Sonar : Jurnal Multidisiplin

https://jurnal.untirta.ac.id/index.php/sonar/index

e-ISSN: XXXX-XXXX

Vol. 1, No. 1, April 2024, Page 23-36

The effectiveness of using a computer-based test (cbt) and paper-based test (pbt) in assessing students' reading skills in SMAN 3 Kota Serang

Mulki Siti Hajar Rezaini*

Universitas Sultan Ageng Tirtayasa, Indonesia

**E-mail: sitihr@gmail.com*

(Received: 15 January 2024; Accepted: 20 March 2024; Published: 30 April 2024)

ABSTRACT

This research aims to describe the process of using a computer-based or computer-based test (CBT) and paper-based system or Paper Based Test (PBT), and the effectiveness of using CBT and PBT at SMAN 3 Kota Serang. Data will be collected using several methods, namely, observation, documentation, tests, and interviews. Data Analysis Techniques, namely all data collected, are analyzed to find the answers to the problems that have been formulated. So that conclusions can be drawn from the results of the study. To find out about the Effectiveness of Computer-Based Tests or Paper-Based Tests in Assessing students' English learning outcomes at SMAN 3 Kota Serang using a statistical approach involving comparative analysis of data on student scores from CBT and PBT exams using Pearson's Coefficient Correlation. The data testing results show that the average value of the class using the CBT method is 50.83, higher than the average value of the class using the PBT method, which is 30.18. These results indicate that students' reading skills in classes that apply the CBT method are better than in classes that use the PBT method, or it can be interpreted that the CBT method is more effective in improving students' reading skills compared to the PBT method

Keywords: Computer Based Test; Paper Based Test; Effectiveness.

DOI:

INTRODUCTION

Information technology has significantly transformed the method of assessment. In many academic domains, educational measurement has been moving from Pencil paper tests (PPT) to the use of computer-based testing (CBT), defined as tests or assessments that are administered by computer in either stand-alone or dedicated network or by other technology devices linked to the internet or World Wide Web most of them using multiple choice questions (MCQs), (Sorana -Daniela and Lorentz, 2007).

Computer-based tests have been used since the 1960s to test knowledge and problem-solving skills, (Peter et al., 2004). Computer-based assessment systems have enabled educators and trainers to author, schedule, deliver, and report on surveys, quizzes, tests, and exams. There are two main types of computer-based testing. The most familiar type is where candidates fill in their responses on a paper form, which is fed into a computer optical mark reader. This reads the form, scores the paper, and may even report on the test's reliability. The second type of computer-based testing is where the computer provides an assessment interface for students ; they input their answers and receive feedback via a computer (Peter et al., 2004).

An effective student assessment technique is necessary to assess student knowledge. Due to the increase in student numbers, ever-escalating work commitments for academic staff, and the advancement of internet technology, computer-assisted assessment has been an attractive proposition for many higher education institutions (Darrell, 2003).

In most institutions, the traditional method (a combination of essay and practical examinations) is used to evaluate students ' knowledge. In the past few years, the number of students has increased drastically, and the conventional examination method has become time-consuming in terms of the examination time for evaluation and assessment. A solution for examination in large classes of students is an automated testing system, which has not yet been entirely introduced by institutions in the country, primarily to address this concern and others.

Generally, the advantages of CBT systems over traditional paper-and-pencil testing (PPT) have been demonstrated in several comparative works. As mentioned by (Peter et al., 2004), CBT is not just an alternative method for delivering examinations. It represents an essential qualitative shift from traditional methods such as paper-based tests. Despite these advantages available in computerized test administration, it was shown that it does not mean that CBTs are intrinsically better than paper - and -pencil tests (John et al ., 2002). Furthermore, while recognizing the system-level advantages associated with CBT, exploring the relationship between assessment mode and the student's behavior is essential. The term “ affordances ” describes what is made possible and facilitated and what is made complex and inhibited by a medium of assessment (Johnson and Green, 2004). It is possible that the affordances offered by computer mediated assessment May affect the perception of students involved in computer-based assessment differently than if they were engaged in paper-based assessment (Johnson and Green, 2004). Several areas appear worthy of investigation, including quality factors that May influence performance and student perception regarding computer-based tests.

