Tersedia secara online
http://journal.um.ac.id/index.php/jptpp/
EISSN: 2502-471X
DOAJ-SHERPA/RoMEO-Google Scholar-IPI

Jurnal Pendidikan:
Teori, Penelitian, dan Pengembangan
Volume: 5 Nomor: 11 Bulan November Tahun 2020
Halaman: 1540—1549

Computerized Dynamic Assessment (C-DA) on Reading Comprehension for L2 Learners of Vocational High Schools

Putri Ismayana[1], Gunadi Harry Sulistyo[2], Primardiana Hermilia Wijayati[3]

[1]Keguruan Bahasa-Universitas Negeri Malang
[2]Pendidikan Bahasa Inggris-Universitas Negeri Malang
[3]Pendidikan Bahasa Jerman-Universitas Negeri Malang

INFO ARTIKEL

Riwayat Artikel:

Diterima: 11-07-2019
Disetujui: 10-11-2020

Kata kunci:

computerized dynamic assessment;
reading comprehension;
vocational high school student

ABSTRAK

Abstract: This study focuses on developing a prototype of an assessment program on reading comprehension based on computerized dynamic assessment. The reading skills to be measure include identifying topic, main idea, the detail of the text, logical inference, an assumption, word meaning and synonym, and a conclusion. The assessment consists of two types of tests including multiple choice type and cloze procedures. Those tests contain prompts as the characteristic of dynamic assessment. The participants in this study were 316 eleventh grade students of vocational high schools. The result reveals that the product was positively agreed by most of the subjects despite the fact that they were not familiar with this assessment program. This indicates that the developed product was acceptable by eleventh grade students on vocational high schools.

Abstrak: Penelitian ini fokus pada pengembangan prototipe program penilaian pada membaca komprehensif yang berdasarkan pada *computerized dynamic assessment*. Kemampuan membaca yang akan diukur pada tes ini, meliputi identifikasi topik, ide pokok, informasi rinci dalam teks, referensi, identifikasi anggapan, memahami makna persamaan kata, dan menyimpulkan. Produk asesmen ini terdiri dari dua tipe teS, yaitu pilihan ganda dan *cloze procedures*. Tes tersebut mengandung petunjuk/saran sebagai karakteristik dari *dynamic assessment*. Peserta yang terlibat pada penelitian ini terdiri dari 316 siswa kelas sebelas sekolah menengah kejuruan. Hasil dari pengembangan produk program penilaian *computerized dynamic assessment* sangat diterima dengan baik dan hampir dari seluruh peserta tidak mengenal program penilaian *computerized dynamic assessment.* Hal ini mengindikasikan bahwa pengembangan produk ini diterima oleh siswa terutama pada sekolah menengah kejuruan.

AlamatKorespondensi:

Putri Ismayana
Keguruan Bahasa
Universitas Negeri Malang
Jalan Semarang 5 Malang
E-mail: pardjomo.puput@gmail.com

Mastery levels of reading comprehension are different (Pourhosein Gilakjani & Sabouri, 2016). Therefore, to find out the level of comprehension in reading, a test is designed to show this competence. Reading assessment designs to reveal students' reading competence may consider the taxonomy of tasks as Brown (2007) divided. These are perceptive reading (recognition of symbols, letter, words), selective reading (focus on morphology, grammar, lexicon), interactive reading (discourse level cloze tasks, reading and comprehension questions, short answer response to reading, discourse editing tasks, scanning, re-ordering sequence of sentence and responding to charts, maps, graphs, diagrams) and extensive reading (skimming, summarizing, responding, to reading through the short essay, note taking, marginal notes, highlighting and outlining). This taxonomy could be used as the guideline for the teacher on designing reading tasks based on students' level. Nilson as attuned by San (2019) suggested teachers to use Bloom taxonomy because the design is appropriately scaffold of the questions development from the basic knowledge (remembering fact) to advanced skill (evaluating and creating). Meanwhile, designing a reading comprehension assessment is quite complex because the task in reading comprehension has some aspects. These are characters, times and events, events and sequences, syntactic structure, the connection between parts of the text, inferences, text sensitivity, text hierarchy, mental model, text flexibility, errors and inconsistencies (Meneghetti, Carretti, & De Beni, 2006). Therefore, reading comprehension assessment needs to be designed by considering those aspects.

Some researches have been carried out on designing a reading comprehension assessment to find out students' ability in reading comprehension, including traditional and alternative assessment. Some of them used a computer-based assessment, while the rest used a paper-based assessment. A research by Wightman and Roney (2013) used performance assessment to discover the students' ability in reading comprehension. There were three treatments conducted in this research including

storytelling, reading aloud, and silent reading. The Silent Reading served as the baseline treatment level to which the two other treatments compared. It was to figure out if Story Telling and Reading Aloud had any significant effect on the narrative and expository comprehension and oral reading fluency of these students compared to the effect of Silent Reading had on. Each performance on every treatment was followed up by the administration of comprehension assessment consisting of eight questions in multiple-choice with a text employed for each. The materials included twenty-eight stories of narrative and expository texts. The treatment was conducted over a seven-week period, and reading comprehension was tested at the end of the treatment period. The study revealed that both reading aloud and telling stories by a teacher did significantly affect students' comprehension of the narrative text as compared to simply having the students read stories silently. That result of the study was obtained in the long period treatment as Wightman and Roney (2013) said the more extensive treatment is more likely to yield significant result. It means the performance assessment study considers the long treatment period to get the significant result.

