Digital Test Instruments Based on *Wondershare-Superitem* for Supporting Distance Learning Implementation of Assessment Course

Dewa Gede Hendra Divayana
Department of Information Technology Education, Universitas Pendidikan Ganesha, Indonesia, *hendra.divayana@undiksha.ac.id*

I Gede Sudirtha
Department of Vocational Education, Universitas Pendidikan Ganesha, Indonesia, *gede.sudirtha@undiksha.ac.id*

I Kadek Suartama
Department of Education Technology, Universitas Pendidikan Ganesha, Indonesia, *ik-suartama@undiksha.ac.id*

This research had the main objective to develop a new form/design in the development of test instruments that valid and reliable. The form of the test instrument that was developed adopted the Superitem concept and was integrated into software called Wondershare. Through the integration, digital format test instruments were realized with the structures of the questions' difficulty level which were tiered from the easiest to the hardest stages. The development of these digital test instruments used a quantitative approach to obtain valid and reliable instruments. The stages of development included: 1) determination of the instrument items, 2) test the validity and reliability of instrument items, 3) determination of final instruments as digital test questions, 4) inputting digital test questions into the Wondershare application, 5) trials of the digital test instruments, 6) revision of digital test instruments, and 7) finalization of digital test instruments. The content validation test of instruments involved two education experts, the reliability test of instruments involved 42 respondents, and the trials of digital test instruments based on Wondershare-Superitem involved 12 respondents. Data collection related to content validation test, instrument reliability, and trials of digital test instruments used the instruments in the form of questionnaires. Analysis of the trial results of digital test instruments using analysis techniques of quantitative descriptive by percentage descriptive formula. The results of this research showed the design of digital format test instruments based on the Wondershare-Superitem that are valid and reliable for supporting distance learning implementation of assessment course.

Keywords: digital test instruments, wondershare, superitem, distance learning implementation, assessment course

INTRODUCTION

The 'Assessment' course is one of the educational field courses that must be given to students who are studying at universities that include *LPTK* (Education Institutions and Educational Personnel) in Indonesia so that students understand the importance of an assessment in an educational organization. The learning process of the 'Assessment' course can be said to have gone well if most students have understood the content of the courses as a whole, starting from the easiest to the hardest levels. To determine the level of understanding of students in studying the 'Assessment' course, so a test is needed.

The fact that often occurs in the field shows that the test items used by lecturers to assess students' ability to understand the 'Assessment' course is still limited to the essay tests with the difficulty level of each test item which is sometimes very easy or difficult. It making difficult to know the level of understanding of students in learning that course from the lowest level to the highest level. Besides, the essay test presented by the Lecturer to the students is still in the form of printouts, so sometimes some lecturers do not present test questions in the form of printouts, but instead, test questions are written on the blackboard, so that it is difficult to realize distance learning later.

Based on that reality, so becomes a necessity in the era of information technology advancement today for lecturers to develop themselves and innovate to develop a form of test used to assess the extent of their students' ability to understand the material being taught in the class or through the distance learning. The form of the test for the 'Assessment' course that needs to be developed is a digital format test that is made using free applications from the internet. Furthermore, the question model inserted in the digital format test application that has been formed must be able to adopt the *Superitem* concept, so that the questions raised in the test application can measure the level of student's understanding about the assessment material from the lowest difficulty level to the highest difficulty level.

Based on the problems found in the field, then several questions proposed of this research, included: 1) What was the digital test instrument items for the 'Assessment' course that adopts the *Superitem* concept?; 2) How validated the digital test instrument items for 'Assessment' course that adopt the *Superitem* concept?; 3) How did the reliability of the digital test instruments for 'Assessment' course that adopts the *Superitem* concept?; and 4) How were the results of the *Wondershare* application trials that used to present the digital test instruments for the 'Assessment' course that adopts the *Superitem* concept?

The long-term goal of this research was to find the right digital test instruments to measure the cognitive abilities of students in the 'Assessment' course. The short-term goal of this research was to be able to develop digital test instruments that valid and reliable by adopting the *Superitem* concept so that it can measure students' cognitive abilities towards the 'Assessment' course presented used the *Wondershare* application at the Department of Information Technology Education, Faculty of Vocational and Technology, *Universitas Pendidikan Ganesha*.

The idea of this research began because it was motivated by the research conducted by Firmasari, Sukestiyarno, and Mariani (2013) that showed the display of textbooks used a valid *SOLO Superitem* taxonomy. But the limitations found in the research of Firmasari, Sukestiyarno, and Mariani (2013) were the lack of reliability in that book. Research conducted by Appulembang (2017) was able to express the thinking process of students in completing the *Superitem* test. The limitations of the research carried out by the Appulembang (2017) were the conventional forms of testing, which only relied on printed paper distributed to students.

Research was conducted by Henke *et al*. (2014) showed the existence of a system called *TEASE* (**T**est, **E**xamination and **A**ssessment **S**yst**e**m). It can be used to conduct a test, examination, and assessment based on the web. It can also be operated remotely or locally and can be accessed anytime and anywhere. The limitations of Henke *et al*.'s research (2014) were it didn't yet show the test instrument items in detail that were used in the *TEASE*. Research was conducted by Davutoglu *et al*. (2014) was related to remote testing systems carried out online to determine the increase in reliability and quality in testing telecommunication networks. The limitations of Davutoglu *et al*.'s research (2014) were it didn't yet show the test instrument items that were used in the remote online test system. Research was conducted by Fellin and Medicus (2015) about a new form of multiple-choice tests that was supplemented comments as feedback on test answers. The results of Fellin and Medicus's research (2015) showed new features in multiple-choice tests consisting of comments for correct or incorrect answers so that students get the right solution to the questions posed in the test. The limitations of Fellin and Medicus's research (2015) were it didn't yet show essay test items that actually can also be used to hone students' cognitive abilities in deepening the concepts they had learned. Research was conducted by Nikou and Economides (2016) about the existence of an assessment activity on motivation and acceptance of students in the learning process indoors and outdoors by using evaluation tools in the form of mobile technology. The results of Nikou and Economides's research (2016) showed that mobile technology was able to be used as an assessment tool both indoors and outdoors. The limitations of Nikou and Economides's research (2016) were it didn't yet fully show the items of instruments in the mobile technology that were used to carry out the assessment.

