International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 15, Issue 4, 2021

An Indonesian Case Study of Computer Operating Familiarity Levels on Computer Based Tests at Vocational High Schools

Riana T. Mangesa[1]; Iwan Suhardi[2]; Jumadi Mabe Parenreng[3], [1,2,3]Department of Electrical Engineering Education, [1,2,3]State University of Makassar, Indonesia

The research was applied to two groups respondent students of class XII Vocational High School with the ability to do the same questions, but different test media. Computer Based Tests (CBT) respondents were assigned random assignments. The results showed the average difference in the results of the assessment between the Paper Based Test (PBT) and CBT was 19.85 with a standard deviation of 2.076. The results of statistical calculations showed a t count (9.562) > t table (1.972) with a probability < 0.05. The level of familiarity in respondents corresponds with operating a computer, but it turns out that the average score of the CBT test model is lower (49.87) compared to PBT (69.72). This shows that testing using CBT requires students to be more than able to operate computers. It requires the habit of using the CBT model so as not to be psychologically burdened by anxiety factors.

Key words: *Familiarity in Operating Computers, PBT, CBT*

1. Introduction

Technological developments around the world support change and the world has seen many changes in various fields due to technological advances. Progress in education is one of the areas that is greatly influenced by these developments (Ozbal and Eski, 2019). Even the development of information technology supports the improvement of skilled human resources.

Empirically, several research results have shown that the world of education is greatly influenced by technological media, because it is effectively used in various activities both as a means of learning and as a source and means of the learning process. Activities measuring the achievement of student learning success can be done through the implementation of various evaluations instruments as measuring tools in collecting student learning outcomes.

However, sometimes there are obstacles in the implementation of tests (evaluations) that can shift the meaning of student learning outcomes. For example, they could cheat by looking at notes, looking at a friend's work or, regarding a test that was still using conventional methods such as using a Paper Based Test (PBT), answer sheets. Therefore, it is better if the test is carried out using an online-based computer. Several educational practitioners have provided solutions to these problems by developing various evaluation tools including the development of computer-based test products.

Along with the improvement and use of information technology, the learning outcome evaluation tool has also undergone rapid changes according to the development of Information Communication and Technology (ICT). Assessment systems that present computer-based exams are better known as Computer Based Testing (CBT). CBT has advantages compared to Paper Based Test (PBT), namely (1) increasing standardisation, (2) increasing test safety, (3) increasing test display capabilities, (4) reducing error of measurement, and (5) accelerating scoring and interpretation (Bunderson, Inouye & Olsen, 1989).

Penetration of the role of CBT began to gradually replace the function of PBT (Bugbee, 1996). A concrete example of the development of CBT is the policy of the National Examination in Indonesia. In the implementation of the National Examination in Indonesia, the test model using paper media is known as the Paper and Pencil-based National Examination (UNKP), while those using computers are known as the Computer-Based National Examination (UNBK).

The research results of Yohanes Adio Balani (2017) showed that the CBT model based on Adobe Flash software is valid, practical and effective for Vocational High School learning. The National Examination provider assumes that the items displayed on the monitor screen have the same level of difficulty when presented with paper media. With this assumption, the test results of the two test models are considered equivalent.

Psychometrically, there is almost no difference in CBT compared to PBT. PBT and CBT use the same number of items for each participant or fix-length test. In scoring, PBT and CBT use Classical Test Theory (CTT). Judging from the context and atmosphere faced by the examinees, PBT and CBT have striking differences. The comparison is presented in Table 1.

