



# Developing Classroom Assessment Tool using Learning Management System-based Computerized Adaptive Test in Vocational High Schools

Samsul Hadi<sup>1</sup>, Haryanto<sup>2</sup>, Muh. Asriadi AM<sup>3</sup>, Marlina<sup>4</sup>, Abdul Rahim<sup>5\*</sup> 

<sup>1,2</sup> Department of Electrical Engineering Education, Engineering Faculty, Yogyakarta State University, Yogyakarta, Indonesia

<sup>3,4,5</sup> Department of Education Research and Evaluation, Faculty of Graduate School, Yogyakarta State University, Yogyakarta, Indonesia

## ARTICLE INFO

### Article history:

Received June 14, 2021

Revised June 15, 2021

Accepted November 24, 2021

Available online February 25, 2022

### Kata Kunci :

Adaptif, Sistem Manajemen Pembelajaran, Penilaian Kelas

### Keywords:

Adaptive, Learning Management System, Classroom Assessment



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright ©2022 by Author. Published by Universitas Pendidikan Ganesha

## ABSTRAK

Komputer telah mengambil peran besar dalam pendidikan, termasuk pengujian dan evaluasi. Tes tradisional yang tidak komprehensif dan tidak membedakan antara bakat awal siswa mengarah pada temuan pengukuran yang tidak mewakili kemampuan mereka yang sebenarnya. Penelitian ini bertujuan untuk mengembangkan dan menguji kelayakan alat penilaian kelas yang digunakan sebagai adaptif berbasis LMS. Jenis penelitian ini termasuk penelitian pengembangan. Responden penelitian ini adalah para ahli yang menilai validasi dan siswa SMK Kompetensi Keahlian Teknik Instalasi Tenaga Listrik. Teknik analisis data menggunakan teori respon butir, teori tes klasik dan statistik deskriptif. Analisis butir soal dengan Model Rasch menunjukkan 10 butir soal tidak sesuai dan sisanya 50 butir soal sesuai. Butir analisis teori tes klasik dengan validitas kurang ada 0 butir, sedang ada 55 butir, dan tinggi ada 5 butir, dengan reliabilitas Alpha sebesar 0,934. Kuesioner sikap yang dikembangkan terdiri dari 8 item. Ada 0 item dengan validitas kurang, 6 item sedang, dan 2 item tinggi, dengan reliabilitas Alpha 0,731. Panduan observasi yang dikembangkan berisi 16 observasi. Item dengan validitas kurang memiliki 0 item, sedangkan item ada 15 item, dan tinggi 1 item. Penggunaan sistem manajemen pembelajaran berbantuan CAT pada perangkat penelitian sangat memudahkan guru dalam melakukan penelitian secara akurat dan praktis.

## ABSTRACT

Computers have taken on a large role in education, including testing and evaluation. Traditional tests that aren't comprehensive and don't distinguish between students' beginning talents lead to measurement findings that aren't representative of their true abilities. This study aims to develop and test assessment tool eligibility class that is used as an LMS-based adaptive. This type of research includes development research. The respondents of this study were experts who assessed the validation and students of SMK Electrical Power Installation Engineering Expertise Competencies. The data analysis technique used item response theory, classical test theory and descriptive statistics. Item analysis using the Rasch Model showed 10 items were not fit and the remaining 50 items were fit. Classical test theory analysis items with less validity there are 0 items, moderate there are 55 items, and high there are 5 items, with an Alpha reliability of 0.934. The attitude questionnaire developed consists of 8 items. There are 0 items with less validity, 6 items being moderate, and 2 items high, with an Alpha reliability of 0.731. The developed observation guide contains 16 observations. Items with less validity have 0 items, while there are 15 items, and 1 item high. The use of CAT-assisted learning management systems on research tools greatly facilitates teachers in conducting research accurately and practically.

## 1. INTRODUCTION

The fourth industrial revolution era (industry 4.0) has impacts on all aspects in humans' life, and as a current issue which becomes critical in the field of education (Devi et al., 2020). The fact is, in the modern concept, education is a mean for the students to face the challenges of modern era (Hadi et al., 2018). In addition, when dealing with the challenges of the fourth industrial revolution era, an appropriate education is needed in order to build a creative, innovative, and competitive generation (Lase, 2019). In industry 4.0, 5 central technologies were used in order to support the education process and as a form of significant development which becomes the main characteristic of education 4.0 (Hussin, 2018). Those five central technologies are: computing hardware, production hardware, software, interfaces, and connections (Abidin et al., 2019). The continuous development on the paradigm of revolution has initiated by the development of sciences and technology as a form of innovation (Liao et al., 2018). Moreover, each level of education utilizes technology in the learning process, and one of them is Vocational High School (SMK) (Yunis & Telaumbanua, 2017). If a Vocational High School does not utilize technology in the teaching and

\*Corresponding author.

E-mail addresses: [samsulhad44@gmail.com](mailto:samsulhad44@gmail.com) (Samsul Hadi)

learning process, then the school will not be able to compete with other Vocational High Schools or other High Schools that have adapted with the technology in 4.0 industry (Batubara, 2017). In the fourth industrial revolution era, education and technology cannot be separated, especially in Vocational Schools (Hardyanto & Surjono, 2016). In addition, the developments in the fourth industrial revolution era are conducted consciously and systematically, which are in accordance with Indonesian' future education goals and have good impacts (Abbasi et al., 2018).

The development of learning program in Vocational Schools needs at least 4 of 5 central technologies in the fourth industrial revolution era, since from 5 central technologies, one of them cannot be conducted in Vocational High Schools since it depends on the expertise program, and it is the production hardware (Putra et al., 2020). Meanwhile, four technology centrals that can be utilized in Vocational High Schools are computing hardwares, interfaces, softwares, and connections, while the utilization of technology in this era will help the students in achieving the learning goals. The four central technologies can be accesses by using gadget, laptop, and WIFI as the facilities provided in the school. The changes are seen from how the conventional technology turns into smart technology (Haryanto, 2013). One of the characteristics of 4.0 industrial revolution is the implementation of artificial intelligence as well as the utilization of 4.0 industry in vocational high schools, by implementing Learning Management System (LMS) (Codish et al., 2019). LMS is an internet-based software designed to control the e-learning program (Hu, X., Ng et al., 2020). LMS has certain characteristics, such as in the management of the learning content, the management of learning process, evaluation, by implementing online examination system, as well as the administration process that can be completed using online chatting and discussion forum (Jana & Khatun, 2021). One of outstanding LMS is Moodle (Modular Object-Oriented Dynamic Learning Environment).