Research was conducted by Korkmaz and Korkmaz (2016) about the validity and reliability of the instruments was used as a measure of basic skills about electronics for electrical engineering students based on the self-efficacy scale. The research results of the Korkmaz and Korkmaz showed that there were valid and reliable instruments that can be used to determine the level of students' basic skills in the field of electronics through their self-perceptions. The limitations of Korkmaz and Korkmaz's research (2016) were it didn't yet show the instruments which were arranged in stages starting from the level of knowledge that was easily understood to the level of knowledge that was difficult.

Research was conducted by Al-Emran and Salloum (2017) about student attitudes towards the presence of mobile technology that was used to realize electronic-

evaluation. The research results of Al-Emran and Salloum (2017) showed that 99% of students already had cellphones or tablets and 94% of students had using their cellphones or tablets to evaluate the performance of their instructors at the end of each semester. The limitations of Al-Emran and Salloum's research (2017) were it didn't yet show in detail the items of the instruments were used to evaluate the instructor's performance. Research that was conducted by Thurner *et al.* (2017) showed the existence of forms of cognitive tests that were used to assess students' initial abilities, mathematics abilities, and computer skills as a basis for knowing their readiness before learning further competencies. The limitations of Thurner *et al.*'s research (2017) were it didn't yet show the calculation of validity and reliability of the items that were used to test students' initial cognitive abilities.

The results of Solihati and Mulyono's research (2018) showed the utilization of Moodle as an online testing platform that allows users to create a quiz/test design by seeing to several aspects, included: time limit, grading method, navigation method, test layouts, questions, review, and feedback layouts. The limitations of Solihati and Mulyono's research (2018) were it didn't yet show the test items which were packed by multilevel structures ranging from the lowest difficulty level to the highest difficulty level. The results of Villányi *et al.*'s research (2018) showed the utilization of tablet technology equipped with attractive illustrations and equipped with language settings makes it easier for students to assess their competencies independently. The limitations of Villányi *et al.*'s research (2018) were it didn't yet show the process of calculating the validity and reliability of the assessment instrument items.

The results of Neitola's research (2019) showed that Moodle Stack allows randomization of quiz questions at various levels ranging from the weight of questions that were easy to difficult questions. The limitations of Neitola's research (2019) were it didn't yet show the process of testing the validity and reliability of quiz question items. The results of research conducted by Sugiharni *et al.* (2018) showed the valid and reliable items that were used to evaluate the implementation of blended learning applied at STIKOM Bali in supporting Discrete Mathematics learning. The limitations of Sugiharni *et al.*'s research (2018) were it didn't yet show the evaluation instrument items which were arranged in stages starting from the highest priority level to the lowest priority level to facilitate the evaluation of each evaluation aspect.

Research conducted by Ridzuan *et al.* (2020) showed that there was a determination of the validity and reliability of the test instruments. In principle, this research has similarities with the research of Ridzuan *et al.* (2020) related to using the concept of *Superitem* in developing test instruments. The difference is in the analysis model used to determine valid and reliable tests. Ridzuan *et al.*'s research (2020) using the *Rasch* model, while this research using the *Gregory* formula to determine the validity, and the *Cronbach Alpha* formula to determine the reliability of the test instruments. Research limitations of Ridzuan *et al.* (2020) were it didn't yet show the calculation process step by step in obtaining scores of validity and reliability of the test instruments. Besides, it had not been shown directly the example of instrument items that had been adopted the *Superitem* concept. Research conducted by Hulukati (2018) showed the use of the

formula of *Pearson-Product Moment Correlation* in determining the validity of the instruments and the *Cronbach Alpha* formula in determining the reliability of the instruments. In principle, Hulukati's research (2018) has similarities with this study in the use of the *Cronbach Alpha* formula to determine the reliability of the instruments. The limitations of Hulukati's research (2018) were it didn't yet show the complete calculation process in determining the validity and reliability of the test instruments. Besides, Hulukati's research (2018) also had not shown the innovations in the form of an application that can be used to create test instruments in digital format.

Referring to the existing problems, innovations that can be realized, and the previous research that background of this research, so the authors were interested to develop a new form/design in the test instruments for the 'Assessment' course. Those instruments contain questions structure arranged by adaptation of the *Superitem* concept and are packaged in the digital format using the *Wondershare* application.

METHOD

Approach and Stages of Research

This research was development research using a quantitative approach. The quantitative approach was used as the basis for determining the validity and reliability of the instruments that had been developed. The instruments developed were in the form of test questions for the 'Assessment' course. Its content was compiled by adopting the Super item concept. Its presentation was packaged in digital form by used the *Wondershare* application. The digital test instruments were made by several development steps shown in Figure 1.

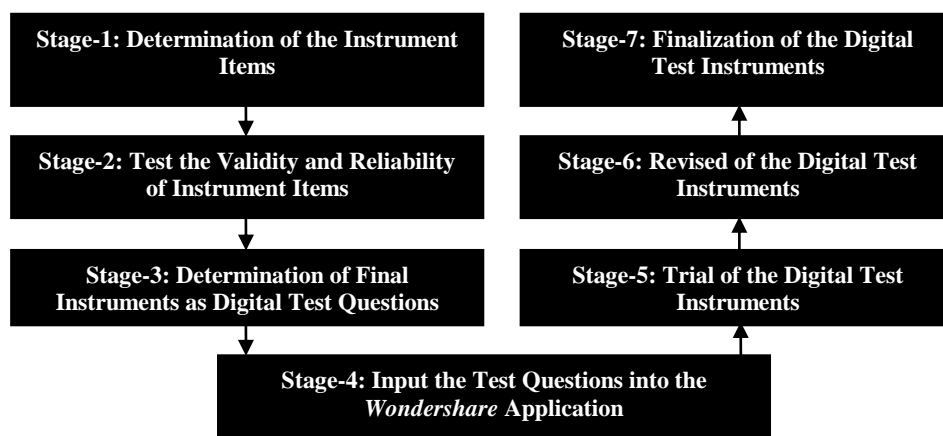|	
Stage-1: Determination of the Instrument Items	**Stage-7: Finalization of the Digital Test Instruments**
Stage-2: Test the Validity and Reliability of Instrument Items	**Stage-6: Revised of the Digital Test Instruments**
Stage-3: Determination of Final Instruments as Digital Test Questions	**Stage-5: Trial of the Digital Test Instruments**
Stage-4: Input the Test Questions into the *Wondershare* Application	