Table 1. Comparison Context and Atmosphere of PBT and CBT

Context and Atmosphere	PBT	CBT
The number of items in the range	Consists of many grain	Usually there is only one question, otherwise it must be scrolled.
Test aid	Paper and pencil	Monitor screen, CPU, keyboard, mouse, and speakers
Form of questions that can be displayed	Text and images	Text, images, audio and video
the model works on the items	Marking the answer choices with a pencil	Select answers with the mouse or keyboard
Item color	Generally black	All colours possible
The habit factor	Aready accustomed	Has not become a habit
Basic knowledge of computer operation	Not required	Is indispensable

Differences in the context and atmosphere aspects between PBT and CBT can influence the results of the test scores and the validity of students' ability estimation. Therefore, the readiness of the CBT system with the support of software and hardware as well as the test taker's computer skills are the main requirements in implementing computer-based tests. In addition, the security and confidentiality of test results must be maintained. Expert psychometrics, such as Rudner (1998), argue that the parameter items used on PBT may not be in accordance with their appearance on computer monitor screen.

The fundamental difference is that CBT requires test takers to be familiar with the basics of operating computer equipment. To see and choose the items, and determine the desired answer from the list of choices, students must be able to operate the keyboard and mouse. The higher the level of familiarity of students in operating computers and the habit of using computer-based testing models will be more supportive when students work on CBT.

Students of Vocational High Schools are accustomed to working on exam items on paper media and have not been accustomed to using CBT. Differences in characteristics between PBT and CBT, and familiarity with using CBT, can affect student psychology. The influence of anxiety and anxiety factors leading up to and on the exam process can cause students to not be able to focus on their tests well. The interaction between individual differences (e.g., previous computer familiarity, attitudes, and reluctance) and CBT performance will affect the equality where score equality and reliability are substituted for CBT with PBT. Along with the demands of globalisation to implement CBT over PBT, the level of familiarity and habits of students in

operating computers needs to be analysed further. This study aims to assess the level of familiarity in operating computers on the results of PBT and CBT exam scores.

Judging from the testing model used in this study, in general there are two test models, namely CBT and PBT. As globalisation demands that CBT replace PBT, CBT has begun to be used by the National Examination in Indonesia. It is assumed that the items displayed on the monitor screen will have the same difficulty level when presented on paper media.

One of the fundamental differences is that the CBT testing model requires the test taker to be familiar with the basics of operating computer equipment. Therefore, the main objective of this study to assess the level of familiarity when operating computers against the results of PBT and CBT exam scores. The sub-objectives are as follows:

1. Knowing the comparison of the results of the PBT and CBT group exams.
2. 2. Assessing the level of computer operation familiarity with the results of PBT and CBT exam scores.

2. Methods

This research uses a combination of development methods and quantitative studies. The development method is carried out to develop the test package test questions and CBT software. As the technical work on the PBT, CBT software is designed so that respondents can choose the number of items desired and can review the response to the answer if you want to correct the answer. CBT software developed using a client-server system is accessed through a LAN network. Each respondent for CBT testing uses 1 (one) complete set of computer equipment. Statistical tests comparing the results of the PBT and CBT values were carried out with SPSS using the T-test of two free samples. This method was chosen with the consideration that the respondents are groups with different respondents.

The sample was formed from 2 *SMK* class XII. The respondents consisted of 200 students and were spread over 2 groups each of 100 respondents who were considered to have the same ability. One group worked on a question package using PBT and the other group using CBT. A total of 68 questionnaires were distributed to respondents in the CBT group randomly to determine their level of familiarity in operating computers and their experience using computer-based testing media.

Furthermore, the results of the test scores of the groups of respondents were further analysed by SPSS to be statistically tested as to whether there were significant differences in the mean between the two test models. Results of the data from the questionnaire were processed in a descriptive form to determine the level of student familiarity in operating computer devices and the habits of using the computer-based testing model. The block diagram in this study is presented in Figure 1.