Another study to find out reading comprehension is using Scenario-Based Assessment as conducted by Sabatini, O'Reilly, Halderman, and Bruce (2014). This study was conducted to high school students while the reading comprehension was measured using a scenario-based assessment approach, which required students to read a range of sources to fulfill a particular reading goal. The results indicated that students, including struggling readers, were able to read, understand, and solve the problem found in complex learning environments. However, students' ability to do so was often tempered by their basic reading skills. The result of this assessment was not only to get the final score but also to reveal the problem, like test anxiety on reading comprehension the students had. Even though the result was good, the time spent in this research was quite long which lasted from the spring of 2012 to the fall of 2012, so that it led11 subjects of the research were not able to complete the test.

A reading comprehension test was also developed in an online reading summative model by Sofa (2016). In this study, the online summative reading test specified three aspects, namely source, content, and format. The task in this assessment model was multiple-choice and looked like the paper-and-pencil test so the students did not find it difficult in reading the passages and the questions since they could just scroll it up and down. This model also helped the teacher score the results since they could be seen once the test was completed; thus, it was not time-consuming. In addition, there are similar computer-based assessments available either online or not, in the market and on the internet. Therefore, teachers can buy and accessed easily. However, the appropriateness is still questionable since by having it teachers might not be able to set the criteria based on the syllabus and the need of the actual students in the classrooms and curriculum. Out of the weakness of the product, this test only consists of 25 items in the item bank. The reading level of the test item is not in the precise percentage, and that has no construct validity process to reveal the psychological quality of the students. However, the product of the online reading summative test in this study shows that the test is not time-consuming in administration. The test is also practical and cost-effective because there is no need for paper. Moreover, it is also an effective test especially in terms of time. It is because the computer will show the scores automatically instead of manually being scored. From the strengths and weaknesses of the study, the result indicates that the online summative reading test for students of reading comprehension 2 can replace the previous conventional, impractical, and time-consuming tests.

Regarding the earlier studies, it seems that the development of reading comprehension test has been changed from the paper-based to the computer-based. The reading comprehension test development above has been conducted to enable the test can be tested in clear results, elaborating the students' comprehension skill, and be effective and efficient in time, cost, and practice. The research has also shown that the test in computer-based is more effective and efficient about the time, cost, and practice. Davidson, Biancarosa, Carlson, Seipel, and Liu (2018) stated that using computer-administrated in reading comprehension tests can provide reliable scores which indicate the type of errors which the students are prone also the information of comprehension rate, and potential indicator. Moreover, the result presented reliability data for correct responses and incorrect responses as well as construct validity data on correct responses. Furthermore, by using the test, the goal would be automatically shown as evidence after a sufficient level of accuracy has been achieved. Even though the research only revealed the final score, even the scenario-based assessment conveyed the problem in reading comprehension, it spent more time. Regarding the result of the previous studies, developing the effective and efficient reading comprehension assessment is needed not only to measure the final score of the test but also to solve the problems the students have in reading comprehension.

Computer-based test (CBT) has also been introduced at schools today based on the static assessment (Sofa, 2016). This provision is based on the government regulations that the final test should be computerized, *Ujian Nasional Berbasis Komputer* (UNBK). In order to attune to this situation, some students could practice using CBT developed by developers and publishers. As explained in the previous research, there are some developments of computer-based assessments. Some of them are to reveal reading comprehension in a computer-based test which is comparable to the paper-based test. Reading assessment types explained above focus on the assessment of the product of the test taker, i.e. reading comprehension ability. Some of them use qualitative instrument to assess students' reading comprehension and some use quantitative ones. Moreover, teacher-made tests tend to be about the grading system rather than promote learning (Naeini & Duvall, 2012). However, the process attempted by the students to complete the assessment was ignored. Now, the assessment must be able to observe students' behavior and potential in reading. Sainsbury (2006) stated that the new models of assessment potentially allow us to see children operate at full potential, be engaged in tasks which command their commitment. Based on those reasons above, it is crucial to develop a

particular model of a reading comprehension assessment which covers product, process, and learning potentials of students instead of just covering students reading ability.

There are some studies on reading comprehension assessment which have been already conducted to show not only the final score of the students, but also covered the product, process, and learning potential of students on high level reading ability and low level of reading ability. The first study was conducted by Naeini and Duvall (2012). This study aimed to reveal the improvements in English Language Training (ELT) university students' reading comprehension performance by applying the mediations of a dynamic assessment approach to instruction and assessment which uses the sandwich method of dynamic assessment. This study used a mixed method study. The result indicated that using dynamic assessment offers a chance for language teachers to measure more accurately a student's level of understanding and awareness and thereby determine the target to promote the level of development of the student in relation to their current level of independent and assisted performance.

