Development of Adaptive Class Assessment Based on Learning Management System for Electrical Engineering Expertise Program in Vocational School

S Hadi\textsuperscript{1} and Haryanto\textsuperscript{2}
\textsuperscript{1,2}Faculty of Engineering, Universitas negeri Yogyakarta, Indonesia

Email: \textsuperscript{1}samsul_hd@uny.ac.id; \textsuperscript{2}haryanto@uny.ac.id

Abstract. This study aims to develop a classroom assessment tool that can be used in LMS-based adaptive assessment and test the appropriateness of the classroom assessment tool. This research is development research. The development research model used is the scale construction process with the following steps: articulate construct and context, choose response format and assemble initial item pool, collect data from respondents, examine psychometric properties and quality so that a final scale is obtained. Respondents of this study were experts who assessed the validation and students of the Vocational High School of Electrical Power Installation Engineering Expertise. The data analysis technique used is instrument analysis using classical test theory and descriptive statistics. The results of the study are: the construct of selected class assessment tools is easy to implement in a computer program. Multiple choice questions with 5 alternative answers. The attitude questionnaire is made in the form of a case assessment with 4 alternative answers, while the observation guidelines are made in the form of a short entry with numbers between 1 and 4. The questions developed have 60 items. Items with less validity are 0, moderate there are 55, and high is 5 with Alpha reliability of 0.934. The attitude questionnaire developed consists of 8 items. The items with less validity are 0, 6, and 2 with high Alpha reliability of 0.731. The observation guide developed contains 16 points of observation. Items with less validity are 0, the medium is 15, and high is 1.

Keywords: adaptive; learning management system; classroom assessment.

1. Background
The Industrial Era 4.0 has penetrated all areas of life, including education. Industry 4.0 relies on five central technologies to support its performance. The five central technologies are computing hardware, production hardware, software, interfaces, and connectivity (Berger, 2016). Vocational high schools (SMK) that do not utilize or teach this technology to their student participants will be outdated and unable to compete with SMK or other schools that are adaptive to the technology used in Industry 4.0.

SMK needs to develop learning programs by at least utilizing four of the five central technologies of Industry 4.0. There is one Industry 4.0 technology that not all SMKs can teach or use for reasons of expertise, this technology is production hardware (Regulation of the Directorate General of Primary and Secondary Education of the Ministry of Education and Culture No: 06 / D.D5 / KK / 2018). Four technologies that can be utilized in SMK to develop learning programs include computing hardware, interfaces, software, and connectivity. The four technologies are already in pockets, bags, or in the learning environment of vocational students in the form of gadgets, laptops, and WIFI in the school.
environment.

The use of Industry 4.0 technology in SMK can be done by implementing a Learning Management System (LMS) or a learning management system. LMS is an internet network-based software designed to manage electronic learning (e-learning) programs. One well-known LMS is Moodle (short for Modular Object-Oriented Dynamic Learning Environment). Moodle is open-source software that can be used freely without having to pay because it is natural that Moodle users are very many. Currently, there are 104,361 sites, 21,000,000 lectures or subjects, 179,000,000 users, and 232 countries that use Moodle (Statistics, taken from https://stats.moodle.org/ on February 6, 2020). The trend of using Moodle continues to increase as the software features improvements and demands for freer learning.

In the Industrial era 4.0, one of the changes in the way learning outcomes are assessed is the use of an internet-based assessment application (Magno, 2019). The advantage of an internet-based assessment application is its ability to present a test with a difficulty level that matches the ability of the test taker. Such tests are called adaptive tests. The items in the adaptive test are taken from a collection of questions (question banks) which have been categorized based on content and level of difficulty (Magis & von Davier, 2017). The adaptive test can accurately determine the ability of the test taker even though the number of questions and processing time varies between test-takers. Also, being a computer-based test, the adaptive test can show a test taker's final score as soon as the test is finished. As a network-based test, the adaptive test can present the results of all test takers to the teacher or system administrator. Moodle, which is installed with the Adaptive Quiz module, can be used to conduct adaptive assessments. Moodle's online quiz capabilities can be upgraded to adaptive by installing the Adaptive Quiz plugin (https://moodle.org/plugins/mod_adaptivequiz). Unfortunately, this LMS ability has not been used in SMK. Because it is necessary to pioneer the use of adaptive class assessment in SMK.

