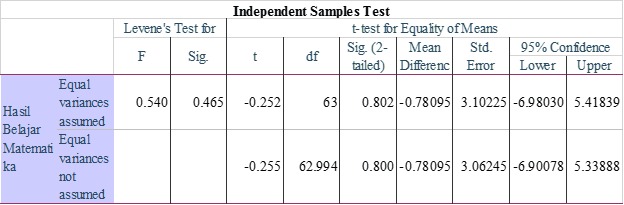
Group A got a mean value of 74.86 and group B with a mean value of 75.67. The mean value in group A was smaller than the mean value in group B (74.86 <75.67). The average score of group A and group B are presented in Figure 6



**Figure 6**. Mathematics Learning Outcomes

To find out whether there is a significant difference in scores in the student groups through the use of the Computer-Based Tests on Essay-type Questions application, the Difference Test is carried out using the Independent Sample t-Test.

**Table 3**. The Different Test of Student Mathematics Learning Results (Independent Sample t-Test)



As shown in table 3, the output of the Independent Sample t-Test yield that the Sig. value on Levene's Test for Equality of Variances is 0.465> 0.05, so it can be said that the data variance between the mathematics learning results of Group A and Group B is homogeneous. Meanwhile, Sig. (2 tailed) on the t-test for Equality of Means of 0.802> 0.05, so it can be concluded that there is no difference in student mathematics learning outcomes in both group A and group B using the Computer-Based Tests fn Essay-type Questions application on Curriculum and Mathematics Learning courses. It shows that CBT is running well and there is no difference in results in trials in two different classes.

1. Discussion

Based on the research result, the CBT-essay type question is effective in the curriculum and mathematics learning courses in the context of mathematics education department students based on expert judgment and test results. Students are prepared to become competent graduates in the aspects of knowledge, understanding, and skills in using technology. This is in line with the research of Appleton, Montero, & Jones [27] who found that a literacy approach needs to be developed through learning employing digital literacy and technology to increase student creativity.

The use of CBT has an important role for lecturers, as part of the application which is friendly-use for lecturers to correct exam results automatically. In addition, online learning currently requires lecturers to have creativity in conducting assessments so that CBT is one of the solutions in solving lecturer problems in the assessment. Furthermore, the implementation of CBT- essay-test questions is demanded in learning to encourage the students' cognitive flexibility abilities related to technology. In other words, it is absolutely an effort from universities and educational institutions in preparing graduates who have digital literacy and technological skills to support professional careers in the near future. Codreanu, et al [28] argue that the skill of using computer-assisted assessment by pre-service teachers is one of the professional requisites and a fundamental requirement for better teaching and learning process in the classroom.

CBT, a test or an assessment administered by a computer has become a sophisticated popular assessment model adapted by the universities and other educational institutions in the last few decades along with the rapid development of technology and information [29][30]. The use of CBT is considered to be able to improve the quality of assessment in education because it is integrated into a system [31][32]. The use of CBT when compared to the Paper-Based Test (PBT), is less time-consuming, easier, and quicker to be administered and scored [33]. Besides, CBT has advantages such as real-time scoring and immediate feedback. The efficiency of the assessment time can be used by teachers to improve their teaching time [34][35][36]. CBT can also facilitate teacher instruction to be more strategically directed to enhance individual student goals [37].

1. Conclusion

Based on the results of the two-class similarity test using the Computer-Based Tests on Essay-type Questions application in the Curriculum and Mathematics Learning course, it can be concluded that there is no difference in student mathematics learning results in both group A and group B. It shows that CBT is running well and there is no difference in result trials in two different classes. The results also show that the application has met the aspects of readability and good scores by experts with an average rating of 0.78. To sum up, the application can be used by lecturers or teachers as part of the solution for online learning assessments, especially during the COVID-19 pandemic which requires online and remote learning.