Shabani (2012) conducted a study on dynamic assessment to explore the feasibility of computerized dynamic assessment (C-DA) in the context of reading comprehension and, more precisely, the effect of electronically delivering textual and visual scaffolding on L2 readers' comprehension processes. The procedure of this study included a short reading text along with its manipulated version and visual prompts which were gradually offered upon the students' failure to provide the correct answer. Afterward, the result demonstrated that this C-DA procedure could discriminate among low-achieving students with reference to their responsiveness to electronic mediation and diagnose quite vividly their underlying abilities in terms of both independent and assisted cognitive functioning. Moreover, the results showed the benefit for the later decision-making processes, like placement and selection as they provided valuable diagnostic information about learners' abilities in terms of both independent and assisted performance functioning.

The effect of using computerized dynamic assessment was also studied by Ebadi and Saeedian (2015). The study was about a process-oriented study of reading comprehension aims at investigating the impacts of applying computerized dynamic assessment (C-DA). The analysis of the results showed that a pretest (unmediated) score was a sufficient indication neither for measuring individuals' ability nor for preparing an effective lesson plan for them. This is related to the proficiency level, even when two learners earned the same score in pretest, it could not necessarily be interpreted as having the same proficiency level (Ebadi & Saeedian, 2015). This proficiency level can reveal once they took the part in the C-DA because the assessment was not only to reveal the final score but also the learning potential from the mediation given. In other words, the subtle distinctions of the learning potentials appeared only after the amount of mediation in different skills was specified by taking the C-DA. Based on the understanding, the learning potential of each individual can lead to a more comprehensive lesson plan.

From the research described above, it is quite possible to develop C-DA at school especially for eleventh grade. This is related to the integration of ICTs (Information and Communication Technology) into educational classroom teaching attaining high attention today (Wilson, Tete-Mensah, & Boateng, 2014). In addition, technology use is common among the students' life, as Gallardo (2016) said that the students use various and different tools (apps) and feel comfortable using technology as a communication means and learning media. Therefore, the development of the C-DA test suits students' life-style nowadays called digital generation. Moreover, the form of C-DA test also relates to the characteristic of the high school students' intellectual development. Meanwhile, Sciarra (2004)stated that the eleventh grade students should have the ability to refine their goal through related information and sources by interacting with others. In academic contexts, Echternacht (1976) stated that vocational high school students show lower score than senior high school students but the formers achieve higher school grade and focused on the work world. Based on that characteristics explained, I believe that C-DA is appropriate for the eleventh grade students of vocational high schools since it is an interactive assessment which constructs their knowledge upon the intervention.

The needs assessment revealed that the teachers have to leave the students who cannot answer the test due to the tight schedule to teach the other materials. Thus, the students have no chance to improve their skill on those given tests. On the other hand, this condition is unfair for the low-level students because they have no chance to overcome the material given. It means that the previous test did not cover all the students from high to low levels of reading comprehension. Moreover, the discussion for the answers is rarely given by the teacher. As a result, they do not understand the reason for the score they obtained. Another fact retrieved from the interview with the teacher and student is the availability of a computer test for the school final exam.

Meanwhile, the exercise conducted in the class is a paper-based test while the final test is conducted in the computer-based test. It means the computer-based test is not common to be used in the classroom and is quite different from the condition and class' practices which only use computer-based test in the final test. Additionally, the preliminary study also indicates that most students prefer computer-based tests to paper-based test and in online to offline. Because of these reasons, it needs to carry out the computer-based assessment containing guidance and discussion which can be used as a daily assessment or called computerized dynamic assessment.

METHOD

This research is conducted to develop an assessment product. In line with this, the research design uses the Research and Development method. As the R&D tenet, the finding of the R&D method is used to design new products and procedure then it is systematically tested, evaluated, and refined until it gets the specified criteria of effectiveness, quality, or similar standards (Gall, Gall, & Borg, 2003). The result of this study is to develop the computer-based dynamic assessment instrument

on reading comprehension for the eleventh grade students. To gain the purpose, it needs a research approach that highlights an effort to produce the product based on the dynamic assessment. The rationale of using the R&D design is because of the steps for development and validation of the educational product related to the steps on computer-based assessment product which accomplishes the previous test. Meanwhile, one of R & D types is the applied research aiming to find the product which fits with the situation and need in the classroom which can help the teacher set up the product used in the classroom. Therefore, this study is conducted in R & D research design expecting that the computerized dynamic assessment product can solve the problem regarding the previous test.

The development model of computerized dynamic assessment should be chosen while the design of the research has been elaborated above. The model of development also took an important role since this research is dealing with the development of test instruments which have to be carried out in certain procedures. The design of the test development model is adopted from Sulistyo (2016) which is divided into conceptual and empirical steps. Sulistyo (2016) stated that the conceptual step is conducted to identify the objective of the test. This step includes that establishment of the test blueprint, review from the experts, item writing, and item review. Meanwhile, the empirical step is the tryout after item revision. The results of this step are responses from the subjects to task which are used to examine internal attributes of the test including reliability, empirical/construct validity, and item analysis.