This study aims to develop a class assessment tool for the SMK Electricity Engineering Expertise Program that can be used in LMS-based adaptive assessment and test the appropriateness of the class assessment tool.

2. Research Method
The type of research used in this research is research and development (Research and Development). The development research model used is the scale construction process (Furr, 2011: 6) with the articulate construct and context steps, choose response format and assemble the initial item pool, collect data from respondents, examine psychometric properties and quality, so that a final scale is obtained.

The LMS-based adaptive classroom assessment tool in this study consists of three (3) types, namely assessment of knowledge aspects, assessment of aspects of attitude, and assessment of aspects of skills (psychomotor). The knowledge aspect assessment tool is made in a multiple-choice form. The attitude assessment tool is made in the form of a rating scale for case assessments. The skills assessment tool is made in the form of an observation guide. The assessment tools for the aspects of knowledge and attitudes were filled by students, while the tools for assessing skills were filled by the teacher.

There are six types of research objects. First is the skill competency of SMK. In this study, the competency of expertise to be selected is Electrical Power Installation Engineering. The skill competency selection technique is done purposively because the development of assessment tools becomes more focused. The second object is the level or class at SMK, between level X to level XII. The level selection is done purposively also because at this level the competence of students' expertise should ideally be formed.

The third object is the name of the subject for which the assessment tool will be developed. In this study, the name of the subjects for which the assessment tool was developed is Electrical Lighting Installation. The fourth object is the SMK where the research was conducted. Determination of the SMK where the research was carried out randomly. The selected SMKs are SMK Muhammadiyah 3.
Yogyakarta and SMK Negeri 3 Yogyakarta. The fifth object is students. The selection of students is done randomly. From SMK Muhammadiyah 3 Yogyakarta, 17 students were obtained, while from SMK Negeri 3 Yogyakarta there were 26. While the sixth object was the teacher, the selection of teachers purposively obtained 3 teachers from each selected school.

The instruments developed include cognitive, affective, and psychomotor domains. The cognitive domain measures the competence of the Electrical Lighting Installation subject. The domain of attitude measures the achievement of IP-21CSS (Indonesian Partnership for 21 Century Skill Standard) which consists of 4C components, namely Creativity, Critical Thinking, Communication, and Collaboration. The psychomotor domain measures work preparation, processes (systematics & work methods), work results, work attitudes, and time.

The quality of the validity of the instrument is proven by calculating the V index of Aiken. If the index is less or equal to 0.4, it is said that the validity is low, 0.4-0.8 is said to have moderate validity, and if it is> 0.8 it is said to be valid (Retnawati, 2015). The empirical quality of the items and the attitude assessment questionnaire was seen with the classical test theory, by calculating the total-item biserial correlation for the test, and the item-total product-moment correlation for the questionnaire. Item is said to be good if the value of the total correlation value> = 0.3. Besides, the quality of the items was also seen by analyzing the Rasch model’s Item Response Theory (IRT). The reliability of the observation guidelines for assessing practice is calculated by calculating inter-rater reliability using the Fleiss Kappa formula (Fleiss, 1971).

3. Research Results and Discussion
The construct of the classroom assessment tool in this study was chosen which is easy to implement in a computer program. The questions are made in the multiple-choice form with 5 alternative answers. Attitude questionnaires are made in the form of case assessments with 4 alternative answers, while the observation guidelines are made in the form of short entries with numbers between 1 and 4.

The results of testing the validity of the questions by the experts show that all items are valid. Of the 60 items, it turns out that the items with less validity are 0 or 0%, 55 or 91.67% are currently available, and the high ones are 5 or 8.34% as shown in Figure 1. The results of empirical reliability calculations using Cronbach’s Alpha with 43 test participants students get a value of 0.934. Thus it can be concluded that the existing questions are valid and reliable.

![Figure 1. Distribution of Question Item Validity](image)

If the quality of the items in terms of the total-item biserial correlation and the correlation coefficient of less than 0.3 is categorized as bad and greater than 0.3 is said to be good, it turns out
that from the 60 items, the bad items are 9 or 15% and the good ones are 51 or 85% as shown in Figure 2. There are 8 bad items but the correlation is positive and small, such items can still be used. One other item has a negative and small correlation, such items need to be corrected or discarded.