Moodle is an Open Source Course Management System (CMC) which is also known as Learning Management System (LMS) or Virtual Learning Environment (VLE) which plays roles in the development of technology-based learning, in the form of website (Chaw & Tang, 2018). Moodle is an open source software that can be used for free, and it increases the number of Moodle users. As noted in 104.361 sites or, 21.000.000 courses or subjects, 179.000.000 users, and 232 countries that use Moodle (Copriady et al., 2020). The Moodle users tend to be increasing as well as the improvement of the features of the software, and the learning demands that are not limited. Moodle has several advantages, such as having complete features of communication media (chatting, messaging, or discussion forum), creating and administrating the learning materials, tracking the data of students' learning outcomes (data tracking), and its function can be extended (extensibility plugin); After that, Moodle is easy to use, since there is a flexibility to arrange the application based on the institutions' needs and policy in the learning process. Currently, in Indonesia, many still use classical test theory to analyze tests and estimate their abilities (Fajrianti et al., 2016), but in the use of classical test theory, and the measurement results are far from the actual ability (Abidin et al., 2019; Asriadi & Istiyono, 2020). Classical test theory has many shortcomings, one of which is that the characteristics of the questions depend on the examinee (Istiyono et al., 2018; Wise et al., 2015). For example, if answered by an intelligent participant, the question will be a simple question, otherwise if it is answered by a less intelligent participant, the question will be difficult. Likewise the score that reflects the ability of the participant depends on the. If this happens, we will never know the true ability level of the examinees. The solution to this problem is to apply item response theory (IRT) (Van der Linden, 2018). This is because the skill level of participants and project parameter estimates are measured accurately.

4.0 industrial revolution era is one of the ways to change the assessment strategy by using internet (Febliza & Okatariyani, 2020). One of the advantages of internet-based assessment application is its ability to present a test whose difficulty level is appropriate with students' capability. In addition, adaptive testing model can be used in dealing with the weakness of educational assessment, and the test is known as adaptive test or tailored test, or a test that has been designed based on the students' capability (Aybek & Demirtasli, 2017). The adaptive test is designed by determining the questions / items based on the participants' initial answers (responses) (Kezer, 2021). Adaptive test can accurately determine the participants' ability, even though there are various questions and working times done by each participant (A. Sahin et al., 2018). In addition, as a computer-based test, an adaptive test can show the final participant value immediately after the test is completed; and as a network-based test, an adaptive test can present all results to the teacher or administrator of the system (Sahin & Gelbal, 2020). Moreover, Adaptive Quiz module installed in Moodle application can be used to perform adaptive assessments (D. I. Sahin & Shelley, 2020). However, LMS has not been used in Vocational High Schools. Therefore, it is necessary to use of Learning Management Systems (LMS)-based Computerized Adaptive Test (CAT) in Vocational High Schools.

Relevant research conducted by (Abidin et al., 2019) in developing a CAT to measure the HOTS of 10th grade high school students in physics. Further research conducted in developing CAT to measure critical physics skills. Research conducted by other researcher in developing CAT with fuzzy logic algorithms in describing students' abilities and inference system models in making decisions to choose the

right test items for students (Haryanto, 2013). Based on previous research, that this study aims to develop a computer-based adaptive classroom assessment tool or Computerized Adaptive Test (CAT) based on Learning Management Systems for the Electrical Engineering Skills Program in Vocational High Schools and test the feasibility of classroom assessment tools for the Electrical Engineering Skills Program in Vocational Schools.

## 2. METHODS

In this study, the researcher implemented Research and Development (R&D) method and scale construction process as the development model (Gustiani, 2019). First, an application of LMS was conducted as a trial in Vocational High Schools, and then the data were analyzed and reported. Second, the application of entire LMS was conducted, and then the data were analyzed and reported again. The whole research processes were presented in Figure 1.

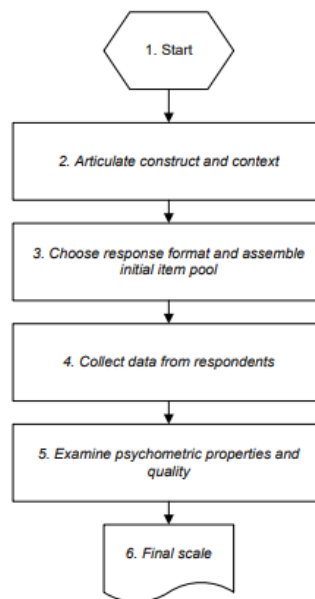


Figure 1. Development Step

There were six types of population of the research. First, the expertise of the students in the Vocational High Schools, since in this research, the expertise that was selected is Electrical Engineering. Hence, the expertise was selected by applying purposive sampling, in order to focus on developing the assessment tool. Second, the population is the students' level in Vocational High Schools, from the tenth grade to the twelfth grade, and the selection was conducted by applying purposive sampling, with the condition that the students have had an ideal capability on the expertise. Third, the population is the course whose assessment will be developed, and in this research is Electrical Installation. The fourth population is the vocational high school where the research was conducted. The determination of the schools was done randomly, and the selected schools were SMK Muhammadiyah 3 Yogyakarta and SMK Negeri 3 Yogyakarta. The fifth population was the student(s). The selection of students was also done randomly, such as 17 students from SMK Muhammadiyah 3 Yogyakarta, and 26 students from SMK Negeri 3 Yogyakarta. Finally, the sixth population was the teacher(s) that were selected purposively, and 3 teachers were selected from each school.

The instruments used in collecting the data were: 1. A test in the form of multiple choices in order to assess the students' knowledge related to Electrical Installation in XI 2 class of Electrical Installation Engineering Expertise Program; 2. Rating scale for assessing the students' attitude; 3. Observation Guidelines for assessing the students' skills; 4. Likert scale for assessing the appropriateness of the instruments on LMS-based Computerized Adaptive Test. In attachment 1, the lattice of the multiple-choice test is presented. The assessment on students' attitude is designed based on IP-21CSS (*Indonesian Partnership for 21 Century Skill Standard*), consist of 4C: Creativity, Critical Thinking, Communication, and Collaboration.