The adaptation was carried out by modifying the steps of the process based on the objective of the research as well as the consideration of time and practicality constraints by eliminating some stages and adding several stages. After the adaptations, the model of the dynamic test development was used as follows: conducting needs assessment, creating content specification/test blueprint, blueprint expert review, developing the prototype, prototype review, field test, item analysis, test assembly, expert review, tryout, final product. The needs assessment in this research was conducted in 2 vocational high schools involving students and English teachers. The result of needs assessment was then used to develop the product related to the students need. The next step is creating test blueprint which adapted from the syllabus. In the blueprint, the reading skills to be assessed include identifying topic, main idea, the detail information of the text, logical inference, assumption, word meaning and synonym, and a conclusion.

The next stage was test blueprint review by the two test experts from Universitas Negeri Malang. The aspects reviewed included the indicators, items, amount of each item from indicator, reading level, and the time allocation. The instrument used to review the blueprint was a set of checklist responses to indicate the appropriateness of the test blueprint. After the test blueprint was reviewed by the test expert, the prototype item was developed. The prototype item developed in this product was based on the indicators stated in the test blueprint. Before the prototype was tested, the prototype was reviewed by the test experts to know the appropriateness of the test with the set of criteria created by the developer. Those criteria include the items based on the objective of the course, there is only one correct answer, the phrase clearly states the problem, only a single idea in each item, the distractors are at least plausible, incorporate common errors in distractors, the position of correct answer should be put randomly, each item includes three to five options, avoid overlapping alternatives, the option length should be about the same, no grammatical clues to the correct answer, list the choices vertically, the intended response options, avoid negative words, avoid excessive use of a negative or double, avoid the excessive use of “all of the above” or “none of the above”.

After being reviewed, the prototype was then tested to the eleventh grade students of vocational high school consisting of 232 students. The result of the field test was then analyzed by using ITEMAN application in item analysis stage to know the reliability of the items in multiple-choice types including item difficulty, item discrimination, and the item which has the suggested answer. The Gunning Fog index was also used to see the readability of the passages especially the passage on cloze procedure. After all the steps were accomplished, the test was then installed to the C-DA program and reviewed by the test experts. The experts consisted of two test experts and one ICT expert. In this stage, the test experts reviewed the blueprint and the test of the final product. The set of criteria used were similar with the criteria of the prototype item. Meanwhile, the ICT expert reviewed the program of the test including the element of design and text. The product was then tried out to know the work and appropriateness of the product. The subjects of this stage were84 eleventh students of vocational high school. The result of all steps in development product was the final product of C-DA on reading comprehension.

The instructional materials used in this study are those from semester 1 of the syllabus for the eleventh grade students. The theme used is formal invitation text. The micro skill of reading based on syllabus include identify topics, the main idea, the detail of the text, logical inference, assumption, word meaning and synonym, and a conclusion. The material includes texts consisting of items in the form of multiple-choice type and cloze procedures. The multiple-choice type consists of 20 items and 1 passage with 20 blanks for cloze procedure. For multiple-choice, every item consists of one right answer and four distractors while the cloze procedure consists of 21 words available in the word bank. Each item of the test consists of 3 hints in the form of instruction followed with text highlighted.

This project involved 316 students of vocational high schools of SMKN 2 Malang and SMKN 5 Malang. The subjects were eleventh grade students of vocational high school. The eleventh grade students were chosen because, in this level, the students start to make group of friend and related of that the vocational high school students tend to solve problems with others’ support (Placklé et al., 2014).The students were at an intermediate level meaning that the students were senior high school sophomore.

The researcher made C-DA program to measure the students' reading comprehension ability. In the multiple-choice type, the C-DA was programmed to present the hint starting from the implicit to explicit instruction followed with the highlighted text. Meanwhile in the cloze procedure, the hint was in the form of an instruction text. The step to administer the C-DA developed in the present study was as follows: first, the students had to log-in to the program by using the username and password from the administrator. The test administration is implied in figure 1 and 2.

The characteristic of the dynamic assessment is an interventionist provided in the assessment. Interventionist is the mediation provided during an assessment which is used in computer-based assessment and is well-adapted to large-scale assessment (Thouësny, 2010). The mediation can be prompts/hints with different step starting from implicit to explicit form of instruction. The different steps provide three hints shown in the table 1 and 2.