Figure 1
Development stages of the digital test instruments based on *Wondershare-Superitem*

The activity carried out at stage-1 was to determine the test instrument items for the 'Assessment' course. The instrument items refer to the material of the 'Assessment' course given to the 5th semester students in the Department of Information Technology

Education, *Universitas Pendidikan Ganesha*. Stage-2 was to test the validity of the items and test the reliability of the instruments. Stage-3 was to determine the valid instrument items so can be used as test questions. Stage-4 was the process of inserting test items into the *Wondershare* application. Stage-5 was to test digital test questions that involved 42 respondents. Stage-6 was to revise the digital test questions if there were test questions that still need to be corrected according to the respondents' suggestions. Stage-7 was to finalize the digital test questions to make sure questions ready to be used to test students' cognitive abilities.

Subject of the Research

The research subjects involved in testing the content validity of the instruments were two experts (one expert in the field of information technology and one expert in the field of educational evaluation). The research subjects involved in testing the reliability of the instruments were two lecturers in the 'Assessment' course and 40 students from the Department of Information Technology Education, *Universitas Pendidikan Ganesha*, who took the 'Assessment' course. The trials of the digital test instruments based on *Wondershare-Superitem* was conducted by two lecturers and also ten students from the Department of Information Technology Education, *Universitas Pendidikan Ganesha*, who took the 'Assessment' course.

Object, Objective, and Location of Research

The object of this research was the design/form of the digital test instruments for the 'Assessment' course based on *Wondershare-Superitem*. The objective of this research was to develop a new form in the development of test instruments that valid and reliable for the 'Assessment' course by adopting the *Superitem* concept and integrated it into the *Wondershare* application. It was used to measure the level of cognitive ability of students to understand the 'Assessment' course. The location of the application of the digital test instruments for the 'Assessment' course based on *Wondershare-Superitem* was located in the Department of Information Technology Education, *Universitas Pendidikan Ganesha*.

Data Collection Instruments

The data collection instruments used in this research were questionnaires that contained the items about the digital test instruments for the 'Assessment' course based on *Wondershare-Superitem* to be tested. Besides, the data collection tools were also in the form of photo documentation as authentic evidence of the research implementation.

Validity and Reliability of Instruments

Testing of Content validity of the test instrument items of the 'Assessment' course used the *Gregory* formula (Sugiharni *et al*., 2018). Reliability testing of instrument items used the *Cronbach Alpha* formula (Namdeo and Rout, 2016; Sugiharni *et al*., 2018). The *Gregory* formula can be seen in equation (1). The *Cronbach Alpha* formula can be seen in equation (2).

Content Validity = D/(A+B+C+D)   (1)
Notes:
D   : cell that showed valid agreement between the two experts

A   : cell that showed disagreement between the two experts
B and C   : cell that showed differences in views between experts

$$\alpha = (n/(n-1)) * (1-(\sum \sigma_i / \sigma_t)) \tag{2}$$

Notes:
α : the reliability coefficient of instruments
n   : number of items
σ_i : the variant of score item-i
σ_t : the variant of total scores

User Interface Design of *Wondershare* Application

In principle, the user interface design of the *Wondershare* application is used to show the process of entering the 'Assessment' course test instruments into the *Wondershare* application so that the test instruments become digital test instruments. The 'Assessment' course test instruments that are entered into the *Wondershare* application must be valid and reliable with the difficulty levels of the items is structured and tiered following a *Superitem* pattern. Besides, the user interface design of *Wondershare* application is also used to display some of the form's design that functions as settings and navigation which facilitate the operation of the digital test instruments for the 'Assessment' course. Several of those user interface designs, included: 1) user interface design of the process of questions randomization setting, 2) user interface design of creating process the user id and password, 3) user interface design of the maximum time-setting to answer each question, 4) user interface design of the process of adding notifications in the sound form for the correct and incorrect answers, 5) user interface design of main menu page, 6) user interface design of questions page, 7) user interface design of test results page, and 8) user interface design of review page.

Data Analysis Techniques

The scores of the results of testing the validity and reliability of the instruments were compared with the categorization scale of the test items that follow the *Guilford* pattern to determine the level of validity and reliability of the test items. The categorization scale intended can be shown in Table 1 (Prihatnawati, Amin, and Muhdhar, 2017; Putra, Husna, and Gunawan, 2018; Suciati *et al.*, 2019; Ardayati and Herlina, 2020)

Table 1
The scale of validity and reliability of test items that follow the *Guilford* pattern

Validity Category	Reliability Category	Scores Interval
Very High	Very High	$0.80 < r_{xy} \le 1.00$
High	High	$0.60 < r_{xy} \le 0.80$
Enough	Enough	$0.40 < r_{xy} \le 0.60$
Low	Low	$0.20 < r_{xy} \le 0.40$
Very Low	Very Low	$0.00 < r_{xy} \le 0.20$
Invalid	Not Reliable	$r_{xy} \le 0.00$

The analysis technique toward trial results of digital test instruments used the calculation of effectiveness percentage. Effectiveness percentage calculation produces a percentage

of assessment that was used to provide an interpretation of the trial results of the digital test instruments being developed. The formula was used for the calculation of the effectiveness percentage can be shown in equation (3) (Triani, 2019; Divayana *et al.*, 2019). Decision making to determine the effectiveness of the implementation of digital test instruments refers to the conversion of five scale effectiveness level. It can be seen in Table 2 (Sujarwo, Sukmawati, and Yahrif, 2019; Divayana, Suyasa, and Abadi, 2019).

$$P = (f/N) * 100\% \tag{3}$$

Notes:
P = Effectiveness percentage
F = the frequency of assessment result
N = the sum of the overall frequencies

Table 2
The conversion of five-scale effectiveness level

Level of Effectiveness	Effectiveness Percentage Range (%)	Decision
Excellent	90-100	Not revised
Good	80-89	Not revised
Moderate	65-79	Revised
Less	55-64	Revised
Poor	0-54	Revised

FINDINGS AND DISCUSSION

Instrument Items

The instrument items that were used as question items on the 'Assessment' course come from task questions in each chapter contained in a textbook that the title "Assessment and Evaluation" (Divayana, 2019). The full details of each item of the instruments can be seen in Table 3.