![Figure 2. Quality of Question Items based on Biserial Correlation](image)

The results of the calculation of difficulty using the IRT with the Rasch model for multiple-choice questions obtained a minimum value = -2.837, maximum = 1.858, average = 0.283, and standard deviation = 0.766. The results of the item fit test show that the items that are not fit = 10, while the fit items = 50 as shown in Figure 3.

![Figure 3. Number of Fit and Unsuitable Test Items Based on Rasch Analysis](image)

The results of testing the validity of the questionnaire by experts show that all items are also valid. However, the level of validity of the items was not the same. If grouped based on the level of validity, items with less validity are 0 or 0%, the validity is 6 or 75.00%, and the high is 2 or 25.00% as shown in Figure 4. The results of empirical testing with respondents as many as 43 people obtained Alpha reliability coefficients, equal to 0.731. Thus it can be concluded that the attitude questionnaire developed in this study is also valid and reliable.
If the quality of the items is viewed from the total-item product-moment correlation and the correlation coefficient is less than 0.3, it is categorized as not good and is greater than 0.3 which is said to be good, from 8 items it is obtained that the items are not good is 1 or 12.50% and the good ones are 7 or 87.50% as shown in Figure 5. These bad items have positive and small correlation, such items can still be used because they do not cause negative judgments.

The results of the validity test by experts on the existing observation guidelines obtained items whose validity was less than 0 or 0%, items with moderate validity were 15 or 93.75%, and the high ones were 1 or 6.25% as shown in Figure 6. Because the observation guidelines had not been tried to assessing students' practical abilities, then its reliability has not been proven.
Figure 6. Distribution of the Validity of the Observation Guidance Items

Seeing the number of valid and good items from the questions and questionnaires as well as the valid items from the observation guidelines as described, it can be concluded that the questions, questionnaires, and observation guidelines can be used in adaptive classroom assessment tools based on the Learning Management System. However, some instruments with poor quality need to be repaired or discarded.

4. Conclusions and Recommendations

The construct of the classroom assessment tool is chosen which is easy to implement in a computer program. Multiple choice questions with 5 alternative answers. The attitude questionnaire is made in the form of a short assessment with 4 alternative answers, while the observation guidelines are made in the form of a short entry with numbers between 1 and 4. The questions were developed to consist of 60 items. The items with less validity are 0, there are 55, and 5 are high with Alpha reliability of 0.934. The attitude questionnaire developed consists of 8 items. Items with less validity are 0, currently there are 6, and high 2 with Alpha reliability of 0.731. The observation guideline developed contains 16 points of observation. Items with low validity are 0, the moderate is 15, and the high is 1.

Multiple choice questions with 5 alternative answers, attitude questionnaires made in the form of case assessments with 4 alternative answers, and observation guidelines made in the form of short entries with numbers between 1 and 4 ready to be implemented in an adaptive classroom assessment system based on the Learning Management System.

References
[1] Adaptive Quiz, downloaded from https://moodle.org/plugins/mod_adaptivequiz on 9 February 2020
[2] Fleiss, J. L. (1971) "Measuring nominal scale agreement among many raters." Psychological Bulletin, Vol. 76, No. 5 pp. 378–382.
[3] Furr, R. M. (2011) Scale construction and psychometrics for social and personality psychology, London: SAGE Publications.
[4] Magis, D., Yan, D. & von Davier, A.A. (2017). Computerized adaptive and multistage testing with R. Switzerland: Springer International Publishing.
[5] Magno, C. (2019). What happens to educational assessment in the 4th industrial revolution. Retrieved from https://www.pemea.org/single-post/2019/02/03/What-Happens-to-Educational-Assessment-in-the-4th-Industrial-Revolution on 19 July 2019.
[6] Regulation of the Director-General of Primary and Secondary Education of the Ministry of Education and Culture No: 06 / D.D5 / KK / 2018 concerning the Spectrum of Vocational High Schools (SMK) / Madrasah Aliyah Vocational Schools (MAK).
[7] Retnawati, H. (2015). Validitas reliabilitas dan Karakteristik Butir. Yogyakarta: Parama Publishing.
[8] Statistics. Retrieved from https://stats.moodle.org/ on February 6, 2020.