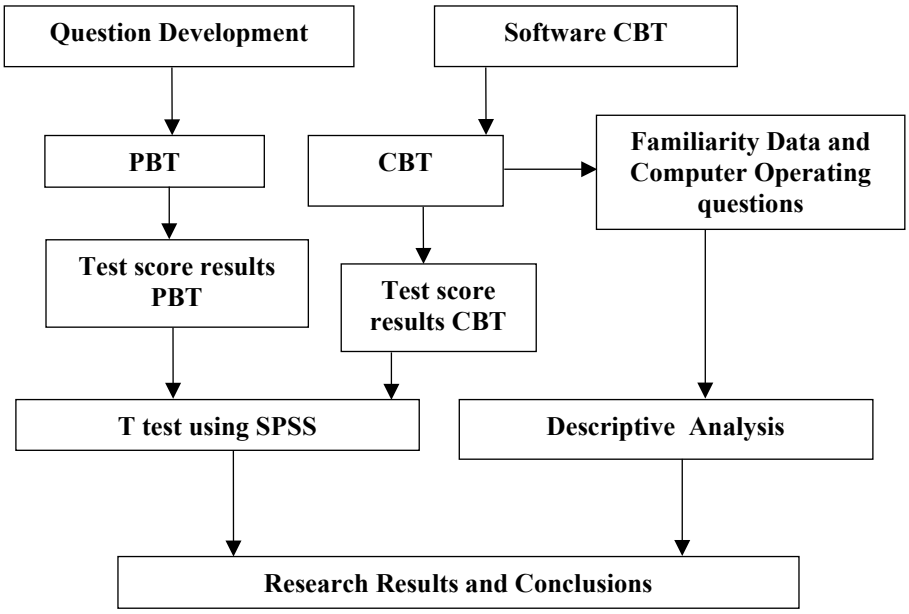

Figure 1. Block Diagram Research

3. Results

This section includes the test findings of the two groups of PBT and CBT tests. The test results show that the average respondent's CBT value is lower than that of PBT. Of the 200 class XII student respondents, the mean difference between the PBT and CBT was 19.85 with a standard deviation of 2.076. The average value of PBT was 69.72 and CBT was 49.87. From the results of the difference between the two scores, it can be concluded that students have more difficulty working with CBT than PBT for the same package questions. The results of statistical calculations can be seen in Table 2, Table 3 and Table 4.

Table 2. Results of PBT and CBT Test Values

Score PBT					Score CBT			
A	B	C	D	No	E	F	G	H
65,00	70,00	72,50	77,50	01	20,00	50,00	55,00	55,00
62,50	70,00	77,50	70,00	02	15,00	52,50	55,00	40,00
60,00	62,50	65,00	92,50	03	25,00	55,00	55,00	40,00
55,00	57,50	47,50	97,50	04	50,00	52,50	55,00	47,50
55,00	60,00	67,50	62,50	05	75,00	65,00	77,50	37,50
67,50	52,50	97,50	62,50	06	42,50	52,50	55,00	45,00
70,00	47,50	100,00	55,00	07	55,00	52,50	47,50	57,50
92,50	55,00	67,50	62,50	08	60,00	22,50	65,00	47,50
72,50	52,50	60,00	75,00	09	57,50	25,00	72,50	42,50
85,00	47,50	67,50	80,00	10	52,50	67,50	60,00	52,50
67,50	62,50	67,50	57,50	11	55,00	55,00	60,00	45,00
60,00	50,00	67,50	87,50	12	47,50	47,50	47,50	42,50

65,00	62,50	70,00	75,00	13	32,50	70,00	47,50	35,00
72,50	62,50	95,00	97,50	14	52,50	60,00	55,00	45,00
60,00	87,50	70,00	77,50	15	55,00	67,50	15,00	30,00
65,00	65,00	60,00	65,00	16	45,00	65,00	20,00	42,50
100,00	52,50	95,00	52,50	17	37,50	67,50	10,00	40,00
70,00	75,00	87,50	50,00	18	72,50	55,00	37,50	52,50
65,00	32,50	92,50	70,00	19	52,50	55,00	30,00	47,50
62,50	32,50	92,50	67,50	20	37,50	62,50	45,00	37,50
60,00	92,50	60,00	82,50	21	37,50	70,00	60,00	45,00
60,00	82,50	97,50	57,50	22	40,00	65,00	50,00	55,00
77,50	82,50	67,50	60,00	23	50,00	60,00	52,50	47,50
70,00	87,50	95,00	67,50	24	37,50	57,50	30,00	87,50
65,00	95,00	57,50	70,00	25	37,50	55,00	30,00	42,50