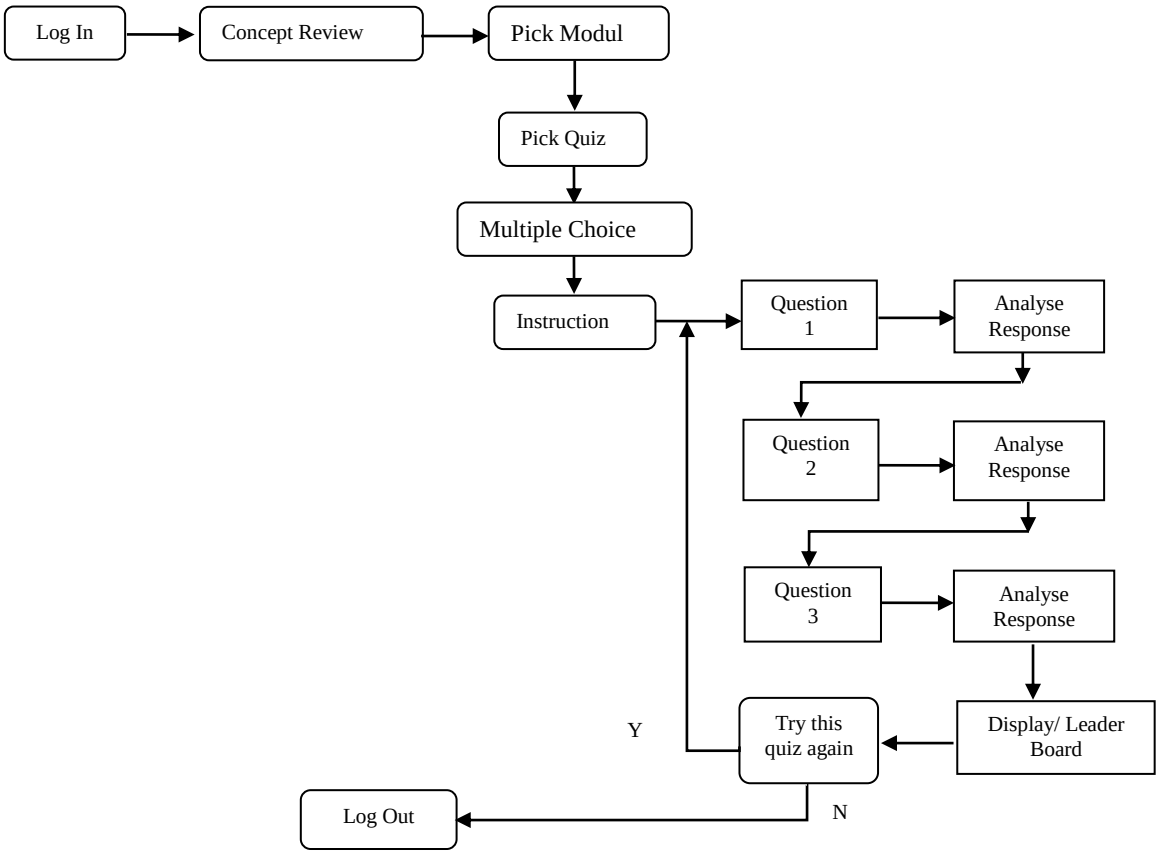

Figure 1. Main Program Flow of Multiple-Choice Type

Table 1. Regulatory Scale for Multiple Choice Types (implicit to explicit) adapted from Shabani (Shabani, 2012)

Level	Prompts/hints
Implicit ↓ Explicit	1. Please read the text again 2. Read the text again but pay more attention in the paragraph highlighted 3. Think about which row are most relevant to the ones you are trying to complete

Table 2. Regulatory Scales for Cloze Procedures (Implicit to explicit)

Level	Prompts/hints
Implicit ↓ Explicit	1. Try again 2. This is kind of .... (singular or plural noun/modal/preposition/conjunction/expression) 3. This is thesynonym of .... (singular or plural noun/modal/preposition/conjunction/expression)