1. References
2. Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education*, 95, 103146.
3. Instefjord, E. J., & Munthe, E. (2017). Educating digitally competent teachers: A study of integration of professional digital competence in teacher education. *Teaching and Teacher Education*, 67, 37-45. <https://doi.org/10.1016/j.tate.2017.05.016>
4. Reisoğlu İ. & Çebi Ayç., How can the digital competences of pre-service teachers Be developed? Examining a case study through the lens of DigComp and DigCompEdu, *Computers & Education* (2020), doi: <https://doi.org/10.1016/j.compedu.2020.103940>
5. Wang, Y., & Lu, H. (2021). Validating items of different modalities to assess the educational technology competency of pre-service teachers. *Computers & Education*, 162, 104081. doi:10.1016/j.compedu.2020.104081
6. Claro, M., Salinas, Á., Cabello-Hutt, T., San Martín, E., Preiss, D. D., Valenzuela, S., & Jara, I. (2018). Teaching in a Digital Environment (TIDE): Defining and measuring teachers' capacity to develop students' digital information and communication skills. *Computers & Education*, 121, 162-174.
7. Lucas M., Bem-Haja P., Siddiq F., Moreira A. & Redecker C., The relation between in-service teachers’ digital competence and personal and contextual factors: What matters most?, *Computers & Education*, <https://doi.org/10.1016/j.compedu.2020.104052>
8. Wang, Y., & Lu, H. (2021). Validating items of different modalities to assess the educational technology competency of pre-service teachers. *Computers & Education*, 162, 104081. doi:10.1016/j.compedu.2020.104081
9. Mouza, C., Karchmer-Klein, R., Nandakumar, R., Ozden, S. Y., & Hu, L. (2014). Investigating the impact of an integrated approach to the development of preservice teachers’ technological pedagogical content knowledge (TPACK). *Computers & Education*, 71, 206–221. <https://doi.org/10.1016/j.compedu.2013.09.020>
10. Instefjord, E., & Munthe, E. (2015). Preparing pre-service teachers to integrate technology: An analysis of the emphasis on digital competence in teacher education curricula. European Journal of Teacher Education, 39(1), 77–93.https://doi.org/10.1080/02619768.2015.1100602 *International Society for Technology in Education* (ISTE).
11. Foulger, T., Graziano, K., Schmidt-Crawford, D., & Slykhuis, D. (2017). Teacher educator technology competencies. Journal of Technology and Teacher Education, 25(4), 413–448
12. Wang, Y., & Lu, H. (2021). Validating items of different modalities to assess the educational technology competency of pre-service teachers. *Computers & Education*, 162, 104081. doi:10.1016/j.compedu.2020.104081
13. Supardi. (2015). Penilaian Autentik Pembelajaran Afektif Kognitif dan Psikomotorik. Jakarta: Raja Grafindo Persada
14. Kaipa, R. M. (2020). Multiple choice questions and essay questions in curriculum. *Journal of Applied Research in Higher Education*
15. Struyven, K., Dochy, F. and Janssens, S. (2005), “Students’ perceptions about evaluation and assessment in higher education: a review”, *Assessment and Evaluation in Higher Education*, Vol. 30 No. 4, pp. 325-341, DOI:10.1080/02602930500099102.
16. Parmenter, D.A. (2009), “Essay versus multiple-choice: student preferences and the underlying rationale with implications for test construction”, *Academy of Educational Leadership Journal*, Vol. 13 No. 2, pp. 57-71.
17. Mullen, K. and Schultz, M. (2012), “Short answer versus multiple choice examination questions for first year chemistry”, *International Journal of Innovative Science and Modern Engineering*, Vol. 20 No. 3, pp. 1-18.
18. Woodford, K. and Bancroft, P. (2004), “Using multiple choice questions effectively in information technology education”, *ASCILITE*, Vol. 4, pp. 948-955
19. Hassmen, P., Hunt, D.P. and Dybeck, C. (2002), “Effects of self-assessment on retention in rule-based learning”, *Perceptual and motor skills*, Vol. 94 No. 1, pp. 296-306, doi: 10.2466/pms.2002.94.1.296
20. Kaipa, R. M. (2020). Multiple choice questions and essay questions in curriculum. *Journal of Applied Research in Higher Education*
