

Utilization of Information Technology in The Development of Cognitive Diagnostic Test Instruments

1st Yuli Prihatni¹, 2nd Supriyoko², 3rd Ayu Fitri Amalia³
{yuli_prihatni@ustjogja.ac.id¹, supriyoko@ustjogja.ac.id², ayufitriamalia@ustjogja.ac.id³}

Physics Education Department, Universitas Sarjanawiyata Tamansiswa^{1,3}, Postgraduate Directorate,
Universitas Sarjanawiyata Tamansiswa²

Abstract. The purposes of this study were to develop diagnostic test instruments by utilizing information technology, find the characteristics of diagnostic tests developed and describe cognitive diagnostic information in science subjects using developed instruments. The methods used in this study were the method of developing diagnostic tests through three stages, namely Test planning stage, Testing phase and Development phase with the use of Information technology. The results showed that the instrument was developed with Computer-Based Testing Software and interpretation of test results using a database type of MySQL. The developed instruments about 28 items, with the Quest program obtained 27 items that were fit in the vertical line or had an MNSQ infit between 0.77 to 1.30. The instrument has an average value of MNSQ infit 1.01 with a standard deviation of 0.09. The instrument reliability value is 0.74.

Keywords: Technology Information, Instrument, Diagnostic Test.

1 Introduction

Information Technology plays an important role in the field of education including in the assessment system. Assessment is not just a collection of student data, but also it is processing to obtain an overview of student learning processes and outcomes. Thus the assessment will be carried out effectively if it is an adequate information system supported. Afaq & Al-Mashari's research presented a computer-based diagnostic tool developed to facilitate the learning process. Assessment is not just giving students questions and then completing, but the teacher must follow them up for the sake of learning[1]. Dunn et al in Prihatni said that the purpose and role of assessment (assessment) in learning are to[2]: 1) diagnose student learning difficulties, 2) measure improvement over time, 3) determine students' mastery of certain knowledge and skills, 4) determine the ranking of students from all students in the class, 5) evaluate learning methods, 6) evaluate the effectiveness of learning programs. In the classroom learning, the assessment done by the teacher aims to find out the progress and learning outcomes of students, compare the ability of students, diagnose learning difficulties, provide feedback/improvement of teaching and learning processes, and determine class increase. Based on Leung in Afitska, the assessment process is an important part of classroom practice and involves both teachers and students in reflection, dialogue and decision making[3].

According to Nitko in Babu, the practice diagnostic assessments in the classroom have two areas, namely 1) to identify learning targets that have not been mastered by students and 2) to

find causes or reasons that have not yet been able to master the targets learning[4]. The assessment that will be carried out in this study uses a test instrument. In particular, DiBello in Javidanmehr stated that the validity of cognitive diagnostic assessments, based on the results of complex and multidimensional substantive students, demands psychometric models that reflect substantive complexity to an adequate level[5]. To assess how far students have mastered a variety of competencies, of course various types of assessments need to be given in accordance with the competencies to be assessed, such performance, assignment (product), work product (product), collection of student work (portfolio) , and written assessment (paper and pencil test). So, the purpose of the assessment is to provide comprehensive information suggestion on the learning outcomes of students, both seen when learning activities take place and from the end of the result, using various assessment methods in accordance with the competencies expected by students.

Therefore, the assessment is a series of activities to obtain data, analyze, and interpret data about student learning processes and results that are carried out systematically and continuously so that it becomes meaningful information in decision making. The assessment in this study was carried out using an instrument in the form of a diagnostic test. In this study assessment was devoted as a tool for cognitive diagnosis of students in junior high school science subjects, the assessment was used to show the location of learning weaknesses and difficulties experienced by students and possible achievements that can be developed. This will help the teacher to provide learning that is appropriate to students' thinking processes.

There are two approaches to cognitive diagnostic assessment, namely the Mislevy evidence-centered model approach and the Embretson cognitive design system approach. The approaching of the central incident model of the Mislevy component is defined as three models namely student model, incident model and task model. Student models characterize weaknesses and students. The incident model describes the behaviour or appearance that determines the occurrence of the student model. Gorin in Kusaeri said that the task model defines tasks specifically covering conditions, materials and work of the Embretson design system approach focusing on the role of cognitive theory and test development[6]. Learning and diagnostic-based learning involve the assessment as a testing process to gather feedback information and as a testing process to gather input information about the difficulties experienced by learners in learning core concepts[7].