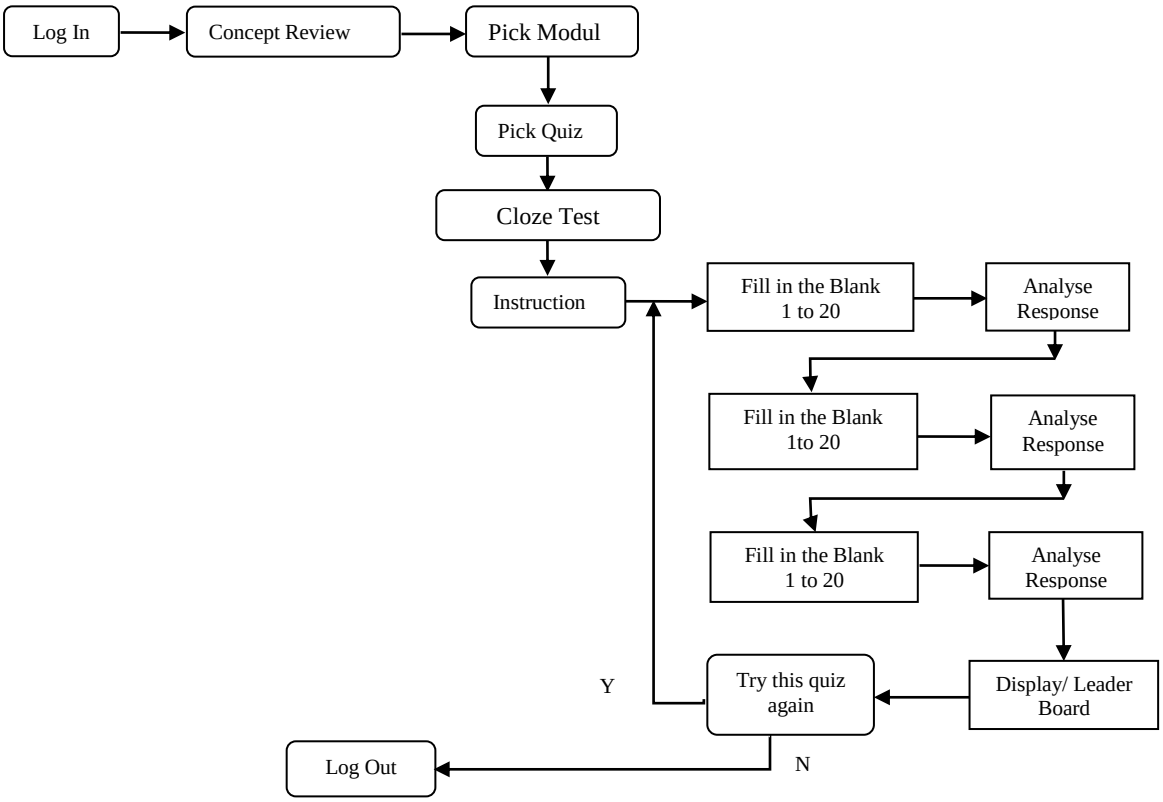

Figure 2. Main Program Flow of Cloze Procedures

The scoring process in this product was calculated by counting the number of hints or prompts offered to help students to arrive at the right answer called ZPD (zone of proximal development) scores. The score system adapted the assumption the more student asked for the clues/hints, the lower their ZPD score would be. This understanding was came from Vygotsky in which children with the same test score and age have different understanding (Shabani, 2012). Therefore, the product and scoring system in this product were developed based on Vygotsky's theory. The scoring system of the product is as follow, if the students answer the question correctly without any clue/hint, he/she will get 4 point for that item. Later, if the answer is incorrect and they need clue to get the correct one, the score will decrease to 3 and so on until the correct answer is revealed. The 0 score were given if the answer was wrong and the clues run out. The score will gradually decrease with more hints that students ask.

RESULTS

Blueprint Review

Validation of the blueprint is the process of evaluating the appropriateness of the indicators synthesized from the syllabus of English in eleventh grade, the reading level suitability, the number of items developed, and the proportion of the items.

There were two test experts reviewing the test blueprint. The validation process of the blueprint from the Expert 1 was carried out through several parts to make the blueprint perfect. After revision, the blueprint was ready to be used to the next stage, i.e. writing the test items or prototype. In the discussion, the advice from expert 1 was about the word banks in the cloze test for the students to choose from. Meanwhile, the validation result from Expert 2 indicates that all the criteria are suitable or good. The only suggestion from her was about to re-check the level and percentage of the item in multiple-choice. The result of the blueprint review indicates the test blueprint is appropriate to use as the guide to develop the prototype items of the product. The items included three types of questions including literal, inferential, and critical.

Test Review

The validation form in this stage used quantitative data analysis. The validation score was in scale ranging from 1 to 3. The experts in this stage consisted of two test experts. The result of the expert validation from Expert 1 shows that 6 items are rarely good, and 14 items are very good. The result of the expert review also shows the test is possible with revision. The addition from expert 1 in the discussion is about the reason for the word used in the cloze test which consists of grammar and vocab. Moreover, it will be very good to add the question from a critical level.

The result from Expert 2 shows that the score given is very good except for the language used which is rarely good and still need correction. The addition from the expert 2 in this result is about the language used in the items related to the grammar. In addition, the expert 2 also helped create the appropriate sentence or instruction used as the clue in the C-DA test program. Moreover, the result from expert 2 shows that the product is possible with revision.

The result indicates that the product needs correction in material principle related to the word choice used as the word bank for the blank. Meanwhile, the items in the assessment have been appropriate with the construction and language principle. After being reviewed, the product was then revised to get the correct one.

ICT Expert

The result of ICT expert review shows that the test is checked positively. The suggestion from the ICT expert is about the next test should probably be more attractive particularly in layout. Additionally, the expert suggested adding one more page for detailed instruction. From the reviewed of the ITC expert, it can be concluded that the product is feasible with revision. The result of ICT expert review indicates that the product has no significant problem meaning that the product work properly related to element of design and text including the appearance, sequence and time allotment.

Item Analysis

From the ITEMAN version 3.50 analysis, it was found that the reliability of the test which was symbolized as alpha was 0,891. The alpha score shows that the reliability of the test is categorized as good. The next analysis was item difficulty. The result indicates most of the items are categorized as moderate. The result of item difficulty analysis is shown in the table 3.

Table 3. The Result of Item Difficulty Analysis

Index Range	Category	Item Number	F
> 0.7	Easy	0	0
0.3—0.7	Moderate	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19, 20, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 57, 58, 59, 60	53
	Difficult	13, 14, 21, 25, 26, 40, 56	7

After item difficulty was analyzed, item discrimination was analyzed. This analysis was carried out in order to know how well the item in discriminating the low and high ability students. Most of the items had very good discrimination to be used in the product. The poor discrimination items included item 6 and 10 which then were dropped from the product. The result of item discrimination is presented in the following table 4.

Table 4. The Result of Item Discrimination

Index Range	Interpretation	Item Number	F
≥ .40	Very good		45
		2, 3, 4, 5, 7, 11, 12, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 37, 38, 40, 41, 42, 43, 45, 47, 48, 50, 51, 52, 55, 57, 58, 60	
.30—.39	Good	9, 26, 35, 36, 39, 46, 49, 53, 54, 56, 59	11
.20—.29	Fair	1, 8, 13, 21, 25, 44	6
≤.19	Poor	6, 10	2

The last analysis was the effectiveness of distractors. Based on the data from ITEMAN result analysis, there are 2 items which have suggested answer key. These items were dropped from the product, namely 10 and 21. From the item analysis, 57 items were valid and used as the test item in the product. Those 57 items were then categorized as a literal, inferential, and critical question in which most of them were inferential, some are literal, and few of them were critical.