Table 3
The test instrument items of the 'assessment' course that compiled based on the *Superitem* concept

No	Chapters	Code	Test Items	Answers
1	I Basic Concepts of Tests, Measurements, Assessments, and Evaluations	C1	What is the mean about the test?	…………
		C2	What is the mean about the measurement?	…………
		C3	What is the mean about the assessment?	…………
		C4	What is the mean about the evaluation?	…………
		C5	Can you explain the differences and the relationship between tests, measurements, evaluations, and evaluations?	…………
		C6	Can you explain the difference between tests, measurements, assessments, and evaluations?	…………
		C7	Can you explain the relationship between test, measurement, assessment, and evaluation?	…………
2	II Instruments	C8	Can you explain the requirements for the cognitive test instruments that are said to be good?	…………
		C9	Can you explain the requirements of the affective measurement instruments that are said to be good?	…………

No	Chapters	Code	Test Items	Answers
		C10	Can you explain the steps for compiling cognitive test instruments?	………..
		C11	Can you explain the steps for compiling affective measurement instruments?	………..
		C12	Can you explain the steps for compiling psychomotor and affective measurement instruments?	………..
		C13	Can you make each one sample of the test instruments is reviewed based on the form of its implementation?	………..
		C14	Can you make each one sample of the affective measurement instruments in the form of questionnaires and interview guidelines?	………..
		C15	Can you make an example of affective measurement instruments in the form of questionnaires?	………..
		C16	Can you make an example of affective measurement instruments in the form of interview guides?	………..
3	III Validity and Reliability	C17	Can you determine the content validity of the instruments using the *Gregory* formula based on the available data?	………..
		C18	Can you determine the validity of the instruments using the *Product-Moment Correlation* formula based on the available data?	………..
		C19	Can you determine the validity of the instruments using the *Aiken* formula based on the available data?	………..
		C20	Can you determine the validity of the instruments using the *Product-Moment Correlation* formula and *Aiken* formula based on the available data?	………..
		C21	Can you determine the reliability of the instrument using *Cronbach Alpha* based on the available data?	………..
4	IV Measurement	C22	Can you determine the index of the difficulty of the items based on available data?	………..
		C23	Can you determine the distinguishing index based on available data?	………..
		C24	Can you determine the parameter estimates using a three-parameter logistic model based on available data?	………..
5	V Assessment	C25	Can you mention and explain the types of assessment?	………..
		C26	Can you explain the scope of the assessment aspects?	………..
		C27	Can you explain the steps for developing assessment tools?	………..
		C28	Can you determine the assessment of the score using the "Standard Reference Assessment" approach based on available data?	………..
		C29	Can you determine the assessment of the score using the "Normative Reference Assessment" approach based on available data?	………..
6	VI Types of the Evaluation Model	C30	What is the mean about the Goal-Free Evaluation model?	………..
		C31	What is the mean about the formative and summative model?	………..
		C32	What is the mean about the *CSE-UCLA* model?	………..
		C33	Mention and explain one of the evaluation models that is the best used to evaluate the 2013 curriculum?	………..
		C34	Why is the *CSE-UCLA* evaluation model suitable for evaluating the digital library program?	………..
7	VII Stages of Making the	C35	Can you determine the evaluation aspects of implementing e-learning in a college?	………..
		C36	Can you make items for evaluating the implementation of e-	………..

No	Chapters	Code	Test Items	Answers
	Evaluation Instruments		learning in a college?	
		C37	Can you determine the content validity of the evaluation instruments of the e-learning implementation?	…………
		C38	Can you determine the validity of the evaluation instrument items of the e-learning implementation?	…………
		C39	Can you determine the validation of the contents and items of the evaluation instruments of the e-learning implementation?	…………
		C40	Can you determine the reliability of the evaluation instruments of the e-learning implementation?	…………
		C41	Can you make the final evaluation instruments after obtaining valid items?	…………

Test of Validity and Reliability

The results of the content validity test toward the test instrument items were obtained from the review results of two education experts, while the reliability test results of instrument items were obtained from the assessment results of 42 respondents that consisting of lecturers and students. The complete process to obtain the results of the content validity test can be seen in Table 4 and Table 5, while the reliability test results can be seen in Table 6.

Table 4
Data of the content validity test results toward test instrument items of the 'assessment' course that compiled based on the *Superitem* concept

Expert-1		Expert-2	
Less Relevant (Score 1 - 2)	Very Relevant (Score 3 - 4)	Less Relevant (Score 1 - 2)	Very Relevant (Score 3 - 4)
C5; C12; C14; C20; C39	C1; C2; C3; C4; C6; C7; C8; C9; C10; C11; C13; C15; C16; C17; C18; C19; C21; C22; C23; C24; C25; C26; C27; C28; C29; C30; C31; C32; C33; C34; C35; C36; C37; C38; C40; C41	C5; C12; C14; C20; C39	C1; C2; C3; C4; C6; C7; C8; C9; C10; C11; C13; C15; C16; C17; C18; C19; C21; C22; C23; C24; C25; C26; C27; C28; C29; C30; C31; C32; C33; C34; C35; C36; C37; C38; C40; C41

The data of content validity test of those test instruments then entered into the form of cross-tabulation. The completely cross-tabulation data can be seen in Table 5.

Table 5
The cross tabulation results from the data of content validity test of the 'assessment' course that compiled based on the *Superitem* concept

		Expert-1	
		Less Relevant (Score 1 - 2)	Very Relevant (Score 3 - 4)
Expert-2	Less Relevant (Score 1 - 2)	A C5; C12; C14; C20; C39 (5)	B - (0)
	Very Relevant (Score 3 - 4)	C - (0)	D C1; C2; C3; C4; C6; C7; C8; C9; C10; C11; C13; C15; C16; C17; C18; C19; C21; C22; C23; C24; C25; C26; C27; C28; C29; C30; C31; C32; C33; C34; C35; C36; C37; C38; C40; C41 (36)

Based on the tabulation data shown in Table 5, then the content validity calculations were carried out on the test instrument items of the 'Assessment' course used the formula of *Gregory* that shown in equation (1). The mechanism for calculating the content validity can be explained as follows.