Actually, the diagnostic tests can help a teacher's assessment in extending their students' conceptual understanding of a particular topic[8]. The purposes of this study were (1) to develop diagnostic test instruments by utilizing information technology, (2) find the characteristics of diagnostic tests developed and (3) describe cognitive diagnostic information in science subjects using developed instruments.

There are six approaches to diagnostic assessment according to Nitko & Brookhart in Fauziah[9]. Approaches related to learning problems are: (1) Profile of students' weaknesses and strengths on a particular material, (2) Approach to identifying lack of prerequisite knowledge, (3) Approach to achieving student competencies, (4) Approach to identifying student performance errors, (5) Approach knowledge structure of students, and (6) Approach to students' competence in solving difficult problems.

2 Method

This study used the diagnostic tests of developing research methods. Five steps of developing diagnostic tests to cognitive assessment based on Nichols are as follows[10]:

- (1) Assessing the construction of a substantive theory as fundamental in the development of the test based on research or research review.
- (2) Selecting a measurement design to construct items responded by test participants based on knowledge, specific skills or other characteristics according to the theory.
- (3) Administering tests which cover several aspects including composing a test form, the technology of making tests, analyzing environmental situations at the time of testing, making test grids, writing tests, etc.
- (4) Make a test score by determining the score of the test carried out through the analysis of the item.
- (5) Conduct revisions the process of adjustment and improvement in accordance with the purpose of the diagnostic test and expected standards.

Five steps for developing cognitive diagnostic tests above will be developed in accordance with the focus of the research through three stages, namely (1). The planning stage of the test, (2) the testing stage and (3) the stage of developing the test by utilizing Information Technology, namely the development of computer-based testing software prototype. So that, this research can be directly applied in schools. General development research, the process includes: build a conceptual framework, followed by designing the system and end with build the prototypes for testing and evaluation [11].

Tests developed with information systems can provide interpretation or interpretation of test results aimed at translating and giving meaning to the scores obtained by the testee (the person being tested). Thus these results can be used as a basis for providing follow-up in the form of appropriate treatment and in accordance with the weaknesses of students. Cognitive diagnostic tests in this study were tested to determine the weaknesses and difficulties of students.

Diagnostic tests are combined with computer software that has been developed using a database type of MySQL that is used to store problem data and other data that is equipped with a password. Every teacher who will apply cognitive diagnostic tests to his students can register with the admin, and after that will be given a password. Students who already have a password can carry out cognitive diagnostic tests on school computers or personal laptops connected to the internet.

The subjects of the tests on the development research of cognitive diagnostic instruments in Science subjects for Junior High School were Junior High School students in Sleman Regency, Yogyakarta Special Region. The tested instrument to determine the feasibility of the cognitive diagnostic test instrument developed. The tests have been carried out in Junior High Schools which have different characteristics in Sleman Regency. Research permits were carried out on principals in 7 schools in Sleman Regency with a total of 484 students.

Data collection by cluster random sampling is sampling based on certain groups. Clusters in this study were various schools that have a certain level, namely schools with the average results of national examination scores are in the category of good, medium and less. Then each school was taken randomly. Because the sample was chosen randomly in a multilevel cluster, this technique is called multi-stage sampling according to Karlton in Nafiu[12].

The type of data collected from the initial study is quantitative and qualitative data, while the data collected from the test instrument is in the form of quantitative data obtained from scoring for each question item. This empirical data was the primary data obtained from primary sources derived from the responses of students, teachers and expert judgment.

Analysis of test instrument development data used product content validation by experts through the FGD process to find out how far the cognitive diagnostic test manuscripts in science subjects for junior high school reflect the development objectives in the first stage. At this stage, the arranged cognitive diagnostic test instruments were discussed in the FGD forum attended by 18 FGD participants consisting of measurement experts, cognitive psychology, Physics Science, peers, and junior high school teachers who were members of the Sleman Regency MGMP. Feedback was delivered orally and through writing. FGD activities were held within 3-4 hours. The next step was to make improvements or additions to cognitive diagnostic test instruments in science subjects for Junior High School based on input and suggestions from FGD participants. The next step was to consult the results of the revision of the cognitive diagnostic test instrument to the supervisor and validated by 4 experts before the instrument was tested in the field.