Table 3. Group Statistics

	Groups	N	Mean	Std. Deviation	Std. Error Mean
Score	PBT	100	69.725	14.8515	1.4852
	CBT	100	49.875	14.5052	1.4505

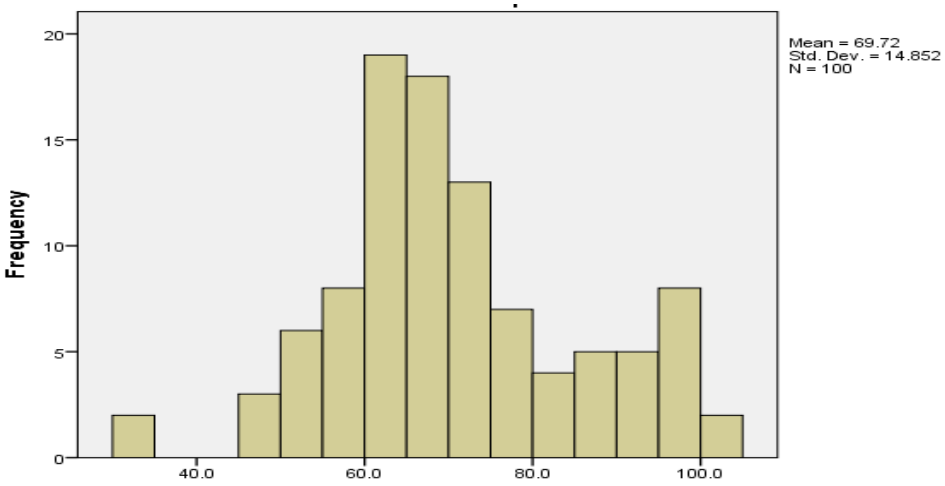

Figure 2. Histogram Group PBT

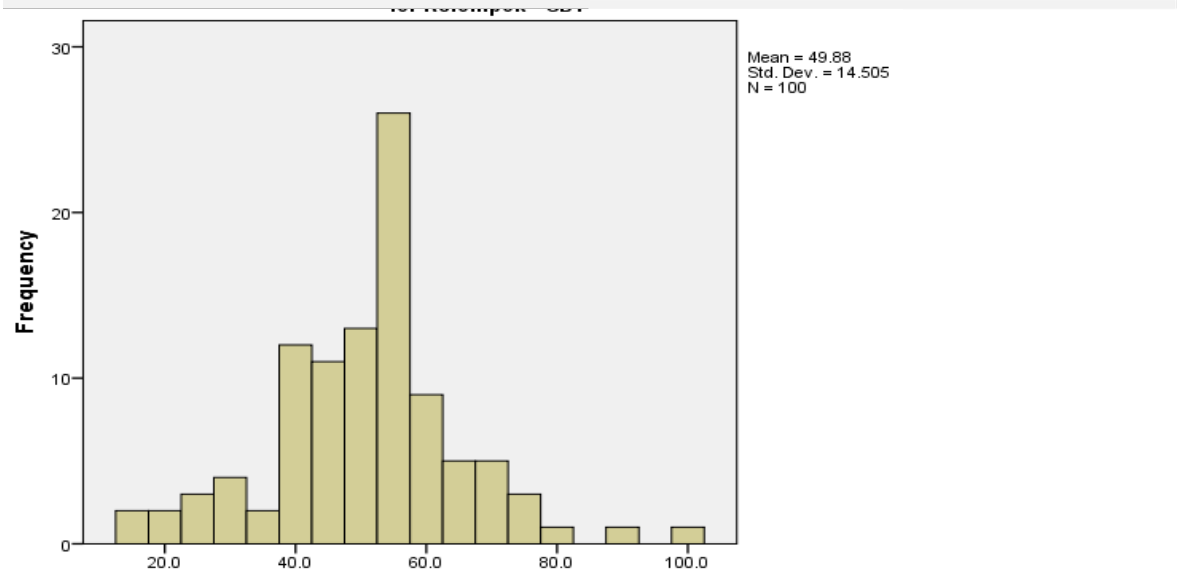

Figure 3. Histogram Group CBT

The statistical test of the comparison of the results of the PBT and CBT values was carried out with SPSS using the two-sample free T test. This method was chosen with the consideration that the respondents are groups with different respondents. To determine the difference in the results of the PBT and CBT test scores, the variance similarity test is carried out using the F test value (Levene's Test).