Furthermore, the analysis for cloze procedure included two invitation letters, namely business event and business meeting invitation. These letters were analyzed using the Gunning Fog index in order to know the level of readability of the text. The result indicates those texts have different level of readability. The readability of business meeting invitation index was15.45 indicating that the reading level was for collage junior, and the readability index of business event invitation was 10.33 indicating it was for high school sophomore level. From the result, the invitation used in the product was the business event invitation.

Try-Out

At the first try-out, the navigation button was error. Therefore, the students had some trouble to access the page. There were also a number of items jumbled. Moreover, the problem also came from the scoring system which did not work as the principle of the researcher created. In the second attempt, the product could run well and there was no problem found after the entire problem fixed, and the result was then used as the final tryout. The try-out of the product was conducted in 45 minutes for each type of the test, so it took 90 minutes for each theme.

The Result of Students' Questionnaire Analysis

The questionnaire was given to the subjects to gain information about the performance of Computer Dynamic Assessment (CDA) from their point of view. The questionnaire consisted of 18 multiple choice items and 1 essay and it was distributed to the 60 subjects after they did the CDA. The questionnaire was valid with $r > 0.215$ from the r table n=84 and the reliability score was 0.833 meaning very reliable.

The result reveals that only a few of the subjects used the computer-based test and the presence of the computer-based test should be introduced more often, particularly in the form of an online one. The result also reveals that only a few subjects often did the computer-based test similar to C-DA. Therefore, it means that the online test model should be introduced more often as the daily tests or daily exercises. In addition, the presence of reading comprehension model of computerized dynamic assessment in school is needed.

Regarding the product, most students said that the instruction of the test was understandable and clear. From the material used in the product, most subjects agreed that the material was appropriate with the material has been taught. Since the test was about reading, there were some passages consisting on the test which indicated the passage in the product was long enough to the test. Related to the amount of items, most students said it was suitable enough with the time allocation to finish the test. Furthermore, most subjects agreed that the passage was fairly easy meaning that the passage in the test was appropriate. Therefore, the result indicates the product is understandable, clear, and appropriate to the students.

The characteristic of C-DA is the clue/hint used in their assessment process. Related to the clue/hint in the CDA, most subjects said that the clue used in the product was quite and fairly appropriate. Moreover, the color of the clue was also suitable enough for the subject to recognize. Meanwhile, for the instruction used in the clue, the subjects had less difficulty in understanding the instruction. Moreover, the result indicates the clue in the product is very helpful for the subjects.

Most students were interested in the product since they had not found out this product before. Moreover, they also felt motivated while operating the product. It indicates that this product needs to be developed since it brings the interest for the subject and can also motivate the subjects to work on the reading comprehension test.

Related the appearance of the product, most subjects said that the appearance was quite interesting and the color chosen in the product was suitable and had no problem. Since more than half of the subjects were quite interested in the product, it means that the presence of the product (C-DA) is needed. The last was a short essay question about the students' opinion of the product C-DA. Most subjects said that the C-DA was good, interesting, effective, practical, motivated, helpful, making curious, and challenging, but it somehow also made frustrated if the internet connection was disconnected.

The Final Product

All stages in conducting Research and Development have been carried out. From 55 items in multiple choices and 1 invitation letter cloze procedure, it remained 20 items for the product after the analysis. The passages used for the product were 6 passages out of 11 passages as used in the try-out. In the cloze test, the product only used 1 passage out of 2 passages from the tryout which consisted of 20 blanks.

The product is an online computerized dynamic assessment for reading comprehension. It is the prototype of reading comprehension assessment consisting of formal invitation themes. The multiple-choice type consists of 20 items and the cloze procedure consists of 1 passage with 20 blanks.

DISCUSSION

This research has described the format and result of an initial attempt at developing Computerized Dynamic Assessment (C-DA) available online. The instrument focuses on reading comprehension. The major contribution of the approach is administering C-DA to the students as the daily test or assessment both multiple-choice types and cloze procedures. Meanwhile, the result is reported in a quantitative format which is easily interpretable. C-DA is the test that brings the students to the correct answer with mediation. Poehner & Lantolf (2010) argued that dynamic assessment provides the ability which brings to the greater insight into the L2 learners' ability in the language also support their continuous development.

The mediation used in this product is a group of instruction arranged in implicit to explicit levels. This mediation is called Graduated Prompt Approach (GPA) formulated by Anne Brown and colleagues. This mediation is offered to the learners whenever they find difficulties to respond to the test item. In this product, the mediation also reflects the level of prompting required for each item, also interpreted as an indicator of learners' performance (Poehner & Lantolf, 2013). Different from other C-DA products, the total mediation of each item needed in this product can be seen by the administrator to know the students' level ability in every item. This product was created not only to recognize the learners' ability in language but the most important was to reveal the learning potential of the learners. Moreover, this product can help the examiner predict the right remedial or feedback for as the follow-up from the test result. The result is in line with J. Carney & Cioffi (1990) explaining that dynamic assessment uses a response-to-instruction pattern to complement the traditional diagnostic assessment of word recognition and comprehension i.e. the process that helps the examiner predict proper remedial or feedback. This process-oriented study aims to reveal the students' learning potential (Ebadi & Saeedian, 2015) so that can lead to a more comprehensive lesson plan and some appropriate future treatment (Harding, Alderson, & Brunfaut, 2015).