### 3. RESULT AND DISCUSSION

#### Results

##### Ensuring the Construction and Context of the Instrument

The classroom assessment tools developed in this study are: a) Electrical Installation test for the twelfth-grade students of Electrical Installation Engineering Expertise Program; b) questionnaires for assessing students' attitude; and c) observation guidelines for assessing students' practical skills. The assessment tools and the practical assessment tools are presented in the following section. The construct of the classroom assessment tool in this study was chosen based on its level of difficulty in order to be applied in a computer program. The questions are in the form of multiple choices with 5 alternative answers. Meanwhile, the questionnaires were designed in the form of case assessments with 4 alternative answers in order to assess the students' attitude, while the observation guideline contents consist of short entries, ranged from 1 to 4.

##### Selecting the format of the instrument

A Mathematical test on Electrical Installation for the twelfth-grade students of Electrical Installation Engineering Expertise Program was analyzed with multiple choices. This type was chosen since it is easier to implement in CAT and the answer is certain. The last participant can choose or click the alternative answer that is considered as the best answer or the most correct. The participant's answer will be corrected directly by the computer and compared with the answer key. The questionnaire on assessing students' attitude uses a rating scale format based on the case study. The context is related to the participants' daily life at school and in the community. Then, the participants were asked to answer by choosing one of the available answers. The alternative for each answer has different weight values from 1 to 4. This model was chosen since it has been implemented in CAT or CBT, the participants' value depends on the alternative answers that are clicked by the students. The observation guideline for assessing students' practical skills uses a format for assessing the subcomponents of students' skills, conducted by the teacher. In addition, the observation guideline is complemented by the assessment grid, so it contains subcomponents of skills that can be assessed by the teacher based on the students' abilities, as presented in the rubric. The observation guideline is also easy to implement in CAT or CBT since because each item is already in the form of numbers.

##### Data Collection

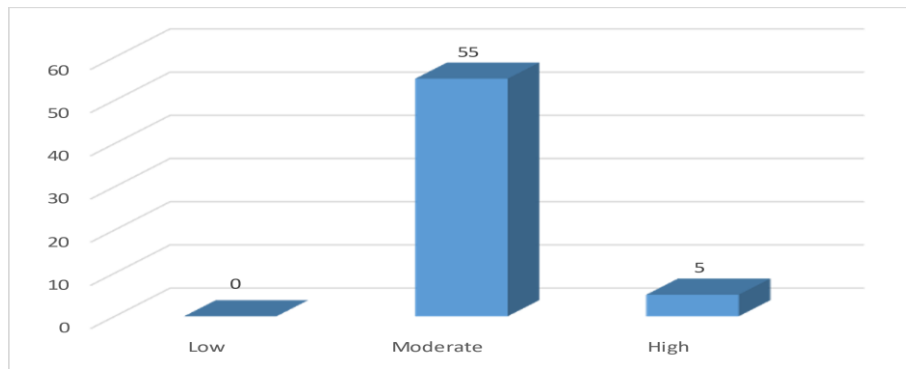
The data collected in this study were obtained from the validation of the experts and students' responses. Before testing the instruments on the real users, the instruments were validated by the experts, who consisted of 6 persons. The validators were asked to assess each item of the instruments, ranged from 1 to 4, while 1 = very invalid, 2 = invalid, 3 = valid, and 4 = very valid. After validating the items of the instruments, then the instruments were given to the users. Questionnaire for assessing students' attitude was given online, while the observation guidelines were not tested, since the research was conducted during Covid-19 pandemic.

##### The Test on Characteristics and Psychometric Quality

###### Validity and Reliability

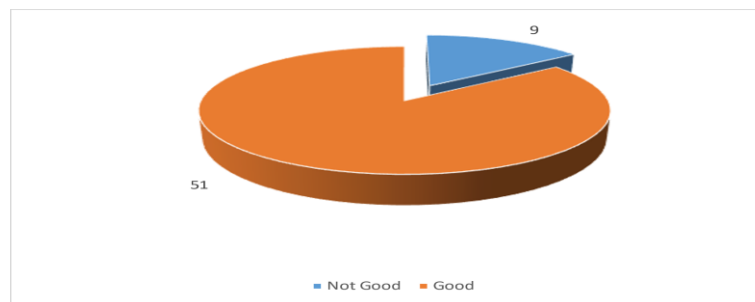
###### Validity and Reliability of the Test

The results of the analysis of the validity using Aiken's formula from 60 items obtained index  $V$  minimum = 0.444, maximum = 0.833, mean = 0.661, and standard deviation = 0.096, if the validity is classified into less, moderate, and high then the items which have less validity are 0 or 0%, the items which have moderate validity are 55 or 91.67%, and the items which have high validity is 5 or 8.34%. The results of the validity of the questions done by the experts indicate that all items are valid. From 60 items, and the items which have less validity are 0 or 0%, the items which have moderate validity are 55 or 91.67%, and the items which have high validity is 5 or 8.34%. Meanwhile, the value obtained from the reliability calculation using Cronbach's Alpha on 43 test participants is of 0.934. Thus, it can be concluded that the questions are valid and reliable, as presented in [Figure 2](#).



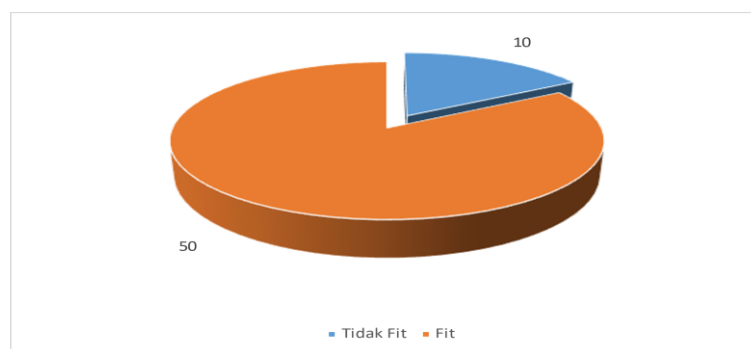
**Figure 2.** The Distribution of Validity of the Test Item

If the quality of the items in terms of total-item biserial correlation and a correlation coefficient is less than 0.3, then it is categorized as bad, and if it is greater than 0.3, it is categorized as good. From 60 items, 9 items or 15% of the items are bad, and 51 items or 85% of the items are good as presented in Figure 3. Even though several items are bad, but the correlation is positive and small, and there are 8 items that can still be used. If an item has a negative and small correlation, such items need to be corrected or discarded.



**Figure 3.** The Quality of the Test Item based on Biserial Correlation

Based on the analysis of multiple-choice items by using Rasch IRT model, the minimum value = -2.837; maximum = 1.858; average = 0.283; and standard deviation = 0.766. In other words, the questions commonly have a moderate difficulty level and no questions are very difficult. In addition, based on the analysis, from 60 items, 10 items are not fit with Rasch model, while 50 items are fit. The proportion of items that are fit and not fit with Rasch model is presented in Figure 4.