Table 4. Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Score	Equal variances assumed	.283	.595	9.562	198	.000	19.850	2.0760	15.76	23.954
	Equal variances assumed			9.562	197.89	.000	19.850	2.0760	15.76	23.954

The steps taken to test the variance similarity are:

1. Formulate research hypotheses

There are differences in variance between the results of PBT and CBT group

2. Formulate operational hypotheses (null and alternative hypotheses):

H0: There were no differences in variance in the results of the assessment between the PBT and CBT groups.

Ha: There are differences in the variance in the assessment results between the PBT and CBT groups.

3. Determine the level of trust used.

The confidence level is 95% or by using alpha 5%.

4. Determine the rules for decision making.

The decision making is to accept H0 if probability or significance is > 0.05 and reject H0 if probability or significance is <0.05.

5. The Decision making and interpretation of results.

F-test numbers that assume the two variants are equal are 0.283 with a probability of 0.595. Because the probability value is > 0.05, then H0 is accepted and Ha is rejected. This means that there is no difference in the variance in the assessment results between the PBT and CBT groups.

Furthermore, because there was no difference in the variance of the assessment results between the PBT and CBT groups, then an average test was carried out between the two test models. Comparing the average test results of the PBT and CBT groups is done by using t-test numbers which assume both variances are equal (equal variances assumed).

The steps taken to test the average value of PBT and CBT are:

1. Formulate research hypotheses.

There are differences in the average results between the PBT and CBT groups.

2. Formulate operational hypotheses (null and alternative hypotheses)

Ho: The average assessment results with PBT and CBT are the same.

Ha: The average results of assessment with PBT and CBT are not the same.

3. Determine the level of trust used.

The confidence level is 95% or by using alpha 5%.

4. Determine the rules for decision making.

The decision-making rule is to accept Ho if t count is smaller than t table and to reject Ho if t count is greater than t table. Based on t table with 5% alpha 2-way test or 2.5%; and the degree of freedom df = 198, a t table value of 1.97202 was obtained. The decision taken is to accept Ho if t arithmetic is smaller than 1.97202 and reject Ho if t arithmetic is greater than 1.97202.

5. Calculate t statistics

To calculate the statistical value, the SPSS assist program is used, t count is 9.562

6. Decision making and interpretation of results.

The average difference in the assessment results between PBT and CBT is 19.85 with a standard deviation of 2.076. The results of the t statistical calculation resulted in a value of 9.562 and a significance of 0,0001. From the SPSS output it was found that t arithmetic (9.562) > t table (1.972) with probability number <0.05, therefore Ho was rejected, and Ha accepted. This means that there is a difference in the results of the exam scores if the same package of questions is displayed with PBT and CBT.

Furthermore, the data distribution can be seen in the normal detrended Q-Q plot image which explains the distribution of the data. Each Group PBT and Group CBT value can indicate a value below or above the normal line which represents how far or high the value is. The normal detrended Q-Q plot can be seen in Figure 4 and Figure 5