Content Validity = 36/(5+0+0+36)
= 36/41
= 0.878

After was founded the results of instruments content validation, the next activity that needs to be done was to test the reliability of the instruments. The data results of the reliability of the complete instruments can be shown in Table 6.

Table 6
The data of the reliability test results toward test instrument items for the 'assessment' course that compiled based on the *Superitem* concept

Items	σ_i^2	Items	σ_i^2	Items	σ_i^2	Items	σ_i^2
C1	0.249	C11	0.230	C21	0.241	C31	0.331
C2	0.249	C12	0.236	C22	0.245	C32	0.248
C3	0.230	C13	0.230	C23	0.249	C33	0.297
C4	0.343	C14	0.230	C24	0.241	C34	0.250
C5	0.222	C15	0.193	C25	0.241	C35	0.245
C6	0.245	C16	0.250	C26	0.236	C36	0.249
C7	0.236	C17	0.249	C27	0.248	C37	0.248
C8	0.245	C18	0.245	C28	0.236	C38	0.222
C9	0.249	C19	0.230	C29	0.245	C39	0.245
C10	0.245	C20	0.236	C30	0.245	C40	0.249
						C41	0.241
						$\sum \sigma_i^2$	10.093

Based on calculations using Microsoft Excel to found the reliability of the instruments, some data were obtained as follows: $\sum \sigma_i = 10.093$; n = 42; $\sigma_t = 44.347$. From that data, the calculating process of reliability can be done using the *Cronbach Alpha* formula shown in equation (2). That reliability calculation process can be seen as follows.

α = (42/(42-1)) * (1-(10.093 / 44.347))
= (42/41) * (1-0.228)
= 1.024 * 0.772
= 0.791

Final Instruments

Instruments were used for digital test items of the 'Assessment' course that were inserted into the *Wondershare* application were valid instruments and must reliable. Valid instruments were obtained from the results of expert content validation, while reliable instruments were obtained from the results of the respondents' assessment. The items of the final instrument intended, included: Item-C1, Item-C2, Item-C3, Item-C4, Item-C6, Item-C7, Item-C8, Item-C9, Item-C10, Item-C11, Item-C13, Item-C15, Item-

C16, Item-C17, Item-C18, Item-C19, Item-C21, Item-C22, Item-C23, Item-C24, Item-C25, Item -C26, Item-C27, Item-C28, Item-C29, Item-C30, Item-C31, Item-C32, Item-C33, Item-C34, Item-C35, Item-C36, Item-C37, Item-C38, Item-C40, and Item-C41.

Input the Final Instruments into the Wondershare Application

After obtaining valid and reliable final instruments, the next step was to enter those instruments into the *Wondershare* application. It can be seen through the display of user interface design in Figure 2.

Figure 2
The display of user interface design of the instrument items in the *wondershare* application

Trials the Digital Test Instruments

The trials of the digital test instruments for the 'Assessment' course based on *Wondershare-Superitem* were conducted by 12 respondents consisting of lecturers and students. The trial results can be seen in Table 7.

Table 7
The trial results data of the digital test instruments

Respondents	Item-											Effectiveness Percentage (%)
	1	2	3	4	5	6	7	8	9	10	11	
Lecturer-1	5	4	4	5	4	5	4	4	4	5	5	89.09
Lecturer-2	4	5	4	4	5	4	5	5	5	4	4	89.09
Student -1	4	4	5	5	4	5	4	4	4	4	5	87.27
Student -2	4	5	4	4	4	5	4	4	4	4	4	83.64
Student -3	5	4	4	5	4	4	5	4	4	5	4	87.27
Student -4	4	4	4	4	4	5	4	4	5	4	4	83.64
Student -5	4	4	5	5	5	4	4	5	4	4	4	87.27
Student -6	4	5	4	4	4	5	5	4	5	4	4	87.27
Student -7	5	4	4	4	4	4	5	4	4	5	4	85.45
Student -8	4	4	4	5	5	5	4	4	5	4	5	89.09
Student -9	5	5	4	4	4	4	5	4	5	5	5	90.91
Student -10	4	4	5	5	4	4	4	4	4	5	4	85.45
Average												87.12

The trial results of the digital test instruments for the 'Assessment' course had shown that digital test instruments were in the category of good effectiveness. It was evidenced by the average effectiveness percentage was 87.12%, so there was no need to make the major revision. Even so, there were some inputs/suggestions given by the respondent while the trials of test instruments. Several suggestions from respondents can be seen in Table 8.

Table 8
Some suggestions that given by respondents when testing the digital test instrument for the 'assessment' course based on *Wondershare-Superitem*

No	Respondents	Suggestions
1	Lecturer-1	Should be set so the questions can be randomized
2	Lecturer-2	Add user ID and password as access rights to enter into the form of the questions
3	Student-4	Add a facility that can be setting the time limit for each question
4	Student-8	Add facilities that can be used to display notifications by voice for the incorrect or correct answers

Revisions of the Digital Test Instruments
Based on the respondents' suggestions as shown in Table 8, several things were able to be done as a follow-up. Several revisions had been made. It can be explained as follows.

a. The revisions made to follow-up on the suggestion of lecturer-1 were randomization settings of the questions by arranging the sequence of questions that were skipped. The display of user interface design of the process of questions randomization setting can be shown in Figure 3.

Figure 3
The display of user interface design of the process of questions randomization setting

b. The revisions made to follow-up on the suggestion of lecturer-2 were the creation of a user ID and password as access rights to enter into the form of the questions. The display of the user interface design of the process for creating the user ID and password can be shown in Figure 4.