**Figure 4.** The Proportion of Fit and Unfit Test Items

### Validity and Reliability of the Questionnaire

Aiken analysis on the validity of the 8 items obtained index  $V$  minimum = 0.5, maximum = 0.833, mean = 0.673, and standard deviation = 0.112. If the validity is classified into less, medium, and high, then the items which have less validity are 0 or 0%, the items which have moderate validity are 6 or 75.00%, and the items which have high validity is 2 or 25.00%. That are shown on Figure 5.

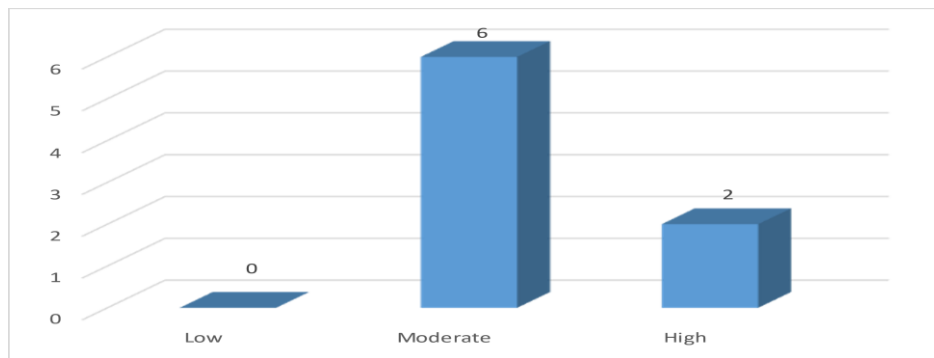


Figure 5. Distribution of Test Item Validity

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 bad, and if it is greater than 0.3, it is categorized as good. Then, the result reveals that from 8 items, 1 item or 12.50% of the items are bad, and 7 items or 87.50% of the items are good, as presented in Figure 6. These bad items have positive and small correlation, but such items can still be used since they do not cause negative judgments.

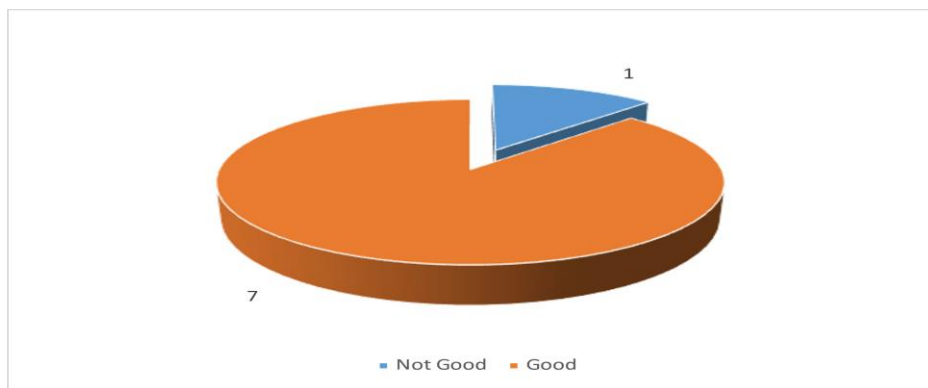


Figure 6. The Quality of the Test Item based on Product Moment Correlation

**Validity of the Observation Guideline**

Aiken's analysis on the validity of the 16 items on the observation guideline obtained index V minimum = 0.5, maximum = 0.833, mean = 0.694, and standard deviation = 0.097. If the validity is grouped into less, medium, and high, then the items which have less validity are 0 or 0%, the items which have moderate validity are 15 or 93.75%, and the items which have high validity is 1 or 6.25%. The result of validity testing by experts on the observation guideline reveals that 0 item or 0% of the items have less validity, 15 items or 93.75% of the items have moderate validity, and 1 item or 6.25% of the items have high validity, as presented in Figure 7. Since the observation guideline has not been tested to assess students' practical skills, the reliability has not been proven.

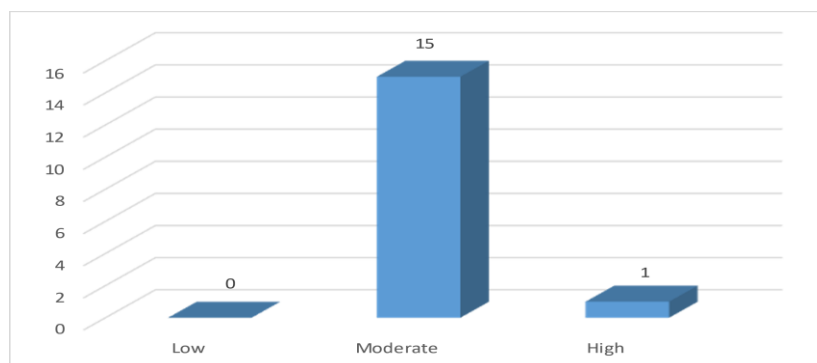


Figure 7. The Distribution of the Validity of Items in the Observation Guideline

By identifying the number of valid and good items from the test and questionnaires as well as the valid and good items from the observation guideline, it can be concluded that the question items of the test, questionnaires, and observation guidelines can be used in Learning Management System-based Computerized Adaptive Test (CAT).

### The Correlation of Item - Total

This analysis aims to identify good and bad question items or statements used in the data collection process. At this stage, the correlation will be analyzed, namely tests and questionnaires. Biserial Correlation of Item Test – Total shown on [Table 1](#).