21. World Economic Forum. (2016). The Global Competitiveness Report 2016–2017. Diakses pada tanggal 8 Januari 2021 dari [http://www3.weforum.org/docs/GCR2016-2017/05 FullReport/TheGlobalCompetitivenessReport2016-2017\_FINAL.pdf](http://www3.weforum.org/docs/GCR2016-2017/05%20FullReport/TheGlobalCompetitivenessReport2016-2017_FINAL.pdf).
22. Appleton, L., Montero, G. G., & Jones, A. (2017). Creative approaches to information literacy for creative arts students. *Communications in information literacy*, Vol. 11, No. 1, 2017, 147–167. <https://doi.org/10.15760/comminfolit.2017.11.1.39>.
23. Claro, M., Salinas, Á., Cabello-Hutt, T., San Martín, E., Preiss, D. D., Valenzuela, S., & Jara, I. (2018). Teaching in a Digital Environment (TIDE): Defining and measuring teachers' capacity to develop students' digital information and communication skills. *Computers & Education*, 121, 162-174.
24. Barnard, J. J. (2018). From simulation to implementation: Two CAT case studies. Practical Assessment, *Research and Evaluation*, 23(14), 1–8. <https://pareonline.net/getvn.asp?v=23&n=14>.
25. Bertling, M., & Weeks, J. P. (2018). Using response time data to reduce testing time in cognitive tests. *Psychological Assessment*, 30(3), 328–338. <https://doi.org/10.1037/pas0000466>
26. Ponce, H. R., Mayer, R. E., Sitthiworachart, J., & López, M. J. (2020). Effects on response time and accuracy of technology-enhanced cloze tests: an eye-tracking study. *Educational Technology Research and Developmen*t, 1-21. DOI: <https://doi.org/10.1007/s11423-020-09740-1>
27. Pressman, R. S. (2005). Software engineering: a practitioner’s approach. Palgrave MacMillan
28. Appleton, L., Montero, G. G., & Jones, A. (2017). Creative approaches to information literacy for creative arts students. *Communications in information literacy*, Vol. 11, No. 1, 2017, 147–167. <https://doi.org/10.15760/comminfolit.2017.11.1.39>
29. Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education*, 95, 103146.
30. Villani, M. The Shift from Paper-Based Test to Computer-Based Test in Italian National Assessment: the INVALSI Case. VOLUME II, 271
31. Yao, D. (2020). A Comparative Study of Test Takers' Performance on Computer-Based Test and Paper-Based Test across Different CEFR Levels. *English Language Teaching*, 13(1), 124-133
32. Jerrim, J. (2016). PISA 2012: how do results for the paper and computer tests compare? *Assessment in Education: Principles, Policy & Practice*, 23(4), pp. 495-518
33. Komatsu, H., Rappleye, J. (2017). Did the shift to computer-based testing in PISA 2015 affect reading scores? A View from East Asia, *Compare: A Journal of Comparative and International Education*, 47(49), pp. 616-23.
34. Yao, D. (2020). A Comparative Study of Test Takers' Performance on Computer-Based Test and Paper-Based Test across Different CEFR Levels. *English Language Teaching*, 13(1), 124-133
35. Eid, G. K. (2005). An investigation into the effects and factors influencing computer-based online math problem solving in primary schools. *Journal of Educational Technology Systems*, 33(3), 223-240
36. Jeong, H. (2012). A comparative study of scores on computer-based tests and paper-based tests. Behaviour & Information Technology, 33(4), 410-422. doi:10.1080/0144929x.2012.710647.
37. McClelland, T., & Cuevas, J. (2020). A comparison of computer based testing and paper and pencil testing in mathematics assessment. *The Online Journal of New Horizons in Education*-April, 10(2)
38. Johnson, M., & Green, S. (2006). On-line mathematics assessment: The impact of mode on performance and question answering strategies. *Journal of Technology, Learning, And Assessment*, 4(5), 1-33.

**Acknowledgments**

The researchers would like to express their gratitude towards the Directorate of Research and Community Service, Ministry of Research and Technology of Republic of Indonesia, for the financial support through the Research Grant Program in the Research and Community Service Office of Universitas Negeri Yogyakarta, 2020.