PBT also has advantages even though it is more visible in the minority than it is lacking. There are in the process of working out questions that require calculations and questions that require accuracy in reading, the PBT provides more flexibility for examinees in completing it. They can scribble on paper and carefully understand the essential points in the matter. If the question is a long text , the examiner can underline or give a sign so it's easy to do it. Compared with CBT, examinees can only look at a monitor, which, when done continuously, can make the eyes hot and even cause dizziness. In addition, the most worrying thing about CBT is the risk of misunderstanding the questions for participants are high enough if the use of language is less firm and straightforward, the risk of system damage can occur and hinder the continuity of the exam is relatively high especially if there is a hacker attack, and with many participants,

if it is not balanced with bandwidth capacity, it can cause system delays and disrupt the continuity of the test.

During the implementation of the CBT, it can be said that fraud can be minimized, which highlights the advantages of this exam. This fraud can be minimized because it is done online when the material or the CBT question is doubled. Likewise, questions are distributed easily and directly through the Internet. If based on a security system, what needs to be considered is security in the computer network from the invasion of hackers or hackers. Therefore, it is necessary to have layered security in existing computer networks.

In addition, the online exam is one of the right strategies in green IT, which will reduce the use of paper worldwide so that timber-producing trees as raw material for making paper can be maintained. Imagine how many tons of paper can be reduced in one year if schools, universities, and educational institutions replace the exams that use paper with computer-based online exams. Reducing paper use will indirectly also preserve the green environment so that it can reduce energy use. Minnesota Pollution Control Agency reports that 40 reams of paper are comparable to 1.5 hectares of pine forest that can absorb carbon in a year; one ream of paper is equivalent to 12 pounds of carbon dioxide, which cannot be removed from the atmosphere. (Minnesota Pollution Control Agency, 2011). The paper industry contributes up to 10% of world emissions in the paper-making process. Energy absorbed by paper mills, too, reached 25.8 billion kWh of electricity and 54.3 billion BTU of world oil in 2010, as stated by J. DeRosa (2007).

Based on observations, SMAN 3 Kota Serang is a favorite school in Serang City, supported by adequate infrastructure. Various learning support facilities at SMAN 3 Serang City can be considered sufficient with learning support facilities installed in environment schools, including computers, laptops, LCDs, CCTV, and internet networks. Almost all schools, especially public schools , have conducted exams using computers to conduct national examinations. In developing technology in the educational environment, trends or habits that can support implementing CBT activities are needed. Why is the application of CBT required in exam activities in the school environment? Carrying out this exam is straightforward in the correction stage. It can minimize fraudulent actions, which include leakage of questions, cheating during the exam, and even changes in the test scores. The implementation of the CBT exam is beneficial for education providers.

In addition to using the CBT program at Serang 3 Public High Schools in UNBK and USBN BK, school teachers also began using CBT in the Daily Test Process (UHBK). With UHBK, teachers can more easily give students grades; teachers don't bother photocopying questions for student exams; teachers only need to listen to questions. Besides, the teacher is more accessible because it doesn't have to fix and immediately get out of value. Still, it's also easier to analyze the problem because it's already on the server. The teacher only checks on the server, values, and data, and analysis is more accessible. The purity of the test score is also higher because if every computer is online, the problem can be random and make it difficult for students to imitate their friends' answers. Teachers are more accessible to monitor because in a computer laboratory, only one sheet of paper is needed, so the opportunity to copy books (chat) is a bit because it is only possible to bring a piece of paper to the lab. Paper

Using CBT as a summative assessment tool carries concrete practical and economic benefits because it provides facilities to test many student cohorts with automated marking of responses (Charman, 1999; Zakrzewski & Bull, 1998). CBT is a mode of testing that acts as a catalyst for change and provides a base for change in the mode of learning, instruction, and curriculum in education institutions (Scheuermann & Pereira, 2008). In most educational institutions, there has been a recent trend of shifting the mode of assessment from PBT to CBT. Administering the CBT mode of assessments becomes predominantly widespread in the education assessment domain because this significant variation in assessment methodology leads to practical changes in pedagogy and curriculum methodology (Chen, 2012; Genc, 2012; Hsiao, Tu, & Chung, 2012; OECD, 2010). Pedagogical advantages of CBT include providing fast and error-free feedback; repeatability of tests consisting of randomly generated test items; unquestionable reliability and fairness; flexibility in the allocation of test timing and venue; and direct responsibility for one's learning and test-taking (Charman, 1999). There is a clear policy statement by the International Guidelines on Computer-Based Testing (International Test Commission, 2006) that in order to administer a valid and reliable CBT, corresponding test scores should be established for conventional paper-based testing (PBT) and its corresponding computer-based methods. A solid support base has been provided to this set of testing standards by the classical true score test theory — the basis of computer- and paper-based testing (Allen & Yen, 1979). As per proposal theory by Allen & Yen (1979), for anyone who takes the same test in the two modes mentioned above (CBT and PBT), it is anticipated that the test taker obtains almost the matching level of test scores. The same idea and theory have also been supported by empirical studies by the OECD (2010); Wilson, Genco, & Yager (1985). In their related study, OECD (2010) stated that no significant discrimination has been found in the mode of test performance between CBT and PBT. Their findings were based on the data collected from the student participants (n = 5,878) from Denmark, Iceland, and Korea.