**Table 1.** Biserial Correlation of Item Test – Total

Item	Item Mean	P Bis	Bis	Alphalf Deleted	Category
B1	0.465116	0.33346	0.41851	0.933607467	Good
B2	0.651163	0.48256	0.62168	0.932659375	Good
B3	0.674419	0.47653	0.61997	0.932703055	Good
B4	0.55814	0.29566	0.37199	0.933845951	Good
B5	0.906977	0.49908	0.8711	0.932951133	Good
B6	0.27907	0.24633	0.32876	0.934055179	Good
B7	0.534884	-0.0003	-0.0004	0.935724544	Not Good
B8	0.488372	0.25952	0.32531	0.934083635	Good
B9	0.325581	0.37389	0.48643	0.933326806	Good
B10	0.232558	0.15771	0.21807	0.93448794	Not Good
B11	0.72093	0.19496	0.2602	0.934349805	Not Good
B12	0.302326	0.0658	0.08661	0.935131269	Not Good
B13	0.186047	0.30665	0.44548	0.933661156	Good
B14	0.511628	0.44383	0.55634	0.932895447	Good
B15	0.302326	0.24713	0.32531	0.934071494	Good
B16	0.627907	0.09606	0.12275	0.935052865	Not Good
B17	0.418605	0.36732	0.46391	0.933384743	Good
B18	0.44186	0.34981	0.44013	0.933499857	Good
B19	0.651163	0.47478	0.61165	0.932707675	Good
B20	0.348837	0.19404	0.24998	0.934431506	Not Good
B21	0.27907	0.40556	0.54128	0.933135313	Good
B22	0.395349	0.56075	0.71186	0.932150859	Good
B23	0.488372	0.4547	0.56997	0.932824858	Good
B24	0.604651	0.57079	0.72462	0.932086487	Good
B25	0.488372	0.27049	0.33906	0.934013387	Good
B26	0.627907	0.38268	0.48901	0.933280623	Good
B27	0.55814	0.20377	0.25638	0.934429949	Not Good
B28	0.534884	0.50855	0.63826	0.932475222	Good
B29	0.325581	0.49218	0.64033	0.932607567	Good
B30	0.395349	0.45798	0.5814	0.932806644	Good
B31	0.55814	0.48889	0.61513	0.932604458	Good
B32	0.55814	0.68886	0.86672	0.931300468	Good
B33	0.534884	0.71557	0.89808	0.931116845	Good
B34	0.511628	0.56676	0.71044	0.932093595	Good
B35	0.511628	0.64938	0.814	0.931550502	Good
B36	0.395349	0.47315	0.60066	0.93271013	Good
B37	0.44186	0.62685	0.7887	0.931706974	Good
B38	0.44186	0.74469	0.93697	0.930932878	Good
B39	0.395349	0.70677	0.89723	0.931210454	Good
B40	0.465116	0.64288	0.80686	0.931596191	Good
B41	0.325581	0.44474	0.57861	0.932896737	Good
B42	0.534884	0.54968	0.68989	0.932206985	Good
B43	0.348837	0.64744	0.83409	0.931629857	Good
B44	0.27907	0.28699	0.38303	0.933821239	Good
B45	0.511628	0.72115	0.90397	0.931075862	Good
B46	0.395349	0.3899	0.49498	0.933238248	Good

Item	Item Mean	P Bis	Bis	Alphalf Deleted	Category
B47	0.255814	0.38064	0.51635	0.93327862	Good
B48	0.372093	0.3303	0.42208	0.933607312	Good
B49	0.348837	0.47125	0.6071	0.932729561	Good
B50	0.534884	0.48616	0.61016	0.932620821	Good
B51	0.395349	0.59902	0.76044	0.931905383	Good
B52	0.302326	0.1759	0.23154	0.93448943	Not Good
B53	0.255814	0.37643	0.51064	0.93330236	Good
B54	0.395349	0.40878	0.51894	0.933118801	Good
B55	0.465116	0.53019	0.66542	0.932334226	Good
B56	0.418605	0.31144	0.39334	0.93373993	Good
B57	0.395349	0.53784	0.68278	0.932297468	Good
B58	0.325581	0.2061	0.26813	0.934336654	Not Good
B59	0.511628	0.48468	0.60755	0.932629846	Good
B60	0.418605	0.48732	0.61547	0.932616819	Good
<b>Mean</b>			<b>= 0.542</b>		
<b>Standard Deviation</b>			<b>= 0.218</b>		

From Table 1. Analysis of biserial correlation on the test items obtained a minimum value = -0.0004, maximum = 0.936, mean = 0.542, and standard deviation = 0.218, while the criteria of biserial correlation coefficient is: <0.300 is bad and> = 0.3 is good, and it can be concluded that from 60 items, 9 items or 15% of the items are bad, and 51 items or 85% of the items are good. Then Bisserial Correlation of Item of the Questionnaire – Total is shown on Table 2.

**Table 2.** Bisserial Correlation of Item of the Questionnaire – Total

Item	Item Mean	Pearson Correlation	Alphalf Deleted	Category
B1	3.51162791	0.484041862	0.303080308	Good
B2	3.39534884	0.409233549	0.344484092	Good
B3	3.55813953	0.649722305	0.4287629	Good
B4	3.41860465	0.606425735	0.216942149	Good
B5	3.37209302	0.607028161	0.215938303	Good
B6	3.58139535	0.529353323	0.436457506	Good
B7	3.41860465	0.606425735	0.216942149	Good
B8	3.53488372	0.282866898	0.432946357	Not Good
<b>Mean</b>		<b>= 0.512</b>		
<b>Standard Deviation</b>		<b>= 0.224</b>		

Based on Table 2, analysis of the product moment correlation item - the total items of the questionnaire obtained the minimum value = 0.282, the maximum = 0.649, mean = 0.512, and the standard deviation = 0.224, with the criteria of biserial correlation coefficient: <0.300 is bad and> = 0.3 is good, and it can be concluded that from 8 items, 1 item or 12.50% of the items are bad, and 7 items or 87.50% of the items are good. The difficulty level of the test item are shown on Table 3.

**Table 3.** Item Difficulty Level of the Test Item

Item	S_X2	df.S_X2	RMSEA.S_X2	p.S_X2	Category
B1	17.6521	7	0.190346	0.013642	Not Fit
B2	8.706704	5	0.132857	0.12135	Fit
B3	4.793633	5	0	0.441581	Fit
B4	1.666938	7	0	0.975962	Fit
B5	#N/A	0	0	0	Not Fit
B6	5.876848	5	0.064618	0.318387	Fit
B7	23.36763	8	0.213862	0.002923	Not Fit
B8	8.139249	7	0.062249	0.320475	Fit
B9	11.10785	5	0.170543	0.049283	Not Fit
B10	6.27729	4	0.116427	0.179375	Fit
B11	8.229092	5	0.124003	0.144054	Fit