Quantitative data analysis with the help of the Quest program, CDM and M Plus Program. The Quest program is used to obtain the parameters of the item, namely the level of difficulty level item to complete cognitive diagnostic information on the compiled instrument, namely by linking the learning continuum stage and the cognitive dimension. Analysis with IRT used the Quest program can compare the unknown abilities and difficulties when analyzed with classical analysis. Thus, the results of the analysis of the Quest program can help provide diagnostic information about students' weaknesses and difficulties. The Quest program can also be used to determine fit items with models. Thus, the Quest program can create standardized items. The limit of an item is declared fit with the model if it has an MNSQ infit between 0.77 and 1.30 according to Adam & Khoo in Istiyono[13]. Fit statistics requirements in the Quest program, namely if the average value of the MNSQ infit approaches 1.0 with a standard deviation of 0.00.

3 Results and Discussions

This study produced a diagnostic test instrument that aimed to determine the weaknesses and excellence of Junior High School students in physics subjects especially heat material. So that, the results can be used as a basis for providing follow-up in the form of appropriate treatment and in accordance with the weaknesses of students. Diagnostic test instruments are equipped with a lattice of cognitive diagnostic tests based on the learning continuum that contains cognitive processes, dimensions of knowledge and the material to be tested.

The form of the test used in this study was an objective test form accompanied by reasons. The advantages of objective tests are as follows: a) the amount of asked material is relatively broader than the covered material by other forms of matter. The number of questions asked is generally relatively large, b) can measure various levels, c) the scoring is easy, fast, objective, and can cover a wide range of materials and materials in one test for a class or level. d) it is very appropriate to use if there are many participants e) the reliability of multiple choice questions is relatively higher than the description problem. Objective tests also have weaknesses, one of which is that it can only measure up to the analysis stage. Therefore in this study the cognitive process that was measured arrived at the analysis stage.

The test form used in this study was an objective test form accompanied by reasons. There was a reason for each question to find out how students work, whether it is in accordance with their knowledge, whether it is the correct steps and reasons. Thus, when the students can determine the choice of answers, the students must know what is the reason. The use of objective questions equipped with reasons is generally not widely used in schools. Therefore, the students are given an explanation firstly by the teacher before working on the questions. Students who can choose answers were they know their thoughts. The developed instruments were validated by experts through FGD activities with the following results.

Table 1. Analysis of FGD results

Contain	Poor	Fairly Good	Good	Very Good
Learning <i>continuum</i>	-	-	25%	75%
Hierarchy of Prerequisite Materials	-	-	75%	25%
Test grid	-	-	50%	50%
Test specifications	-	-	-	100%
Alternative answers	-	-	25%	75%
Diagnostic information	-	-	50%	50%
Problem form	-	-	50%	50%
Answer sheet and answer key	-	-	-	100%

The data that has been collected in the validation empirically use Rasch models through the Quest program. Grain selection with the Quest program was used to obtain the level of difficulty of the item and fit the test. Analysis with the Rasch Model was used to measure one theory construct at a time in the logit scale hierarchy. In this study, the number of items arranged was 28 items. The trial data analyzed by Rasch Model used the dichotomous Quest program are presented in Table 2.

Table 2. Fit items with the Quest program

Item Fit					5/11/14 14:14				
all on all (N = 484 L = 28 Probability Level= .50)									

INFIT									
MNSQ	.56	.63	.71	.83	1.00	1.20	1.40	1.60	1.80
-----+-----+-----+-----+-----+-----+-----									
1 item 1			.	*	.				
2 item 2			.	*	.				
3 item 3			.	*	.				
4 item 4			.	*	.				
5 item 5			.	*	.				
6 item 6			.	*	.				
7 item 7			.		.*				

8 item 8	.	*		.
9 item 9	.	*		.
10 item 10	.		*	.
11 item 11	.	*		.
12 item 12	.	*		.
13 item 13	.		*	.
14 item 14	.		*	.
15 item 15	.	*		.
16 item 16	.	*		.
17 item 17	.	*		.
18 item 18	.	*		.
19 item 19	.	*		.
20 item 20	.	*		.
21 item 21	.	*		.
22 item 22	.			* .
23 item 23	.			* .
24 item 24	.	*		.
25 item 25	.	*		.
26 item 26	.	*		.
27 item 27	.			* .
28 item 28	.			* .