Figure 4
The display of user interface design of creating process the user id and password

c. The revisions made to follow-up on the suggestion of student-4 were the settings of maximum time to answer each question. The display of user interface design of the process of maximum time limit setting can be seen in Figure 5.
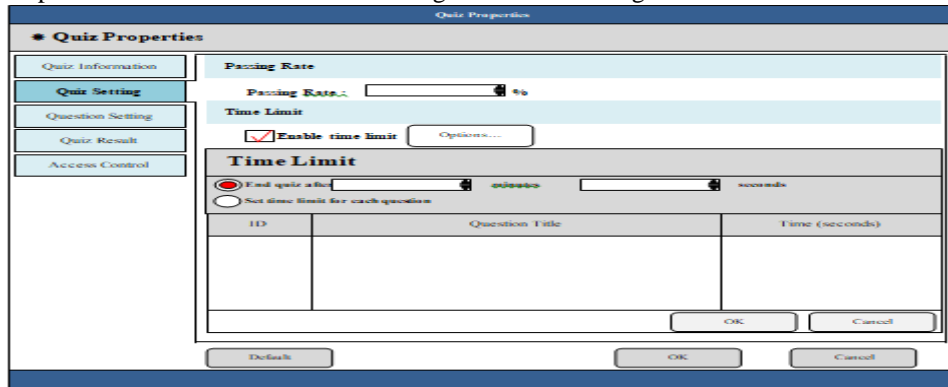

Figure 5
The display of user interface design of the maximum time-setting to answer each question

d. The revisions made to follow-up on the suggestion of student-8 were the adding notifications in the form of sound for incorrect or correct answers. The display of user interface design of the process of adding notifications in the form of sound can be seen in Figure 6.
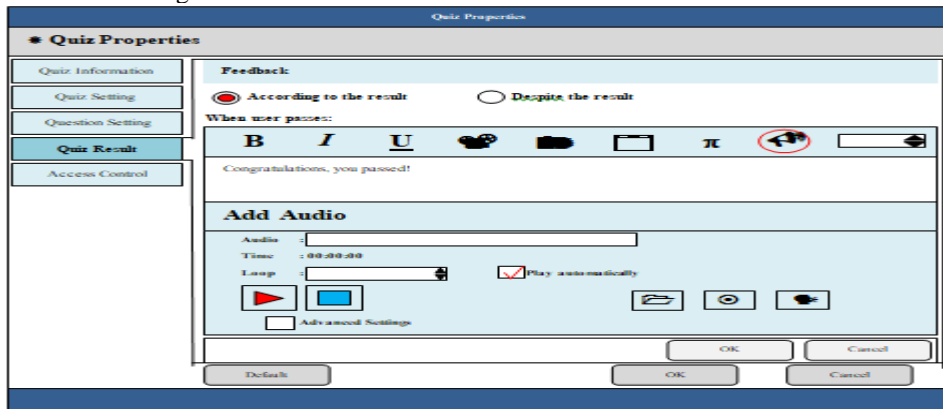

Figure 6
The display of user interface design of the process of adding notifications in the sound form for the correct and incorrect answers

The Final Digital Test Instruments

Based on the revisions that had been carried out in the previous stage, then it was obtained the final test instrument which was ready to be applied in the learning process in the Department of Information Technology Education, *Universitas Pendidikan*

Ganesha. The final display of the user interface design of the digital test instruments for the 'Assessment' course based on *Wondershare-Superitem* can be seen completely in Figure 7 to Figure 10.
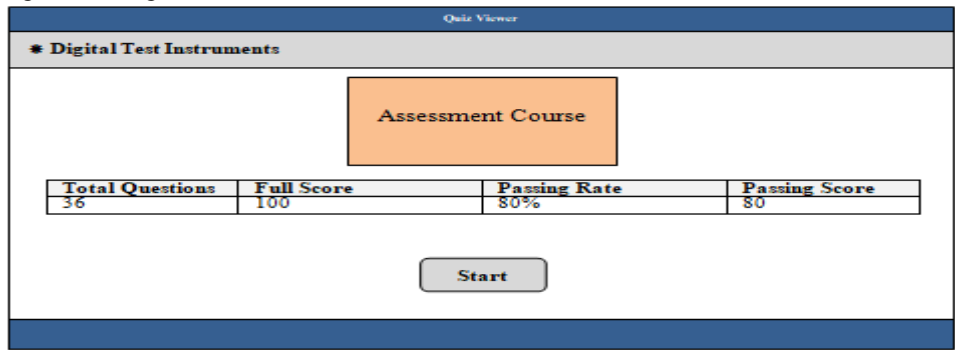

Figure 7
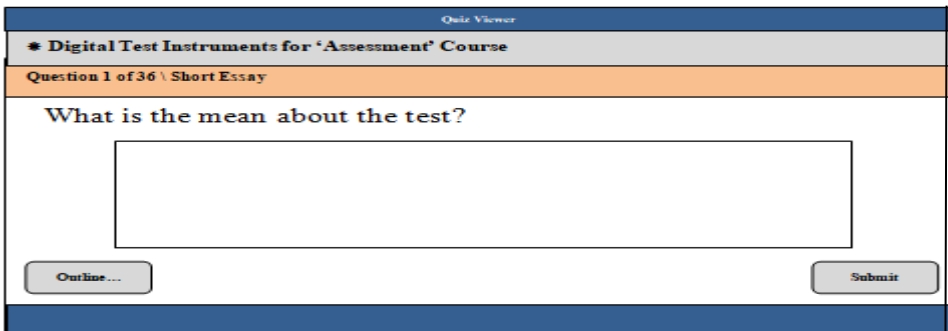The display of user interface design of main menu page

Figure 8
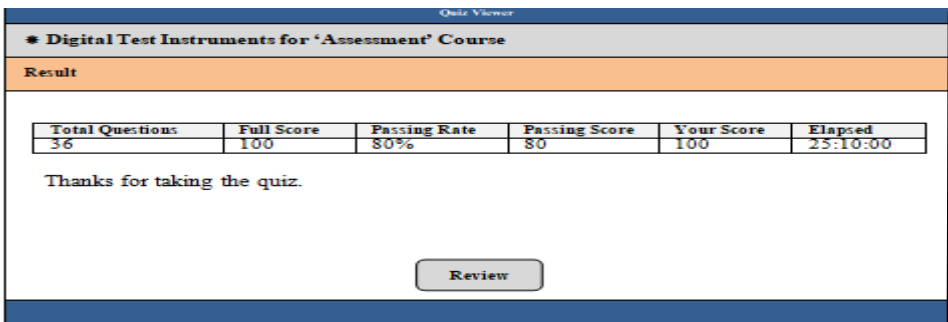The display of user interface design of questions page