Item	S_X2	df.S_X2	RMSEA.S_X2	p.S_X2	Category
B12	11.22261	5	0.172138	0.047141	Not Fit
B13	4.217551	3	0.098301	0.238911	Fit
B14	15.54175	7	0.170451	0.029649	Not Fit
B15	6.063697	5	0.07117	0.30007	Fit
B16	18.20928	6	0.220113	0.00573	Not Fit
B17	11.05106	8	0.095292	0.198823	Fit
B18	10.18902	7	0.104149	0.178113	Fit
B19	5.917389	5	0.066095	0.31434	Fit
B20	7.280882	6	0.071294	0.295651	Fit
B21	0.172813	5	0	0.999379	Fit
B22	5.730141	7	0	0.571589	Fit
B23	7.427425	7	0.038129	0.385772	Fit
B24	5.706671	6	0	0.456838	Fit
B25	13.58395	7	0.149648	0.059096	Fit
B26	4.262223	6	0	0.641238	Fit
B27	7.506919	7	0.041524	0.378071	Fit
B28	5.740929	8	0	0.676226	Fit
B29	3.11773	5	0	0.681841	Fit
B30	5.954574	7	0	0.545063	Fit
B31	5.923508	7	0	0.548709	Fit
B32	13.37427	7	0.147245	0.063498	Fit
B33	10.2428	8	0.081701	0.248397	Fit
B34	14.23542	7	0.156877	0.047149	Not Fit
B35	9.787565	7	0.097373	0.200935	Fit
B36	10.41164	7	0.107723	0.166422	Fit
B37	8.403594	7	0.069095	0.298353	Fit
B38	12.2101	7	0.133122	0.093858	Fit
B39	9.077308	7	0.084058	0.247153	Fit
B40	5.862853	7	0	0.555852	Fit
B41	7.339135	5	0.10554	0.196615	Fit
B42	15.14244	8	0.145799	0.056434	Fit
B43	10.15488	6	0.128404	0.118279	Fit
B44	7.535264	5	0.109876	0.183777	Fit
B45	16.05028	7	0.175452	0.024661	Not Fit
B46	11.60972	7	0.125217	0.114146	Fit
B47	3.844361	4	0	0.427479	Fit
B48	6.685809	6	0.052168	0.350883	Fit
B49	10.20438	6	0.129167	0.116305	Fit
B50	7.867164	8	0	0.446553	Fit
B51	10.82151	7	0.11401	0.146596	Fit
B52	10.62461	5	0.163658	0.059352	Fit
B53	1.045281	4	0	0.902854	Fit
B54	1.492804	7	0	0.982558	Fit
B55	9.528943	7	0.092746	0.216876	Fit
B56	19.5958	8	0.185772	0.011978	Not Fit
B57	5.303469	7	0	0.622983	Fit
B58	8.739286	5	0.13344	0.119926	Fit
B59	8.238586	7	0.064907	0.312027	Fit
B60	8.896436	8	0.051652	0.351108	Fit
<b>Mean</b>			<b>= 0.283</b>		
<b>Standard Deviation</b>			<b>= 0.766</b>		

Base on Table 3, the IRT analysis using Rasch model on the participants' responses in the multiple choice test obtained a minimum value = -2.837; maximum = 1,858; mean = 0.283; and standard deviation = 0.766. and it can be concluded that from 60 items, 10 items are not fit with the Rasch model, while 50 items are fit.

## Discussion

The development of an adaptive classroom assessment device based on a learning management system is a set of assessment forms consisting of three types, namely the assessment of knowledge aspects in the form of multiple choice, assessment of aspects of attitudes made in the form of a graded scale (rating scale), case assessment, and assessment of skill aspects (psychomotor). It was made in the form of an observation guide. The assessment tool is used in assessing learning activities in the electrical engineering expertise program in SMK. This learning device is made based on a learning management system by applying the CAT system. Currently the assessment no longer has to be done manually but can be made in a digital system (Boonprasom & Sintanakul, 2020). A learning management system is a platform that offers a variety of integrated tools for delivering and managing online instruction via computer devices (Fearnley & Amora, 2020). The implementation of the LMS has made it easy and provides the right digital learning environment for a large number of distance lessons (Noreen, 2020). Assessment with a learning management system serves as the right platform to assess their self-learning skills, so that they become more useful and relevant in today's world of work (Biney, 2020). Adaptive classroom assessment tools based on the learning management system with the CAT system provide the opportunity for teachers to assess comprehensive student competencies in an easy way (Rizal et al., 2020). The standardized learning management system supports an inclusive learning environment for academic progress with an intermediary structure that promotes online collaborative grouping, professional training, discussion, and communication (Bradley, 2020). Learning management system-based assessment forms approach can produce objective values in a practical way (Ngafeeson & Gautam, 2021). So that this assessment tool can be used by teachers or schools in mapping students' abilities more honestly and transparently. The assessment component that is the main emphasis in this assessment tool is the 4C component, namely Creativity, Critical Thinking, Communication, and Collaboration. These components are the focus of competencies that must be achieved in 21st century learning (Asriadi & Istiyono, 2020).

Based on the feasibility analysis, all developed instruments (tests, questionnaires, and observation sheets) have met the valid and reliable criteria. From the results of the Aiken V index calculation, an item or device can be categorized based on its index. If the index is less than or equal to 0.4, it is said to be less valid, 0.4-0.8 is said to be moderately valid, and if it is greater than 0.8 it is said to be very valid. (Retnawati, 2017). Based on the findings from data analysis, the test instrument has 55 test items in the valid category and 5 test items in the very valid category. The questionnaire instrument obtained 6 statement items that were in the valid category and 2 statements that were in the very valid category. In the observation sheet instrument obtained 15 items that are in the valid category and 1 item which is in the very valid category. The results of Cronbach's Alpha reliability calculation from the questions obtained a value of 0.934 which was in the very reliable category, while the questionnaire instrument obtained a value of 0.731 which was in the reliable category. It is supported by previous study that uses validity and reliability to ensure that the measurement of the assessment tool developed is appropriate and implementation in the learning process program runs well (Rudibyani et al., 2020). Thus, the concepts of validity and reliability are needed so that the required information can be obtained accurately and precisely.

A good instrument in addition to meeting the valid and reliable criteria, must also have the conformity of the item function with the overall function of the instrument so it is necessary to do Biserial Correlation analysis. Instrument items are said to meet good criteria if they have a biserial correlation coefficient greater than 0.3 (Dewanti et al., 2020). Based on the results of the Biserial Correlation analysis on the test instrument, 51 test items were in the good category and 9 test items were in the bad category. The questionnaire instrument obtained 7 good statements and 1 bad statement. Especially for test instruments, it is necessary to know the index of difficulty of each item because the index of difficulty of the items will not change even if the questions are done by different respondents. Therefore, it is necessary to carry out an IRT analysis of the Rasch model of the responses to multiple choice questions to determine the level of difficulty of the items. The test item is said to be fit/good if it has a difficulty level item greater than 0.5 (Dewanti et al., 2021). Based on the results of the IRT analysis of the Rasch model, there were 50 items that fit and 10 items that did not fit.