=====

The limit of an item is declared fit with the model if it has an MNSQ infit between 0.77 to 1.30 (Adam & Khoo, 1996; 30). Fit statistics requirements in the Quest program, namely if the average value of the MNSQ infit approaches 1.0 with a standard deviation of 0.00. The instrument for cognitive diagnostic tests in science subjects has a total of 28 items analyzed. From the analysis results, it obtained that 27 items fit. In this study, 27 items were in a vertical line or have an MNSQ infit between 0.77 and 1.30. 7th item was on infit MNSQ 1.31. But because item number 7 is not far from the vertical line, and has a difference of infit MNSQ 0.01, in this study 7th point was still used with revision. This research instrument has met the Fit statistics requirements in the Quest program, namely with an average value of MNSQ infit 1.00 with a standard deviation of 0.11. The reliability of the samples in this study amounted to 0.99. The reliability score based on item estimation is called sample reliability[14]. Therefore, the higher the value, the more items fit the model.

While the test reliability in this study obtained at 0.72 by looking at the score of the case estimate of reliability. Test reliability is the case estimate of reliability[14]. Test reliability of 0.72 provided information that the measurements of cognitive diagnostic test instruments in Science lessons provided consistent results. Gronlund in Prihatni stated that the magnitude of the reliability index of a measurement instrument, especially in the educational measurement such as the teacher-made test of about 0.6-0.85[2]. The higher the test reliability score, the more trials samples provided the expected information and vice versa. The lower the test reliability score, the fewer trials samples provided the expected information.

The results of cognitive diagnostic tests tested in seven Junior High Schools in Sleman Regency for heat material presented as shown in Figure 1.

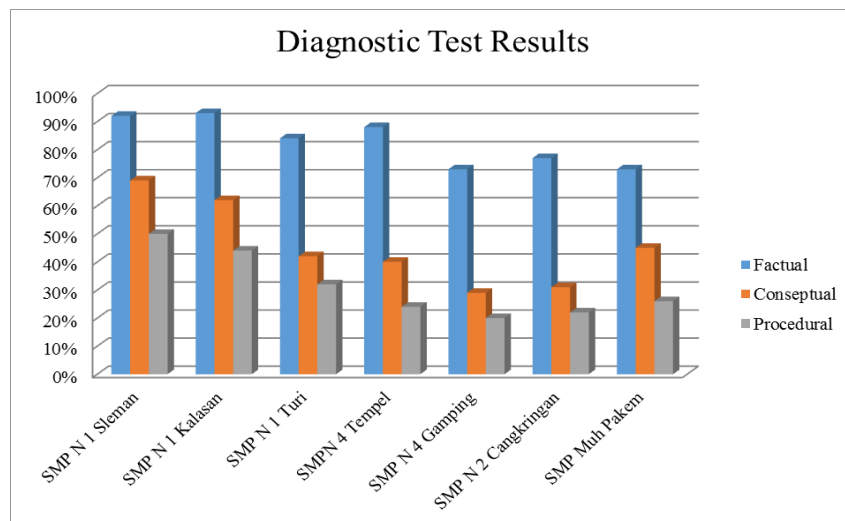
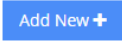


Fig. 1. Comparison of productive questioning skills for natural science teacher prospective in the 1st cycle and 2nd cycle.

The developed information system can see the diagnostic test results for each student. Development of Computer-Based Testing Prototype Software as an interpretation of Test Results used a database type of MySQL that was used to store question data and other data equipped with a password obtained by registering with the admin. To open the application of cognitive-diagnostic data, students can click the browser icon (Mozilla, opera, google chrome etc.) on the computer. Students who already have a password can carry out cognitive diagnostic tests on computers or personal laptops connected to the internet at school. How to use it, students can type the application address www.dagnostik-kognitif.com in the URL section then press enter, if the address is written correctly then the computer screen as shown in Figure 2.

Students can fill in the username in the field with the  icon and fill in the password in the field with the  icon. Then click  button to verify the username and password. If the username and password entered are correct, then the main menu appears from the diagnostic-cognitive application dashboard.

Students who carry out cognitive diagnostic tests can only access the main menu for testing, and those who can open the results are the teacher and admin. In its management, the admin is the system controller and access network in cognitive diagnostic tests. Admin can log in using the application administrator account

By selecting the menu, you will see a list of existing system users as shown in Figure 3. If the admin wants to add a user to the system, press the  button, then the form appears as shown in Figure 4. The developed cognitive diagnostic tests are equipped with time settings.