The ideas related to the results of the correspondents of both PBT and CBT are also reinforced by many studies in a particular subject area, and clear outcome discrimination in achievement tests such as science, language, and mathematics, as well as the very same ascertained. By a series of psychological tests such as personality and neuropsychological assessment (for example, Friedrich & Bjornsson, 2008; Choi, Kim, & Boo, 2003; DeAngelis, 2000). In their findings on a review of educational and psychological measurement approaches, Bunderson, Inouye & Olsen (1989) have determined that 48% of previous studies revealed negligible differences between the two modes of testing (PBT & CBT) in the field of test performance, while 13% of studies have reported that the performance of a CBT test is better than PBT and 39% of the Findings prove PBT is better than CBT.

There May be a direct explanation for the specific differences mentioned above from test performance, both the proposed CBT has weak validity as an assessment tool for educational assessment and related to psychological mode, or there May be various other factors that overshadow the positive impact of CBT mode on test performance by the pattern of repeated action studies applied. In their parallel study, as established by Yu & Ohlund (2010), a possibility shadow variable is the effect of testing; according to the process, it has a pretest preceded by a posttest analytically confusing the impact of CBT treatment on performance tests.

English entered into the Academic Potential Abilities Test. This means that English is one of the fields that must be studied. Students must face the Department of MIPA and IPS English language Proficiency tests. In general, four aspects will be measured by these questions. The four aspects include listening, reading (speaking), speaking (speaking), and grammar (grammar). This is where the types of questions that appear will vary. However, in implementing daily tests, the teacher at SMAN 3 made questions to test students in various ways; sometimes, not all English language skills were tested, focused more on reading skills and listening skills alone in the CBT and PBT tests methods. Then, what about the test results and the readiness of students to use computer-based exams? Can students get better scores compared to paper-based exams, especially in their reading skills?

Based on the above assumptions, a computer-based National Examination (CBT) system problem is applied; what about the test results and the readiness of students to use computer-based exams, and whether students can get better scores compared to paper-based exams, especially in their reading skill. With the background above, based on these reasons, the researcher intends to raise the title " Effectiveness Using Computer Based Test (CBT) or Paper Based Test (PBT) in assessing students reading skills in SMAN 3 Serang City.".

RESEARCH METHODS

This study uses a quantitative approach. Quantitative research is systematic scientific research on parts and phenomena and their relationships. The quantitative research method in this study uses comparative research. Research comparative is research that compares two or more symptoms. Comparative research can be descriptive comparative (descriptive comparative) and comparative correlation (comparative). Descriptive comparative compares the same variables for the sample differently.

In this study, researchers wanted to compare the effectiveness of using computer-based or paper-based tests in assessing students ' English subjects by conducting comparative analyses. Comparative research is directed at knowing whether differences exist between two or more of the two groups in the aspects or variables studied. In this research, there is no variable control, manipulation, or treatment of the researcher. The study was carried out naturally; researchers collected data using measuring instruments. The result was statistically analyzed to look for differences between the variables studied.

Data collection will be carried out using several methods, including observation. This method is used to observe class situations when testing using computer-based test methods and paper-based test methods are ongoing; documentation: This method is used to obtain data about matters relating to this research, examples of the process of carrying out computer-based and paper-based tests at SMAN 3 Serang City, test: This test is conducted for English Language Education lessons so that effectiveness can also be known from the Computer-based test and paper-based test methods, interview: is used to find out more in-depth information from respondents regarding the efficacy of using the Computer Based Test and Paper Based Test methods on reading student skills. Respondents for this interview techniques are the head of the new UNTIRTA student entrance examination test, where this year there were changes in the method from the Paper based test to the computer based test method.

After all the data is collected, the step is to analyze the data to find the answers to the problems that have been formulated so that a conclusion can be drawn from the study results. In this research, to find out about the Effectiveness of Computer Based Test or Paper Based Test in Assessing students' reading skills in SMAN 3 Kota Serang use, a statistical approach that involves comparative analysis of data on students' scores from CBT and PBT exams will be performed using Pearson Coefficient Correlation.

