Development of computer adaptive testing on the subject of Digital and Communication Simulation using item theory responses

D Y Prasetyo1* and M Khairudin2

1School of Graduates Studies, Universitas Negeri Yogyakarta, Indonesia
2Department of Electrical Engineering Education, Universitas Negeri Yogyakarta, Indonesia
*dianyugo10@gmail.com

Abstract. This article describes the development of an adaptive assessment model. This adaptive assessment model aims to determine the level of students’ abilities based on the difficulty level of the items. This type of research uses Research and Development (R & D). It uses the Borg& Gall model that explains, educational research and development is a process used to develop and validate educational product. This development aims to determine the characteristics of the items used as the basis for making computer adaptive testing products by using item theory responses to determine the level of difficulty of the questions, the distinguishing power of each item, and false guesses.

1. Introduction

Technological developments in the world of education increase rapidly, especially in the education of Vocational High Schools (SMK) which are required to improve the competence of teachers and students through the development of educational media, technology, and facilities that support the learning needs in vocational education [1]. The need for the right educational media to support teacher’s performance in the student learning process is an accurate and precise the assessment processes and suitable the student’s abilities level. Learning media are sources of communication, information, and evaluation in in the form of diagrams, printed material, computer programs, videos, internet and instructors [2], [3]. Learning media is everything related to software and hardware that can be used to convey the contents of teaching material from learning resources to group / individual students, which stimulates thoughts, feelings, attention, and interest in learning [4], [5].

Media is a physical mediator which is planned and designed to convey information and build interaction [6]. The effectiveness of learning media in the learning process can affect student achievement results [7], agree with the study. Learning media is needed to help students in the effective learning process [8]. The Physical equipment includes original objects, printed materials, visuals, audio, audio visual, multimedia and web. The equipments can be used to convey information that contains learning a messages, so that students can construct knowledge effectively and efficiently. The lack of the accurate assessment applications used to determine the level of students’ abilities causes them lack in mastering the material taught to students. The student assessment model on Digital and Communication Simulation subjects still uses classical theory in determining the
difficulty level of the items, different powers, false predictions where there are several weaknesses in the assessment, those are: the difficulty level of the item depends on the sample used in the analysis. The students’ average ability level, range, and distribution as sampled in the analysis greatly influenced the value of statistics obtained [9]. The difficulty level of the question will be high if the sample used has higher ability than the average ability of students in their population. The question’s discrimination power will be high if the ability level of the sample varies or has a large range of abilities. Reliability of the second weakness test is that student’s scores obtained from a test are very limited to the tests used [10]. Conclusion test results cannot be generalized beyond the tests used. A person’s score depend very much on the selection of the test used, not on the ability of the test participant. Limited use of test scores, classical test theory does not have a basis for studying the development of student abilities [11], [12]. The items/questions made on digital and communication simulation are from the materials found in syllabus and Lesson Plan, but all of them have not used the Bloom taxonomic Model. Based on the observations on the questions made for evaluation, that have been made using adaptive tests, none have ever been done, especially at the vocational school level for the subject of Digital and Communication Simulation.

2. Research methods
This type of research uses Research and Development (R & D). The development of test models using Computerized Adaptive Testing (CAT) is a new thing considering that it will be based on the ability of each test participant. The development model (R & D) uses the Borg & Gall model that explains, educational research and development is a process used to develop and validate educational product, the development of a good education system can help the process of implementing better quality learning as expected [13]. The the Borg & Gall model process divides the stages into ten kinds, namely, (1) Research & Data Collection, (2) Planning, (3) Early Product Development, (4) Initial Product Trial, (5) Early Product Improvement, (6) Broader Field Trials, (7) Product Improvement on Field Extensive Test Results, (8) Final Product Trials, (9) Final Product Revisions / Improvements (10) Dissemination and Implications.

The stages of the Borg & Gall model development research are the basis of the procedures used then they were applied in the planning process in creating a CAT development research scheme. The software development process using pressman model/incremental model is the development of a waterfall model designed to build a software with stages adapted to the concept of the waterfall model. The stages of designing the waterfall model can be seen in Figure 1.

![Figure 1. Pressman waterfall development model](image)

3. Results and Discussion
The design in the preparation of the CAT examination system refers to the steps as shown in Figure 2 below.
The development of this research begins with a preliminary study by analyzing the initial needs which are used as the basis for development research, then initial analysis is carried out at the institutions that will be the object of research by matching appropriate developments that can address the students' more accurate assessment problems. With an accurate assessment, the learning objectives can be achieved well. The achievement of learning objectives begins with a good learning process with the selection of material according to the agreed syllabus, making items with the basis of bloom taxonomy with an analysis of questions using the theory of item respond theory used for the purposes of calibrating items, so they get the characteristics of the items, which are used as a basis for determining the suitability between the students who work so that the questions will adjust the abilities of the students have.

Item characteristics that have been known are inputted into the computer adaptive testing application by grouping the level of difficulty of the students' grains automatically. Application testing is done to check the server computer device and client computer intended to check the application network whether it is connected to the server by checking the results can be used as evaluation of deficiencies seen in terms of software or hardware on the application of adaptive tests. With the evaluation and improvements, the weakness of adaptive test applications can be overcome.

3.1. Analysis

The stage of needs analysis in software is needed, so the software needed can be in accordance with the plan. Data collection can be done using techniques of interview, observation, questionnaire and formula planning in accordance with the development of computer adaptive testing, using of algorithms, calibration of items with IRT measurement models in order to find out the characteristics of each item, carrying out the software simulation in order to be able to describe the software which is going to be developed.