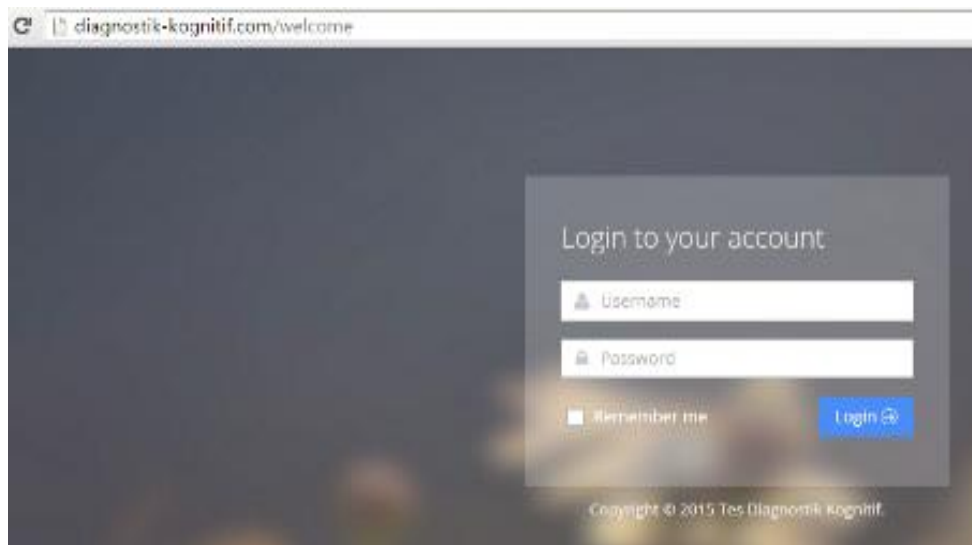


Fig. 2. Initial display login Diagnostic Test.

User & Student Management

Dashboard > User & Student Account MARCH 23, 2015 - APRIL 21, 2015

User & Student Management

Add New +

Number	Username	Email	Group	Status	
1	teacher	teacher@teacher.com	Teacher	Active	Edit Delete
2	student	student@student.com	Student	Active	Edit Delete
3	Afifa	afifa_ps@gmail.com	Student	Active	Edit Delete

Add New +

Fig. 3. System user view.

The teacher and admin can see the results of cognitive diagnostic tests that students have implemented. The results of cognitive diagnostic tests consist of two information, namely 1) the number of items answered correctly and items that students can answer based on their type of knowledge, 2) information about the material that has been mastered and not mastered by students based on the learning continuum and unknown material prerequisites by students.

Hasil Test

Nama Guru	Khasbi Khamdan, S.Pd
Asal Sekolah	SMPN 1 Sleman
Alamat Sekolah	Jl. Magelang Sleman
Nomor Induk	02
Jawaban Benar	26
Jawaban Salah	2

Nomor Soal	Learning Continuum	Jenis Pengetahuan	Proses Kognitif	Benar
1	Menjelaskan pengertian kalor	Faktual	Ingatan	✓
2	Menjelaskan konsep perpindahan kalor	Konseptual	Pemahaman	✓

Fig. 4. Display of Student Test Results based on cognitive processes.

The display of cognitive diagnostic test results contained information about students' abilities based on the learning continuum, the type of knowledge and the process of students knowledge. The teacher or admin can print the information. The information used as a source for planning follow-up and improving learning. Computer software will display more specific diagnostic information as shown in Figure 5.

Thus the results of the interpretation of this diagnostic test can be used by the teacher in knowing the weaknesses of students and can provide appropriate treatment for students' needs.

3	Menentukan konversi satuan pada besaran kalor	C	Siswa dapat menentukan konversi satuan pada besaran kalor	
4	Menjelaskan sifat-sifat termometrik suatu zat	C	Siswa dapat menjelaskan konsep kalor jenis	
5	Menghitung jumlah kalor pada zat yang diketahui kalor jenisnya	A	Siswa dapat menjelaskan satuan kalor	
6	Menghitung jumlah kalor pada zat yang diketahui kalor jenisnya	A	Siswa dapat menentukan konversi satuan pada besaran kalor	
7	Menghitung kapasitas kalor pada zat yang diketahui massa dan kalor jenisnya	D		Siswa belum dapat menjelaskan satuan kalor dalam SI
8	Menjelaskan pengaruh kalor dalam merubah suhu dan wujud zat	D	Siswa dapat menjelaskan pengaruh kalor dalam merubah suhu dan wujud zat	
9	Menyebutkan perubahan wujud zat yang terjadi dari cair ke padat	B	Siswa dapat menyebutkan perubahan wujud zat yang terjadi dari cair ke padat	
10	Menyebutkan perubahan wujud zat yang terjadi dari gas ke padat	D	Siswa dapat menyebutkan perubahan wujud zat yang terjadi dari gas ke padat	
11	Menyebutkan perubahan wujud zat	B	Siswa dapat menyebutkan	

Fig. 5. Display of Student Test Results based on learning continuum and mastery of the prerequisite material.