Pearson moment Correlation (r) signifies the degree of relationship that exists between the dependent variable is the Paper Based Test (PBT), denoted as X, While the independent variable is the Computer Based Test (CBT), denoted as Y. Equation 1 represents the Pearson correlation coefficient formula, the valid result for r lies between -1 and +1. If r lies between 0 and 1, it shows a positive correlation: X increases as Y increases. If r = 1, it shows that the result is perfectly positive. If r is between 0 and 0.49, it shows a low positive correlation. When r = -1, it shows a perfect negative correlation; that is, the rate at which the dependent variable increases is precisely equal to the rate between -0.49 and -1, and it exhibits a strong negative correlation. Below is the Pearson:

Coefficient correlation formula:

$$r = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{N}}{\sqrt{(\Sigma x^2 - \frac{(\Sigma x)^2}{N})\,(\Sigma y^2 - \frac{(\Sigma y)^2}{N})}}$$

Where :

N = Number of pairs of students

Σxy = sum of the products of paired scores

Σx = sum of x scores

Σy = sum of y scores

Σx^2 = sum of squared x scores

Σy^2 = sum of squared y scores

X represents the students' paper-based Test, Y represents the students Computer Based Test score, and N is the number of students analyzed.

RESULTS AND DISCUSSION

A. Descriptive Statistics

The following is a description of students ' reading Skills using the CBT (Computer Based Test) and PBT (Paper Based Test) methods:

Table 1. Descriptive Statistics on Student Reading Skills

	N	Minimum	Maximum	Mean	Std.Deviation
Computer Based Test	40	64	100	85.23	10,531
Paper Based Test	40	60	100	75.25	10,157

Based on the data processing results, it is known that the reading skills score of students in the class using the CBT method has the lowest score of 64, the highest is 100, with an average score reaching 85.23. In the class using the PBT method, students ' reading skills had the lowest score of 60; the highest was 100 with an average score reaching 75.25. With a KKM (Minimum Completion Criteria), class XI students' score is 72.

The distribution of score data on the reading skills of students in class using the CBT method can be illustrated in the following diagram:

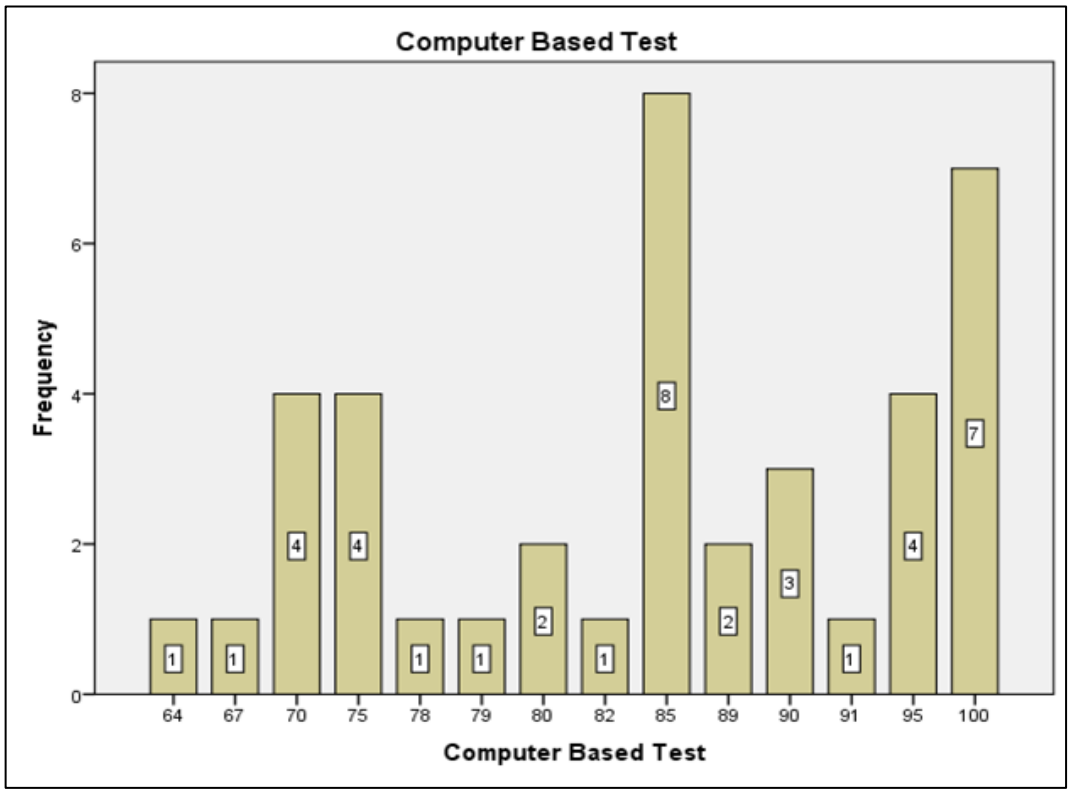

Figure 1. Distribution of Reading Skills Scores in CBT Classes

Based on the data shown in figure 1, it is known that the most achievable score is 85 (with a frequency of 8 students), then as many as seven students have a score of 100, 4 students each get a score of 95, 57 and 70, 3 students got a score of 90, 2 students each got a score of 89 and 80, while the smallest number what was the score of 64, 67, 78, 79, 82 and 91 (1 student each).

The distribution of score data on the reading skills of students in class using the PBT method can be illustrated in the following diagram:

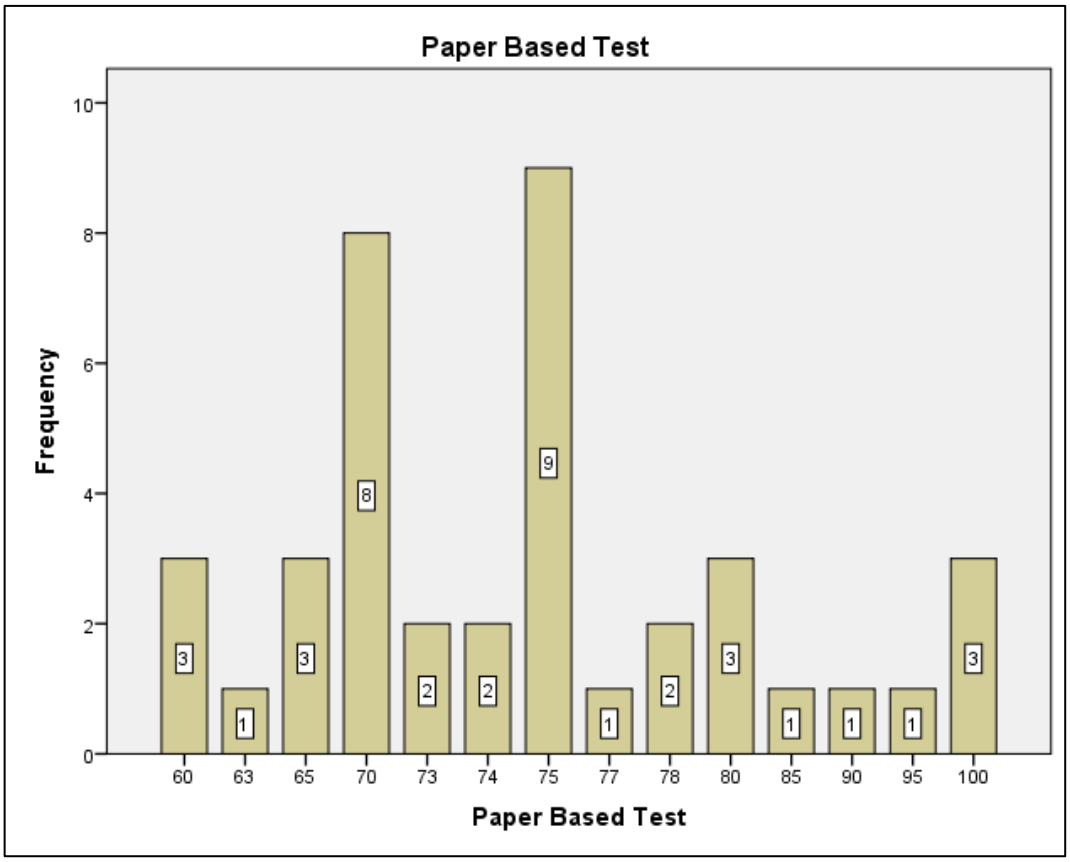

Figure 2. Distribution of Reading Skills Scores in PBT Classes

Based on the data displayed in Figure 2, it is known that the most achieved score is 75 (with a frequency of 9 students), then as many as eight students have a score of 70, 3 students each score 60, 65, 80 and 100, 2 students each scored 73, 74 and 78, while the smallest number what was the score of 63, 77, 85, 90 and 95 (1 student each).