3.2. Design

Application design is a stage that focuses on the completeness of the application itself, among others, in terms of material, appearance, theme, image, sound, initial start, stop in rule. In the stage of software design, it is made as easy as possible so that in the process of an adaptive test there is no
problem in terms of software. Software is designed as flexible as possible so that it can be used in various places and inadequate conditions.

3.3. Code
The code stage is the implementation stage of the development that is built in the process of making software & hardware that is used as a tool to facilitate the application of adaptive tests by performing software testing, program evaluation, transferring from simulation programs into the application. The process of coding can be interpreted as the process of applying design designs that are made and transformed into a form of programming language that can be understood by computers.

3.4. Test
In the test process, an application needs to be used in order to see the compatibility of the software to be used so as to minimize program errors. The software test can be applied to respondents so that there is input from various sources such as from media experts, material experts, programmers, and the respondents, it is expected to be input in improving the software.

The type of testing instruments for the subject of subject Digital and Communication simulation used by the experts in the assessment are several questionnaires, aspects assessed is material, language construction. Types of CBT product testing instruments which are used to collect CBT application trials consist of two types, namely: questionnaire and software on CBT applications that are carried out online through network access. The CBT application database is tested and it will record each transaction of the CBT testing process, including the identity of the examinee, the items presented, the response of the answers, and the scores of the test participants. Types of CAT product testing instruments similar to the instruments used to collect CBT product trial data, trials for CAT products consist of two types, namely: questionnaires and CAT software or programs that are conducted online through network access.

The CAT program database that was tested will record every CAT testing process transaction, including the identity of the examinee, the items tested, the response of the answers, esthetic theta, IFF, TIF, SEM, and scores of the test participants.Instruments on the CAT product trial were used for CAT readability testing, teacher assessment sheets, and teacher assessment sheets and student assessment sheets. To obtain student responses to the CAT program that was developed, questionnaire was provided, which was distributed after students conducted trials by working on simulation tests and digital communication with the CAT program developed. The aspects of the CAT program evaluation were developed by three aspects, namely: (1) familiarity with computers (2) general assessment of the CAT program, and (3) assessment of the CAT program used. To obtain teacher responses to the CAT program developed, the questionnaire was provided, distributed after the teacher explained about the CAT program developed. Teachers are welcome to conduct trials by working on digital simulation and communication tests with the development of the CAT program. The evaluation aspect of the CAT program developed to see the teacher's response to the CAT program included four aspects: (1) aspects of performance (2) aspects of display performance, (3) aspects of benefit, (4) aspects of relevance of test material. The display of the development of the CAT rating system is as shown in Figure 3. The initial view contains the main view that there are menus that make it easier for users in the learning process and manage all the calories online.
In assessing using this CAT system, the scores and test results can be immediately recognized by test participants and test management managers as shown in Figure 4. In figure 4, it can be seen that the student database is entered automatically so that the ability of students can be recapitulated quickly and automatically visible.

Figure 5 shows class management that can be accessed by students and teachers, managing the class can be used to regulate the assignment of assignments to students, assessment, tests that can be accessed on android or computer devices.
4. Conclusion
This study has described the development of an adaptive assessment model. It has presented the characteristics of the items used as the basis for making computer adaptive testing products using item theory responses to determine the level of difficulty of the questions, the discrimination of each item, and false guesses. The CAT rating system that has been prepared has been able to display the results of the assessment shortly after the test participants completed the test according to the ability of the test participants. The long term test time for each participant is determined by how many items can be completed by the test participants. Each test participant gets test services according to their respective talents and abilities.

5. References
[1] Haryanto & Khairudin 2012 Jurnal Pendidikan Teknologi dan Kejuruan 21
[2] Robert Heinich 2002 Instructional media and technologies for learning (Upper Saddle River, N.J. : Merrill)
[3] M Khairudin, A K Triatmaja, W J Istanto, M N A Azman 2019 International Journal of Interactive Mobile Technologies 13
[4] Nizwardi Jalinus Ambiyar 2016 Media pembelajaran dan sumber belajar (Jakarta: Kencana)
[5] A K Triatmaja & M Khairudin 2018 Journal of Physics: Conference Series 1140
[6] Muhammad Yaumi, Sitti Fatimah Sangkala Sirate, & Andi Anto Patak 2018 Investigating multiple intelligence-based instructions approach on performance improvement of indonesian elementary madrasah teachers (SAGE Open)
[7] Dini Aria Farindhni 2018 Pengembangan media video animasi untuk meningkatkan motivasi belajar dan karakter siswa
[8] Hiskia Kamang Manggopa, Christine Takarina Meitty Manoppo, Peggy Veronica Togas, Alfrina Mewengkang, & Johan Reimon Batmetan 2019 Web-based learning media using hypertext markup language as course materials
[9] M J Allen & W M Yen 1979 Introduction to measurement theory (Monterey, CA: Brooks/Cole Publishing Company)
[10] Walter Dick, Lou Carey, and James O Carey 2005 The systematic design of instruction (Boston: Pearson)
[11] Rini Agustiningsih, Dwi Rahdiyanta 2019 Jurnal Pendidikan Teknologi dan Kejuruan 25, p 103
[12] R K Hambleton & H Swaminathan 1985 Item response theory (Boston MA:Kluwer Inc.)
[13] W R Borg, & M D Gall 1989 Educational research: An introduction 5th Edition. (New York: Longman)