4 Conclusions

From this study, it can be concluded that the developed instrument with Computer-Based Testing Software and interpretation of test results used a MySQL database type. The developed instruments amounted to 28 items, with the Quest program obtained 27 fit items in the vertical line or had an MNSQ infit between 0.77 to 1.30. The instrument had an average value of MNSQ infit 1.01 with a standard deviation of 0.09. The instrument reliability score was 0.72. Cognitive diagnostic information generated by cognitive diagnostic test software in the form of (a) the number of correctly answered items and students can answer based on their type of knowledge items, (b) information about mastered and not mastered material by students based on the learning continuum and pre-requisite materials that have not controlled by students.

Acknowledgments. We are grateful to teamwork. This research is based upon work supported by DRPM DIKTI.

References

- [1] A. Ahmad, A. Al-Mashari, and A. Al-Lawati, "On the development of a computer based diagnostic assessment tool to help in teaching and learning process.," *Int. J. Educ. Dev. using Inf. Commun. Technol.*, vol. 6, no. 1, pp. 1–15, 2010.
- [2] P. Yuli, Kumaidi, and Mundilarto, "PENGEMBANGAN INSTRUMEN DIAGNOSTIK KOGNITIF PADA MATA PELAJARAN IPA DI SMP," vol. 20, no. 1, pp. 111–125, 2016.
- [3] O. Afitska, "Use of Formative Assessment , Self- and Peer-Assessment in the Classrooms : Some Insights From Recent Language Testing and Assessment (Lta) Research," *J. English Lang. Teach.*, vol. 4, no. 1, pp. 29–39, 2014.
- [4] R. Babu, "Teaching Science in Bangladesh : Expectation versus Reality," vol. 10, pp. 245–255, 2016.
- [5] Z. Javidanmehr, M. Reza, and A. Sarab, "Cognitive Diagnostic Assessment: Issues and Considerations," *Int. J. Lang. Test.*, vol. 7, no. 2, pp. 73–98, 2017.
- [6] Kusaeri and Kumaidi, "MENENTUKAN UKURAN MATRIKS Q PADA MODEL DINA UNTUK DIJADIKAN DASAR," *J. Ilmu Pendidik.*, vol. 21, no. 2005, pp. 39–44, 2015.
- [7] N. S. Æ. P. Pirnay-dummer, "Model based assessment of learning dependent change within a two semester class," *Educ. Tech Res. Dev.*, vol. 57, pp. 753–765, 2009.
- [8] F. Gulf and F. Myers, "Diagnosing students ' misconceptions in algebra : Results from an experimental pilot study," *Behav. Res. Methods*, vol. 41, no. 2, pp. 414–424, 2009.
- [9] D. Fauziah, Mardiyana, and D. R. S. Saputro, "Mathematics authentic assessment on statistics learning : the case for student mini projects Mathematics authentic assessment on statistics learning : the case for student mini projects," *J. Phys. Conf. Ser.*, vol. 983, pp. 1–5, 2018.
- [10] P. Nichols and S. Ferrara, "The Learning Diamond : A Systemic Perspective on Student Learning Research Report," 2014.
- [11] T. J. Ellis and Y. Levy, "A Guide for Novice Researchers : Design and Development Research Methods What Design and Development Research Is –," in *Proceedings of Informing Science & IT Education Conference (InSITE) 2010*, 2010, pp. 107–118.
- [12] L. A. Nafiu, I. O. Oshungade, and A. A. Adewara, "Alternative Estimation Method for a Three-Stage Cluster Sampling in Finite Population," *Am. J. Math. Stat.*, vol. 2, no. 6, pp. 199–205, 2012.
- [13] E. Istiyono, W. B. Dwandaru, and F. Rahayu, "PENGEMBANGAN TES CREATIVE THINKING SKILLS FISIKA SMA (PhysCreTHOTS) BERDASARKAN TEORI TES MODERN," *J. Cakrawala Pendidik.*, vol. Juni 2018, no. 2, pp. 190–200, 2018.
- [14] B. Wright and M. Stone, *Measurement Essential*. Wilmington, Delaware: Wide Range, Inc, 1982.