B. Normality Test

The normality test is conducted to determine the distribution of research data and the type of statistical analysis used in bivariate tests (Correlation Test and Comparison Test). The normality test uses the Kolmogorov-Smirnov Test. Data is declared to be generally distributed if the Sig > 0.05.

Table 2. Normality Test Results

	Kolmogorov-Smirnov [a]			Shapiro Wilk		
	Statistics	df	Sig	Statistics	df	Sig
Computer Based Test	,116	40	,185	,943	40	,044
Paper Based Test	,210	40	,000	,883	40	,001

a. Lilliefors Significance Correction

Based on the test results, the significance value for the class using the CBT method was 0.185, while the class using the PBT method was 0,000. Because in the PBT class, the data is not normally distributed (Sig. <0.05), the overall data is not normally distributed. Bivariate testing will use non-parametric statistics because it does not meet normality assumptions.

C. Correlation Test

The correlation test for non-parametric statistics is the Spearman Rho Test. The following are the results of testing the correlation between classes using the CBT method and PBT:

Table 3. Correlation Test Results

			Computer Based Test	Paper Based Test	Coefficient of Determination
Spearman's rho	Computer Based Test	Correlation Coefficient	1,000	,249	0.062
		Sig (2-tailed)	.	,121	
		N	40	40	

Based on the test results, it is known that the value of the correlation coefficient between CBT and PBT is 0.249. The coefficient value is in the range of 0.200 - 0.400 or is in the criteria of a low relationship. This significance is 0.121> 0.05, indicating that the relationship between CBT and PBT is insignificant. The Determination Coefficient value is 0.062, indicating that the closeness between CBT and PBT is 6.2%.

D. Comparison Test Between Application of CBT and PBT Method

Because the previous test results show the data is not normally distributed, the hypothesis testing uses non-parametric statistical tests, namely the Mann-Whitney Test.

Table 4. Mean Rank for Each Group

	Group	N	Mean Rank	Sum of Ranks
Learning methods	CBT	40	50.83	2033.00
	PBT	40	30.18	1207.00
	Total	80		

The data testing results show that the average value of the class using the CBT method is 50.83, which is higher than the average value of the class using the PBT method, which is 30.18. These results indicate that the reading skills of students in the class that apply the CBT method is better than the classes using the PBT method, or it can be interpreted that the CBT method is more effective in improving students ' reading skills than the PBT method.

Furthermore, the significance of differences in skills scores between classes using CBT and PBT can be seen in the following table:

Table 5. Mann-Whitney Test Results

	Learning methods
Mann -Whitney U	387,000
Wilcoxon W	1207,000
Z	-3,998
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: Group

Based on the test results in Table 4.5, the U value of 387 and a W value of 1207 are shown. If converted to the value of Z, then the magnitude is -3.998. Asymp Value. Sig. (2-tailed) or PValue of 0.000 <0.05. Because the value of < value < critical value of 0.05, H0 is rejected and Ha is accepted, which means a significant difference exists between the classes applying the CBT method and the class using the PBT method.

Attachments

Description Statistics

	N	Minimum	Maximum	Mean	Std . Deviation
Computer Based Test	40	64	100	85.23	10,531
Paper Based Test	40	60	100	75.25	10,157
Valid N (listwise)	40				

Computer Based Test

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	64	1	2.5	2.5	2.5
	67	1	2.5	2.5	5.0
	70	4	10.0	10.0	15.0
	75	4	10.0	10.0	25.0
	78	1	2.5	2.5	27.5
	79	1	2.5	2.5	30.0
	80	2	5.0	5.0	35.0
	82	1	2.5	2.5	37.5
	85	8	20.0	20.0	57.5
	89	2	5.0	5.0	62.5
	90	3	7.5	7.5	70.0
	91	1	2.5	2.5	72.5
	95	4	10.0	10.0	82.5
	100	7	17.5	17.5	100.0
	Total	40	100.0	100.0	

Paper Based Test

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	60	3	7.5	7.5	7.5
	63	1	2.5	2.5	10.0
	65	3	7.5	7.5	17.5
	70	8	20.0	20.0	37.5
	73	2	5.0	5.0	42.5
	74	2	5.0	5.0	47.5
	75	9	22.5	22.5	70.0
	77	1	2.5	2.5	72.5
	78	2	5.0	5.0	77.5
	80	3	7.5	7.5	85.0
	85	1	2.5	2.5	87.5
	90	1	2.5	2.5	90.0
	95	1	2.5	2.5	92.5
	100	3	7.5	7.5	100.0
	Total	40	100.0	100.0	