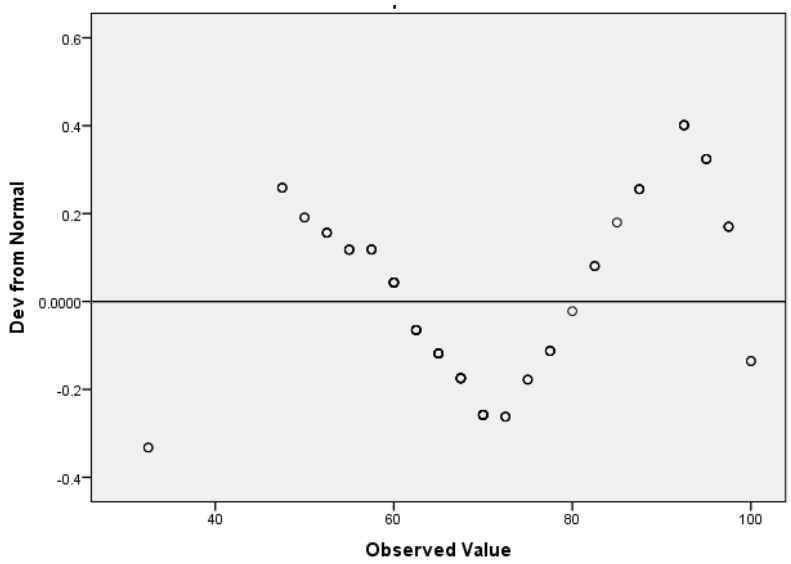

Figure 4. The Detrended Normal Q-Q PLots Group PBT

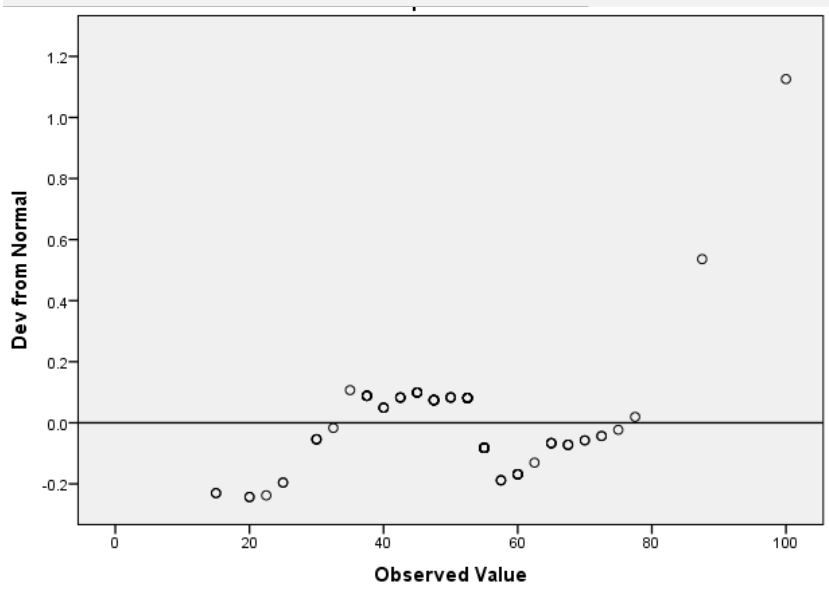

Figure 5. Detrended Normal Q-Q Plots Group CBT

The results of the study indicate that the question packages displayed by CBT monitor using a mouse and keyboard have different levels of difficulty when presented on a sheet of paper using a pencil. Based on the difference between the two scores, it can be concluded that students find it more difficult when working with CBT compared to PBT for the same question package.

Along with the demands of globalisation to implement CBT to replace PBT, further testing is carried out for the level of familiarity and habits of students in operating computers. In this study the sample was taken randomly from respondents in the CBT group. In the CBT group, 68 questionnaires were distributed randomly. This questionnaire contains students' responses to aspects of familiarity in operating computers and their responses to CBT. The results of the questionnaire analysis can be displayed in tabular form as follows:

1. Familiarity Aspects of Computers

Table 3. Respondents' Familiarity Responses to Computers

No	Familiarity Aspects of Computers	Criteria	Percentage (%)
1	How long have you been using a computer?	more than 3 years	72
		between 2 - 3 years	23
		between 1 - 2 years	5
		less than 1 year	0
2	How often do you do internet activities?	more than 3 years	72
		between 2 - 3 years	25
		between 1 - 2 years	3
		less than 1 year	0
		very familiar	32