Figure 9
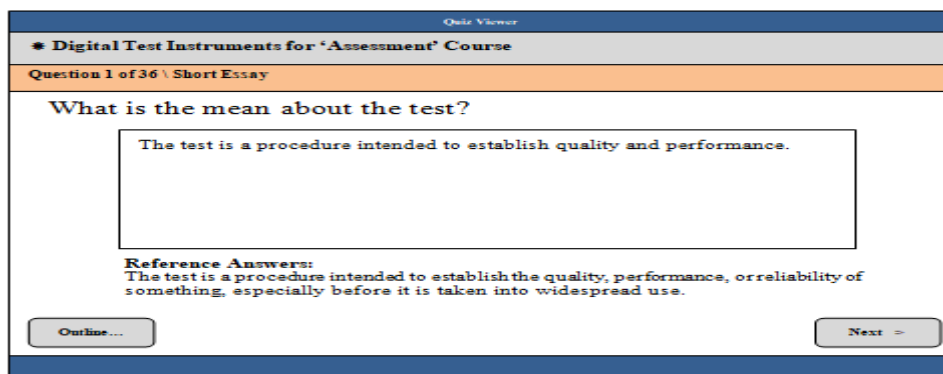The display of user interface design of test results page

Figure 10
The display of user interface design of review page

The validity test results on the item test instruments classified as a very high validity category with a validity score were 0.878, reliability test results classified as high reliability with reliability score were 0.791, and the trial results toward digital test instruments based on *Wondershare-Superitem* showed good effectiveness with average percentage were 87.12%. The validity and reliability categorization refers to the *Guilford* pattern scale (Fazlina, 2018; Suciati *et al*. 2019), while the trial results categorization refers to the five-scale effectiveness level (Cheng *et al*., 2019; Mantasiah, Yusri, and Jufri, 2020).

The novelty of this research is the existence of digital format test instruments that are created by the *Wondershare* application and the content of the questions arranged by adopt the *Superitem* concept and sourced from the material of the 'Assessment' course. Through the adopting of the *Superitem* concept in the arranged of test questions, then the test instruments can present the difficulty weight of the questions from the easiest to the hardest level. Through the *Wondershare* application, then the instruments are presented on paper can be converted into digital form so that the assessment process carried out by lecturers will be faster and easier.

This research had been able to answer the limitations of the research of Henke *et al*. (2014), research of Davutoglu *et al*.(2014), research of Nikou and Economides (2016), and research of Al-Emran and Salloum (2017), by showing in detail the items of the digital test instruments. This research also had answered the limitations of Fellin and Medicus' research (2015) by showing the essay test items.

The results of this research also had been answered the limitations of Korkmaz and Korkmaz's research (2016), Solihati and Mulyono's research (2018), and research of Sugiharni *et al*. (2018) by showing the instrument items which was packaged in stages starting from the easy level to the hardest level by adopting the *Superitem* concept. Furthermore, the limitations of Thurner *et al*.'s research (2017), Villányi *et al*.'s research (2018), Hulukati's research (2018), Neitola's research (2019), and Ridzuan *et al*. (2020) were also had answered through this research by showing detailed calculation processes to determine the validity and reliability of the items.

The limitations found in this research were the form of matchmaking questions had not been shown so that there were no variations in the questions presented in the test instruments. Other limitations found in this research were digital test instruments had not been made in the online format. Students only can answer the test in the form of stand-alone through their personal computer/laptop if available the *Wondershare* application.

CONCLUSION

This research has succeeded in realizing a new form of test instrument development. The new form is in the form of valid and reliable digital test instruments for the 'Assessment' course. The instrument items are arranged by adopting the *Superitem* concept and are packaged in digital format through the *Wondershare* application. The presence of these test instrument forms can stimulate students' thinking and logic in a structured manner starting from the basic level to the high level related to the assessment material. This instrument form can support the success of the implementation of distance learning in the 'Assessment' course especially viewed from the cognitive domain. Opportunities for further research that can be done to solve the limitations in this research are by adding items in the form of matchmaking questions, so there are form variations of digital test questions. Besides, digital test instruments need to be developed online so that tests can be carried out within the campus or outside the campus.

REFERENCES

Al-Emran, M., & Salloum, S.A. (2017). Students' attitudes towards the use of mobile technologies in e-evaluation. *International Journal of Interactive Mobile Technologies*, *11*(5), 195-202. https://doi.org/10.3991/ijim.v11i5.6879

Appulembang, O.D. (2017). A profile of guided algebra problem solving using the SOLO taxonomy and the cognitive conceptual tempo style of students at the SMA Negeri 1 school in Makale, Tana Toraja. *A Journal of Language, Literature, Culture, and Education POLYGLOT*, *13*(2), 133-149. https://dx.doi.org/10.19166/pji.v13i2.336

Ardayati, & Herlina. (2020). Teaching reading comprehension by using herringbone technique to the eighth-grade students of SMP Negeri 11 Lubuklinggau. *Journal of English Education, Literature and Linguistics*, *3*(1), 79-85. https://doi.org/10.31540/jeell.v3i1.884

Cheng, Q., Kularatna, S., Lee, X.J., Graves, N., & Pacella, R.E. (2019). Comparison of EQ-5D-5L and SPVU-5D for measuring quality of life in patients with Venous Leg Ulcers in an Australian Setting. *Quality of Life Research*, *28*, 1-9. https://dx.doi.org/10.1007/s11136-019-02128-6

Davutoglu, C., Durakbasa, N.M., Bas, G., & Guclu, E. (2014). Improved quality and reliability in telecommunication network testing by means of a remote online test system. *International Journal of Online and Biomedical Engineering*, *10*(4), 66-69. https://dx.doi.org/10.3991/ijoe.v10i4.3830

Divayana, D.G.H. (2019). The implementation of blended learning with Kelase platform in the learning of assessment and evaluation course. *International Journal of Emerging Technologies in Learning*, *14*(17), 114-132. https://doi.org/10.3991/ijet.v14i17.8308

Divayana, D.G.H., Suyasa, P.W.A., & Abadi, I.B.G.S. (2019). Digital library evaluation application based on combination of CSE-UCLA with weighted product. *Journal of Engineering and Applied Sciences*, *14*(4), 1318-1330. https://dx.doi.org/10.36478/jeasci.2019.1318.1330