In addition, the use of CAT-assisted learning management systems on research tools greatly facilitates teachers in conducting research accurately and practically. This is in accordance with research conducted by other researcher that learning using the Learning management system makes for a better and more professional understanding (Liu & Geertshuis, 2021). So in general, this assessment tool is suitable for use on different respondents but at the same level. The novelty of this research lies in the research output, namely the output of the assessment tool that is oriented to the knowledge aspect, attitude aspect and skill aspect. This study focuses on the development of an assessment system based on the Learning management system with the CAT system. In the midst of the current situation of technological development, of course, the use of digital devices in assessing the learning process must be applied.

#### 4. CONCLUSION

The construct of the class assessment tool in this study was chosen which is easy to implement in a computer program. The questions are made in the form of multiple choice with five alternative answers. The attitude questionnaire was made in the form of a case assessment with 4 alternative answers, while the observation guidelines were made in the form of short entries with numbers between 1 to 4. All assessment tools had met the valid and reliable criteria based on validity analysis and reliability tests so that the class assessment tool was suitable to be used as a reference. Then teachers to evaluate in measuring the achievement of indicators at the end of learning to achieve predetermined learning objectives. This is seen from the aspects of readability, construction, and content suitability. The limitation in this study lies in the material presented on the test equipment which focuses on the electrical system installation engineering material. As a recommendation in the future for researchers to develop assessment tools made in various formats such as test instruments can be formed in essay format with more diverse material. In addition, it can also be applied to focus on different materials which are expected to increase knowledge and variety and their impact on the achievement of learning objectives.

#### 5. REFERENCES

- Abbasi, S., Moeini, M., Shahriari, M., Ebrahimi, M., & Khoozani, E. K. (2018). Designing and manufacturing of educational multimedia software for preventing coronary artery disease and its effects on modifying the risk factors in patients with coronary artery disease. *Electronic Journal of General Medicine*, 15(3), 1–6. <https://doi.org/10.29333/ejgm/85942>.
- Abidin, A. Z., Istiyono, E., Fadilah, N., & Dwandaru, W. S. B. (2019). A computerized adaptive test for measuring the physics critical thinking skills in high school students. *International Journal of Evaluation and Research in Education (IJERE)*, 8(3), 376. <https://doi.org/10.11591/ijere.v8i3.19642>.
- Asriadi, M., & Istiyono, E. (2020). Exploration of Creative Thinking Skills of Students in Physics Learning. *Journal of Educational Science and Technology (EST)*, 6(2), 151–158. <https://doi.org/10.26858/est.v6i2.12737>.
- Aybek, E. C., & Demirtasli, R. N. (2017). Computerized Adaptive Test (CAT) Applications and Item Response Theory Models for Polytomous Items. *International Journal Of Research In Education And Science*, 3(2), 475–475. <https://doi.org/10.21890/ijres.327907>.
- Batubara, H. H. (2017). pengembangan Situs E-Learning Dengan Moodle Versi 3.1 Sebagai Media Pembelajaran Pada Program Studi Pendidikan Guru Madrasah Ibtidaiyah. *AL-BIDAYAH: Jurnal Pendidikan Dasar Islam*, 9(1), 1–10. <https://doi.org/10.14421/al-bidayah.v9i1.116>.
- Biney, I. K. (2020). Experiences of adult learners on using the Sakai Learning Management System for learning in Ghana. *Journal of Adult and Continuing Education*, 26(2), 262–282. <https://doi.org/10.1177/1477971419864372>.
- Boonprasom, C., & Sintanakul, K. (2020). The Development of Collaborative Learning Management System Using Problem-Based on Cloud Learning to Enhance Critical Thinking. *2020 7th International Conference on Technical Education (ICTechEd7)*, 13–18. <https://doi.org/10.1109/ICTechEd749582.2020.9101249>.
- Bradley, V. M. (2020). Learning Management System (LMS) Use with Online Instruction. *International Journal of Technology in Education*, 4(1), 68–92. <https://doi.org/10.46328/ijte.36>.
- Chaw, L. Y., & Tang, C. M. (2018). What Makes Learning Management Systems Effective for Learning? *Journal of Educational Technology Systems*, 47(2), 152–169. <https://doi.org/10.1177/0047239518795828>.
- Codish, D., Rabin, E., & Ravid, G. (2019). User behavior pattern detection in unstructured processes – a learning management system case study. *Interactive Learning Environments*, 27(5–6), 699–725. <https://doi.org/10.1080/10494820.2019.1610456>.
- Copriady, J., Iswandari, S. N., Noer, A. M., & Albeta, S. W. (2020). Pengembangan E-Modul Berbasis Moodle Pada Materi Hidrokarbon. *Edusains*, 12(1), 81–88. <https://doi.org/10.15408/es.v12i1.11503>.
- Devi, M., Annamalai, M. A. R., & Veeramuthu, S. P. (2020). Literature Education and Industrial Revolution 4.0. *Universal Journal of Educational Research*, 8(3), 1027–1036. <https://doi.org/10.13189/ujer.2020.080337>.
- Dewanti, S. S., Ayriza, Y., & Setiawati, F. A. (2020). The Application of Item Response Theory for Development of a Students' Attitude Scale Toward Mathematics. *The New Educational Review*, 60(2), 100–122. <https://doi.org/10.15804/tner.2020.60.2.09>.
- Dewanti, S. S., Hadi, S., Nu'man, N., & Ibrahim. (2021). The Application of Item Response Theory in Analysis