3	Are you used to using computer equipment and supporting equipment (such as keyboard and mouse)?	familiar	65
		less familiar	1
		unfamiliar	2
4	How often do you use computer in your activities?	very familiar	28
		familiar	52
		less familiar	20
		unfamiliar	0
5	How often do you use internet access in your activities?	very familiar	35
		familiar	58
		less familiar	7
		unfamiliar	0

Table 3 is the result of the analysis of the respondent's data, which shows that 72% of the respondents are familiar with computers and have used computers for a long time. Furthermore, 65% are familiar with computers and 58% of respondents are familiar with and frequently use the internet on computer devices. Therefore, there should be no constraints on using computers for CBT.

2. General Assessment Aspects of CBT

Table 4. General Assessment Aspects of CBT

No	General Assessment Aspects of CBT	Criteria	Percentage (%)
1	Before you used this CBT, how often did you use the computer for testing purposes?	very often	0
		often	15
		rarely	77
		never	8
2	Do you feel happy and interested when using this CBT?	very often	23
		often	65
		rarely	12
		never	0
3	What do you do when you use this CBT?	very often	18
		often	33
		rarely	49
		never	0
4	How do you think CBT is more fun than the regular test (using paper and pencil)?	very often	17
		often	76
		rarely	7
		never	0

Table 4 shows that most respondents rarely (77%) use computers for testing purposes. However, 76% agreed to use CBT instead of paper and pencil. The fact is that there are still

many schools that have not implemented testing applications using computers due to constraints on school facilities or constraints on the ability of teachers. From the results of the questionnaire it can also be concluded that 65% of the majority of respondents feel happy and interested in using the CBT program, and consider testing with CBT more attractive than PBT. Only 49% of respondents stated that they are still not happy about doing CBT just because they are not used to it.

If the comparison of the results of the test scores produced on PBT and CBT obtained significant differences, it means that students feel a different atmosphere when working on the same problem between PBT and CBT. If it is found that the average level of difficulty of items on PBT is greater than the average level of difficulty of items on CBT, then students are stated not to have a good level of habit in using computers for testing facilities. In this criterion, students still feel the psychological burden of the different testing models from paper-based to computer-based.

4. Conclusion and Discussion

Based on the results of the study and analysis, it was concluded that statistically there were differences in the results of the scores if the items were displayed with PBT and CBT. The difference in test results can be caused by the unfamiliarity of students using computer-based testing models. The admission of students who are familiar with operating computers does not guarantee that these students are familiar with computer-based testing models. From the findings of the differences in the two scores between the average PBT score of 69.73 and CBT 49.87, it can be concluded that students find it more difficult to do CBT compared to PBT for the same package questions.

Along with the research results of Patel et al. (2014) who pointed out that there was an overwhelming disruption between PBT and CBT exams, students who scored "A" dropped dramatically from 29% on exams using PBT to 19% on CBT exams. The percentage of students failing the test increased to 26% on the on-screen test from 3% on the on-paper test. This shows that testing using CBT requires students to be more than just being able to operate computers but requires the habit of using computer-based testing models so that students are not burdened psychologically by anxiety factors. According to Bellotti, et al (2013), two goals of computer-based learning and testing are that they should be (1) fun and entertaining, (2) educational and challenging.

Furthermore, the results of the study show that the question packages displayed on the monitor screen of the CBT software using a mouse and keyboard have different levels of difficulty when presented on a sheet of paper using a pencil. In addition, differences in context and atmosphere, the familiarity factor of operating computers and the unfamiliarity of students working on questions with a computer-based testing model can result in differences in score

results between PBT and CBT. One of the crucial limitations in implementing CBT is that it relates to aspects of the test taker's familiarity in using a computer device. Along with the statements of Rudner (1998) and Grist (1989), one of the crucial limitations in the application of CBT is related to the familiarity of test takers in using computer devices. Therefore, according Sumiati (2018), students have to practice more to work on questions on CBT. Teachers should motivate students and familiarise themselves with CBT to face the UN.