Divayana, D.G.H., Suyasa, P.W.A., Ariawan, I.P.W., Mahendra, I.W.E., & Sugiharni, G.A.D. (2019). The design of digital book content for assessment and evaluation courses by adopting superitem concept based on Kvisoft Flipbook Maker in Era of Industry 4.0. *Journal of Physics: Conference Series*, 1165, 1-6. https://dx.doi.org/10.1088/1742-6596/1165/1/012020

Fazlina, A. (2018). An analysis of college entrance test. *English Education Journal (EEJ)*, *9*(2), 192-215. http://jurnal.unsyiah.ac.id/EEJ/article/view/11528

Fellin, W., & Medicus, G. (2015). Multiple choice tests: More than a time saver for teachers. *International Journal of Engineering Pedagogy*, *5*(3), 15-24. https://dx.doi.org/10.3991/ijep.v5i3.4376

Firmasari, S., Sukestiyarno, Y.L., & Mariani, S.C. (2013). The development of teaching materials using taxonomy of solo superitem with wingeom-assisted peer tutors. *Unnes Journal of Mathematics Education Research*, *2*(1), 184-189. https://journal.unnes.ac.id/sju/index.php/ujmer/article/view/1244

Henke, K., Debes, K., Wuttke, H.D., & Katzmann, A. (2014). Mobile assessment tools. *International Journal of Recent Contributions from Engineering, Science & IT*, *2*(3), 9-14. https://dx.doi.org/10.3991/ijes.v2i3.3816

Hulukati, E. (2018). The effect of guided discovery learning model with superitem test on students' problem-solving ability in mathematics. *Journal of Social Science Studies*, *5*(2), 210-219. https://doi.org/10.5296/jsss.v5i2.13406

Korkmaz, Ö., & Korkmaz, M.K. (2016). A validity and reliability study of the basic electronics skills self-efficacy scale (BESS). *International Journal of Engineering Pedagogy*, *6*(4), 30-35. https://doi.org/10.3991/ijep.v6i4.6168

Mantasiah, R., Yusri, & Jufri. (2020). Semantic feature analysis model: Linguistics approach in foreign language learning material development. *International Journal of Instruction*, *13*(1), 185-196. http://www.e-iji.net/dosyalar/iji_2020_1_12.pdf

Namdeo, S.K., & Rout, S.D. (2016). Calculating and interpreting Cronbach's alpha using Rosenberg assessment scale on paediatrician's attitude and perception on self esteem. *International Journal of Community Medicine and Public Health*, *3*(6), 1371-1374. http://dx.doi.org/10.18203/2394-6040.ijcmph20161448

Neitola, M. (2019). Circuit theory e-assessment realized in an open-source learning environment. *International Journal of Engineering Pedagogy*, *9*(1), 4-18. https://doi.org/10.3991/ijep.v9i1.9072

Nikou, S.A., & Economides, A.A. (2016). An outdoor mobile-based assessment activity: Measuring students' motivation and acceptance. *International Journal of Interactive Mobile Technologies*, *10*(4), 11-17. https://dx.doi.org/10.3991/ijim.v10i4.5541

Prihatnawati, Y., Amin, M., & Muhdhar, M.H.I.A. (2017). The effect of module implementation with STAD cooperative learning toward process skills in science and cognitive achievement of 8th grade students. *Advances in Social Science, Education and Humanities Research*, *164*, 111-116. https://doi.org/10.2991/icli-17.2018.22

Putra, I.D.G.T.T., Husna, I.E., & Gunawan, K.I.A. (2018). Computer-based application as a social media to improve safety traffic knowledge of junior high school students (Case Study of SMP N 1, SMP N 4 and SMP N 5 Ciamis). *International Conference on Technology for Sustainable Development 2018, KnE Social Sciences*, 2019, 114-125. https://dx.doi.org/10.18502/kss.v3i23.5142

Ridzuan, M.F., Lian, L.H., Fozee, F.A.A., & Nasser, S.N.A.M. (2020). Rasch analysis model: Reliability and validity of Superitem Test Instrument. *International Journal of Academic Research in Progressive Education and Development*, *9*(4), 1-11. https://dx.doi.org/10.6007/IJARPED/v9-i4/8166

Solihati, N., & Mulyono, H. (2018). Designing and evaluating the use of smartphones to facilitate online testing in second-language teacher education (SLTE): An Auto-ethnographic study. *International Journal of Emerging Technologies in Learning*, *13*(1), 124-137. https://doi.org/10.3991/ijet.v13i01.7683

Suciati, Munadi, S., Sugiman, & Febriyanti, W.D.R. (2019). Design and validation of mathematical literacy instruments for assessment for learning in Indonesia. *European Journal of Educational Research*, *9*(2),865-875. https://dx.doi.org/10.12973/eu-jer.9.2.865

Sugiharni, G.A.D., Setiasih, N.W., Mahendra, I.W.E., Ardana, I.M., & Divayana, D.G.H. (2018). Development of Alkin Model Instruments as evaluation tools of blended learning implementation in discrete mathematics course on STIKOM Bali. *Journal of Theoretical and Applied Information Technology*, *96*(17), 5803-5818. http://www.jatit.org/volumes/Vol96No17/15Vol96No17.pdf

Sujarwo, Sukmawati, & Yahrif, M. (2019). Improving students' english learning outcomes through PQ4R (Preview, Question, Read, Reflect, Recite, Review) Learning Model at the eighth grade students of SMP Tunas Bangsa Makassar. *Celebes Education Review*, *1*(2), 48-55. https://doi.org/10.37541/cer.v1i2.228

Thurner, V., Zehetmeier, D., Hammer, S., & Böttcher, A. (2017). Developing a test for assessing incoming students' cognitive competences. *International Journal of Engineering Pedagogy*, *7*(4), 35-50. https://doi.org/10.3991/ijep.v7i4.7433

Triani, D.A. (2019). Mathematical teaching development with contextual approaches to improve student learning outcomes. *Factor M: Focus Action of Research Mathematic*, *2*(1), 51-65. https://doi.org/10.30762/f_m.v2i1.1690

Villányi, D., Martin, R., Sonnleitner, P., Siry, C., & Fischbach, A. (2018). A tablet-computer-based tool to facilitate accurate self-assessments in third and fourth-graders. *International Journal of Emerging Technologies in Learning*, *13*(10), 225-251. https://doi.org/10.3991/ijet.v13i10.8876