Tests of Normality

	Kolmogorov-Smirnov [a]			Shapiro Wilk		
	Statistics	df	Sig .	Statistics	df	Sig .
Computer Based Test	,116	40	,185	,943	40	,044
Paper Based Test	,210	40	,000	,883	40	,001

a. Lilliefors Significance Correction

Correlations

			Computer Based Test	Paper Based Test
Spearman's rho	Computer Based Test	Correlation Coefficient	1,000	,249
		Sig. (2-tailed)	.	,121
		N	40	40
	Paper Based Test	Correlation Coefficient	,249	1,000
		Sig. (2-tailed)	,121	.
		N	40	40

Ranks

	Group	N	Mean Rank	Sum of Ranks
learning methods	CBT	40	50.83	2033.00
	PBT	40	30.18	1207.00
	Total	80		

Statistics Test

	learning methods
Mann -Whitney U	387,000
Wilcoxon W	1207,000
Z	-3,998
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: Group

Based on the data processing results, it is known that the scores of reading skills of students in the class using the CBT method score averaged 85.23. In class using the PBT method, students' reading skills have an average score of 75.25. With a KKM score (Minimum Completion Criteria) students of class XI is 72. This shows that students in SMAN 3 Serang City feel more comfortable doing tests with the help of computer functions. Compared to the Paper Based Test (PBT) model, Computer Based Test (CBT) has advantages such as when compared to written tests, in this test participants can immediately find out the results of the test, if compared to oral tests, this test can be done simultaneously with many participants with a relatively short time, students feel more free and confident in working on problems, reducing occurrence cheating during the exam, because every participant will get different questions with the same level of difficulty, and test-based computer or CBT is more objective than PBT because the question is given directly by the computer and also corrected by computer. The test computer is done using computer software to submit test questions, accommodate participants responses test, and then store and analyze electronically. There is also an exam using a directly connected computer with internet networks, which are often called online tests. Computer Based Test exams also have weaknesses, especially if done online, which is the risk of misunderstanding for participants high enough if the use of language is less firm and straightforward; the risk is system damage can occur and hinder the continuity of testing, especially if there are hacker attacks, sniffing or attacks on others networks and with many participants if not equipped with adequate capacity can cause delays system and interfere with the continuation of the test.

CONCLUSION

From the results of the study, it can be concluded that The results of testing the data show that the average value of the Class using the CBT method is 50.83 which is higher than the average value in the Class using the PBT method of 30.18. These results indicate that the reading skills of students in the class that apply the CBT method are better than the classes using the PBT method, or it can be interpreted that the CBT method is more effective in improving students ' reading skills than the PBT method. Based on the results of testing the significance of the difference in skill scores between CBT and PBT classes, the U value is 387 and the W value is 1207. If converted to the value of Z, then the magnitude is -3.998 Asymp Value . Sig . (2-tailed) or PValue of 0.000 <0.05. Because the value of < value < critical value of 0.05, H0 is rejected, and Ha is accepted, which means a significant difference exists between the classes applying the CBT method and the class using the PBT method.

REFERENCES

Allen, M. J. & Yen, W. M. (1979). Introduction to Measurement Theory . Monterey : Brooks /Cole.

Ary, et al. 2010. 2006. Introduction to Research in Education. Wadsworth : Cengage Learning

Bennett , R. E. (2002). Inexorable and inevitable : the continuing story of technology and assessment . The Journal of technology , Learning , and Assessment , 1(1), 1-23.

Bunderson , C.V., Inouye , D.K., & Olsen , J.B. (1988). The four generations of computerized educational measurement . ETS Research Report Series , 1988 (1). https://doi.org/10.1002/j.2330-8516.1988.tb00291.x

Bunz , Ulla. (2002). Growing From Computer Literacy Towards Computer Mediated Communication Competence : Evolution of a Field and Evaluation of a New Measurement Instrument , Department of Communication , Rutgers , The State University of New Jersey , New Brunswick , NJ

Burt , M., Peyton, J. K., & Adams , R. 2003. Reading and Adult English Language Learners : A Review of the Research . Washington, DC: Center for Applied Linguistics .

Charman , D. (1999). Issues and impacts of using computer based assessments (CBAs) for formative assessment . In S. Brown , J. Bull , & P. Race (Eds .), Computer-Assisted Assessment in Higher Education (pp . 85-94).

Chen, K. T. C. (2012). Elementary EFL teachers ' computers phobia and computer self-efficacy in Taiwan. TOJET: The Turkish Online Journal of Educational Technology, 11 (2).