- of Characteristics of Mathematical Literacy Test Items. *İlköğretim Online*, 20(1), 1226–1237. <https://doi.org/10.17051/ilkonline.2021.01.119>.
- Fajrianthi, F., Hendriani, W., & Septarini, B. G. (2016). Pengembangan Tes Berpikir Kritis Dengan Pendekatan Item Response Theory. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(1), 45–55. <https://doi.org/10.21831/pep.v20i1.6304>.
- Fearnley, M. R., & Amora, J. T. (2020). Learning Management System Adoption in Higher Education Using the Extended Technology Acceptance Model. *IAFOR Journal of Education*, 8(2), 89–106. <https://doi.org/10.22492/ije.8.2.05>.
- Febaliza, A., & Okatariani, O. (2020). The Development of Online Learning Media by Using Moodle for General Chemistry Subject. *Journal of Educational Science and Technology (EST)*, 6(1), 40. <https://doi.org/10.26858/est.v6i1.12339>.
- Gustiani, S. (2019). Research And Development (R&D) Method As A Model Design In Educational Research And Its Alternatives. *Holistics*, 11(2). <https://jurnal.polsri.ac.id/index.php/holistic/article/view/1849>.
- Hadi, S., Retnawati, H., Munadi, S., Apino, E., & Wulandari, N. F. (2018). The Difficulties of High School Students in Solving Higher-Order Thinking Skills Problems. *Problems of Education in the 21st Century*, 76(4), 520–532. <https://doi.org/10.33225/pec/18.76.520>.
- Hardyanto, R. H., & Surjono, H. D. (2016). Pengembangan dan Implementasi E-Learning Menggunakan Moodle dan Vicon Untuk Pelajaran Pemrograman Web di SMK. *Jurnal Pendidikan Vokasi*, 6(1), 43. <https://doi.org/10.21831/jpv.v6i1.6675>.
- Haryanto, H. (2013). Pengembangan Computerized Adaptive Testing (Cat) dengan Algoritma Logika Fuzzy. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 15(1), 47–70. <https://doi.org/10.21831/pep.v15i1.1087>.
- Hu, X., Ng, J., Tsang, K. K. Y., & Chu, S. K. W. (2020). Integrating Mobile Learning to Learning Management System in Community College. *Community College Journal of Research and Practice*, 44(10), 722–737. <https://doi.org/10.1080/10668926.2019.1640146>.
- Hussin, A. A. (2018). Education 4.0 Made Simple: Ideas For Teaching. *International Journal of Education and Literacy Studies*, 6(3), 92–98. <https://doi.org/10.7575/aiac.ijels.v.6n.3p.92>.
- Istiyono, E., Dwandaru, W. B., & Rahayu, F. (2018). The Developing Of Creative Thinking Skills Test Based On Modern Test Theory In Physics Of Senior High Schools. *Jurnal Cakrawala Pendidikan*, 37(2). <https://doi.org/10.21831/cp.v37i2.19233>.
- Jana, S., & Khatun, S. (2021). Towards Selecting Effective Open Source Learning Management System/Software (OSLMS) for Higher Education Domain. *Higher Education for the Future*, 8(1), 47–70. <https://doi.org/10.1177/2347631120970083>.
- Kezer, F. (2021). The Effect of Item Pools of Different Strengths on the Test Results of Computerized-Adaptive Testing. *International Journal of Assessment Tools in Education*, 8(1), 145–155. <https://doi.org/10.21449/ijate.735155>.
- Lase, D. (2019). Pendidikan di Era Revolusi Industri 4.0. *Sundermann: Jurnal Ilmiah Teologi, Pendidikan, Sains, Humaniora Dan Kebudayaan*, 12(4), 28–43. <https://doi.org/10.36588/sundermann.v1i1.18>.
- Liao, Y., Loures, E. R., Deschamps, F., Brezinski, G., & Venâncio, A. (2018). The impact of the fourth industrial revolution: A cross-country/region comparison. *Production*, 28(1), 1–18. <https://doi.org/10.1590/0103-6513.20180061>.
- Liu, Q., & Geertshuis, S. (2021). Professional identity and the adoption of learning management systems. *Studies in Higher Education*, 46(3), 624–637. <https://doi.org/10.1080/03075079.2019.1647413>.
- Ngafeeson, M. N., & Gautam, Y. (2021). Learning Management System Adoption: A Theory of Planned Behavior Approach. *International Journal of Web-Based Learning and Teaching Technologies*, 16(1), 27–42. <https://doi.org/10.4018/IJWLTT.2021010104>.
- Noreen, S. (2020). Implementation of Learning Management System: A Way Ahead on the Digital Journey in Distance Learning. *Open Praxis*, 12(3), 329–342. <https://doi.org/10.5944/openpraxis.12.3.1086>.
- Putra, E. A., Sudiana, R., & Pamungkas, A. S. (2020). Pengembangan Smartphone Learning Management System (S-LMS) Sebagai Media Pembelajaran Matematika di SMA. *Kreano, Jurnal Matematika Kreatif-Inovatif*, 11(1), 36–45. <https://doi.org/10.15294/kreano.v11i1.21014>.
- Retnawati, H. (2017). The Dif Identification in Constructed Response Items Using Partial Credit Model. *International Journal of Assessment Tools in Education*, 5(1), 73–89. <https://doi.org/10.21449/ijate.347956>.
- Rizal, R., Rusdiana, D., Setiawan, W., & Siahaan, P. (2020). Students Perception of Learning Management System Supported Smartphone: Satisfaction Analysis in Online Physics Learning. *Jurnal Pendidikan IPA Indonesia*, 9(4), 600–610. <https://doi.org/10.15294/jpii.v9i4.25363>.
- Rudibyani, R. B., Perdana, R., & Elisanti, E. (2020). Development of Problem-Solving-Based Knowledge

- Assessment Instrument in Electrochemistry. *International Journal of Instruction*, 13(4), 957–974. <https://doi.org/10.29333/iji.2020.13458a>.
- Sahin, A., Hurtado Grooscors, H. A., & Góngora-Cortés, J. J. (2018). Review of Fast Test: A Platform for Adaptive Testing. *Measurement: Interdisciplinary Research and Perspectives*, 16(4), 256–263. <https://doi.org/10.1080/15366367.2018.1492867>.
- Sahin, D. I., & Shelley, D. M. (2020). *Educational Practices during the COVID-19 Viral Outbreak: International Perspectives*. USA: ISTES Organization Monument.
- Sahin, M. D., & Gelbal, P. D. S. (2020). Development of a Multidimensional Computerized Adaptive Test based on the Bifactor Model. *International Journal of Assessment Tools in Education*, 7(3), 323–342. <https://doi.org/10.21449/ijate.707199>.
- Van der Linden, W. J. (2018). *Handbook of Item Response Theory*. Boca Raton London: CRC Press Taylor & Francis Group.
- Wise, S. L., Kingsbury, G. G., & Webb, N. L. (2015). Evaluating Content Alignment in Computerized Adaptive Testing. *Educational Measurement: Issues and Practice*, 34(4), 41–48. <https://doi.org/10.1111/emip.12094>.
- Yunis, R., & Telaumbanua, K. (2017). pengembangan E-Learning Berbasiskan LMS untuk Sekolah, Studi Kasus SMA/SMK di Sumatera Utara. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, 6(1), 33–36. <https://doi.org/10.22146/jnteti.v6i1.291>.