On the other hand, one way to reduce test anxiety using a computer is to improve students' computer experience and confidence in taking computer-based exams (Zeidner & Matthews, 2003; Liebert & Morris, 1967). Providing opportunities for students to become familiar with CBT is important (Russell, 1999). Familiarising students with increasing CBT trials before test day can reduce anxiety factors. Familiarising students with computer-based exams will benefit students who are economically disadvantaged and do not have computers at home to enhance their computer operating experience.

Having the basic ability to operate a computer device is not a guarantee that students are familiar with the computer-based testing model. This is considering that the habit of using testing with paper media has been going on for years, while the socialisation of testing using a computer is done in a short time. The unfamiliarity of taking computer-based exams makes students unable to show their best abilities when taking exams. The habitual factor of students working on items using PBT unwittingly has a psychological effect that is less supportive when using a computer-based testing model. Bellotti, et al (2013) recommends that in the future there is a need for an increase in computer-based testing.

In the use of CBT, it is important to consider aspects of computer self-efficacy, namely how confident a student sees himself in working on a computer-based test. Computer self-efficacy helps reduce student anxiety levels in taking computer-based exams (Compeau & Higgins, 1995; Surej, 2013; Sam, Othman, & Nordin, 2005). With the reduced level of anxiety, students can focus more on working on problems and can show their best abilities.

REFERENCES

Bates, A. (2005). Technology, e-learning and Distance Education. London: Routledge, https://doi.org/10.4324/9780203463772

Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of serious games: An overview. *Advances in Human-Computer Interaction*, *2013*(February). https://doi.org/10.1155/2013/136864

Bunderson, C. V., Inouye, D. K., & Olsen, J. B. (1988). the Four Generations of Computerized Educational Measurement. *ETS Research Report Series*, *1988*(1), i–148. https://doi.org/10.1002/j.2330-8516.1988.tb00291.x

Čandrlić, S., Katić, M. A., & Dlab, M. H. (2014). Online vs. Paper-based testing: A comparison of test results. *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2014 - Proceedings*, 657–662. https://doi.org/10.1109/MIPRO.2014.6859649

Coffman, T. (2009). *Engaging Students Through Inquiry-oriented Learning and Technology*. Rowman & Littlefield Education. https://books.google.co.id/books?id=pP3soZoH6Y4C

Compeau, B. D. R., & Higgins, C. A. (1995). *Computer Self- Efficacy ’. Development of a Measure and Initial Test*. *June*.

DeSouza, E., & Fleming, M. (2003). A comparison of in-class vs. online quizzes on student exam performance. Journal of Computing in Higher Education, 14, 121-134.

Ghaderi, M. (2014). *Comparing Between Computer based Tests and Paper-and-Pencil based Tests*. *2*(4), 2005–2007. https://doi.org/10.7575/aiac.ijels.v.2n.4p.36

Gordon, A. M. (2015). *Paper Based Testing vs . Mobile Device Based Testing in an EFL Environment : What ’ s the Difference ?*

Heafner, T. (2004). Using technology to motivate students to learn social studies. *Contemporary Issues in Technology and Teacher*, *4*(1), 42–53. http://www.citejournal.org/vol4/iss1/socialstudies/article1.cfm

John, S. P. (2013). Influence of Computer Self-Efficacy On Information Technology Adoption. *Internation Journal of Information Technology*, *19*(1), 1–13.

Liebert, R.M., & Morris, L.W.(1967). *Cognitive and emotional components of test anxiety: a distinction and some initial data*. Psychological Reports20975–978. http://dx.doi.org/10.2466/pr0.1967.20.3.975

Patel, Ayyub Ali, Mohammed Amanullah, Khalid Mohanna and Sarah Afaq (2014). *E-exams under E-learning System: Evaluation of On-Screen Distraction by First Year Medical Students in Relation to On-Paper Exams.* ISBN: 978-1-4799-3166-8 ©2014 IEEE

Sumiati. (2018). *Students Perception Towards National Examination:Computer-Based Test Or Paper-Based Test*. JIPI S V o. 2 7 No. 1