Choi , I.C., Kim, K.S., & Boo , J. (2003). Comparability of a paper based language test and a computer based language test . Language Testing, 20 (3), 295-320. https://doi.org/10.1191/0265532203lt258oa

Choi , S. W. , & Tinkler , T. (2002). Evaluating comparability of paper - and -pencil and computer based assessment in a K-12 setting . Paper presented at the 2002 annual meeting of the National Council on Measurement in Education .

Clariana , R. & Wallace , P. (2002) Paper- based versus computer-based assessment : key factors associated with the test mode effect . British Journal of Educational Technology. 33(5) 593-602.

Darrell LB (2003). The Impact of Computer Based Testing on Student Attitudes and Behavior . Technol . Source.Retrieved from http://ts.mivu.org/default.asp?show=article&id=1034.

Fleming , S. & Hiple , D. (2004). Foreign language distance education at the University of Hawai'i . In CA Spreen , (Ed.), New technologies and language learning : issues and option (Technical Report #25) (pp . 13-54). Honolulu , HI: University of Hawai'i , Second Language Teaching & Curriculum Center.

Grabe , William and Fredricka L. Stoller . 2011. Teaching and Research Reading . London: Pearson Education Limited .

Harmer , Jeremy. 1998. How to Teach English . Edinburgh : Addison Wesley Longman .

Harrison , Colin. 2004. Understanding Reading Development. London: Sage Publications.

International Test Commission . (2006). International guidelines on computer based and internet -delivered testing. International Journal of Testing, 6 (2), 143-171.

https://doi.org/10.1207/s15327574ijt0602_4

J. DeRosa , The Green PDF: Reducing Greenhouse Gas Emissions One Ream at a Time , 2007, http://www.scribd.com/doc/6077919 5/The-Green-PDF- Revolution Journal of LLT, 11, 3-9

Kim, D.H., & Huynh , H. (2007). Comparability of computer and paper - and -pencil versions of Algebra and Biology assessments . Journal of Technology, Learning , and Assessment , 6(4), 1-35.

Kingston, N.M. (2009). Comparability of computer - and paper-administered multiple choice tests for K-12 populations : A synthesis . Applied Measurement in Education , 22(1), 22-37.

Laborda , J. (2007). On the Net Introduction Standardized EFL/ESL Exams .

Leahy , S., Lyon , C., Thompson , M., & William, D. (2005) . Classroom assessment minutes by minute , day by day . Educational Leadership , 63(3), 19-24.

Lottridge , S., Nicewander , A., Schulz , M. & Mitzel , H. (2008). Comparability of Paper-based and Computer based Tests : A Review of the Methodology .Pacific Metrics Corporation 585 Cannery Row , Suite 201 Monterey , California 93940

Noyes , J.M., and Garland , K. J. (2008). Computer - vs. paper-based tasks : Are they equivalent ? Ergonomics . 51(9), 1352 – 1375.

Nuttal , Christine. 2000. Teaching Reading Skill in a Foreign Language (New Edition). London: MacMillan Heinemann .

O'Malley , K.J., Kirkpatrick , R., Sherwood , W., Burdick , H.J., Hsieh , MC&, Sanford , E.E. (2005, April). Comparability of a Paper Based and Computer Based Reading Test in Early Elementary Grades . Paper presented at the AERA Division D Graduate Student Seminar, Montreal , Canada .

OECD. (2010). PISA Computer-based assessment of student skills in science . https://doi.org/10.1787/9789264082038-en

Peter C, Bill I and David S (2004). Using computer for Assessment in medicine . British Med . J., 329(7466):606-609.

Pommerich , M. (2004). Developing computerized versions of paper - and -pencil tests : Mode effects for passage-based tests.The Journal of Technology, Learning , and Assessment , 2(6), 3-44.

Russell , M., & Haney , W. (2000).The gap between testing and technology in schools . Education Policy Analysis Archives,8(19), 1-10.

Scheuermann , F., & Pereira , A. G. (2008). Towards a research agenda on Computer based Assessment

Sorana -Daniela B, Lorentz J (2007). Computer-based testing on physical chemistry topic : A case study. Inter . J. Educ . Develop . using Info. Com . Technol ., 3(1): 94-95.

Subali, Bambang. 2012. Principles of Assessment and Evaluation of Learning. Yogyakarta: UNY Press

Sudaryono . 2012. Basics of Learning Evaluation. Yogyakarta: Graha Ilmu.

Suryabrata , Sumadi. 2003. Research Methodology. Jakarta: Raja Grafindo Persada