Regarding the effectiveness and efficiency of the product, this product is an online program test so it does not need any paper to copy the test. Moreover, the teachers can get the score of the students' work and the detail of the item which have not accomplished or finished trough the administrator account to see the difficulties of the students' have based on the amount of the clue they need to solve the item.

CONCLUSIONS

Based on the process and the result in developing the computerized dynamic assessment, there are two domains of the conclusion that can be drawn. The conclusion covers also see the questions which have not been accomplished, so it can be used as the decision to arrange the remedial. In addition, this test is efficient since the teachers do not need to copy the paper and make a waste. Moreover, this CDA prototype is also a good model of the test in this era, specifically for daily assessment since the materials used in this test cover those in the syllabus.

Meanwhile, the only weakness of the product is related to the stability of the internet at school. The product of this research is an online-based test which needs a stable network to run. Based on the strengths and the weakness of this study, it can be concluded that the developed computerized dynamic assessment for reading comprehension can be used as the test which replaces the previous impractical and uncovered test in vocational high schools.

After conducting the process in this research, there are some suggestions for the future completion. This product can be a model of online-based test for reading comprehension since in this era the students need more interaction with technology especially by having an online test as their daily assessment in class. As this online test needs a high-quality network, it is expected that the principal can provide faster and more stable network in the future, so that the use of online test can be maximized. From the teacher, this product can be a model for other reading themes as well as decision of the students' performance on C-DA reading comprehension test. Moreover, this test product also can cover all of the students from low to high-level abilities since this product is developed with mediation. Therefore, it can be used as a reference to create a follow-up action of the students' work. Besides the researcher development, it still remains the future development. This product can be an insight into the effectiveness of the online reading test in enhancing students' reading motivation with better qualifications. As the reference, this study can also be used to develop other language skills and another type of test based on computerized dynamic assessment. Another issue concerns is that the item in this program can collaborate with the item response based test.

REFERENCES

Brown, H. D. (2007). *Teaching by Principles: An Interactive Approach to Language Pedagogy*. New York: Pearson Education

Davison, M. L., Biancarosa, G., Carlson, S. E., Seipel, B., & Liu, B. (2018). Preliminary Findings on the Computer-Administered Multiple-Choice Online Causal Comprehension Assessment, a Diagnostic Reading Comprehension Test. *Assessment for Effective Intervention*, *43*(3), 169–181. https://doi.org/10.1177/1534508417728685

Ebadi, S., & Saeedian, A. (n.d.). The Effects of Computerized Dynamic Assessment on Promoting At-Risk Advanced Iranian EFL Students' Reading Skills. *Issues in Language Teaching* 26. 10.22054/ILT.2015.7224

Echternacht, G. (1976). Characteristics Distinguishing Vocational Education Students from General and Academic Students. *Multivariate Behavioral Research*, *11*(4), 477–491. https://doi.org/10.1207/s15327906mbr1104_8

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational Research an Introduction* (Seventh). Boston: Longman.

Gallardo-Echenique, E. E., Bullen, M., & Marqués-Molías, L. (2016). Student Communication and Study Habits of First-Year University Students in the Digital Era. *Canadian Journal of Learning and Technology*, *42*(1). https://doi.org/10.21432/T2D047

Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic Assessment of Reading and Listening in a Second or Foreign Language: Elaborating on Diagnostic Principles. *Language Testing*, *32*(3), 317–336. https://doi.org/10.1177/0265532214564505

J. Carney, J., & Cioffi, G. (1990). Extending Traditional Diagnosis: The Dynamic Assessment of Reading Abilities. *Reading Psychology*, *11*(3), 177-192. https://doi.org/10.1080/0270271900110302

Meneghetti, C., Carretti, B., & De Beni, R. (2006). Components of Reading Comprehension and Scholastic Achievement. *Learning and Individual Differences*, *16*(4), 291–301. https://doi.org/10.1016/j.lindif.2006.11.001

Naeini, J., & Duvall, E. (2012). Dynamic Assessment and the Impact on English Language Learners' Reading Comprehension Performance. *Language Testing in Asia*, *2*(2), 22. https://doi.org/10.1186/2229-0443-2-2-22

Placklé, I., Könings, K. D., Jacquet, W., Struyven, K., Libotton, A., Van Merriënboer, J. J. G., & Engels, N. (2014). Students' Preferred Characteristics of Learning Environments in Vocational Secondary Education. *International Journal for Research in Vocational Education and Training*, *1*(2), 107-124. https://doi.org/10.13152/ijrvet.1.2.2

Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's Teaching-Assessment Dialectic and L2 Education: The Case for Dynamic Assessment. *Mind, Culture, and Activity*, *17*(4), 312–330. https://doi.org/10.1080/10749030903338509

Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the Equation: Capturing L2 Development during Computerized Dynamic Assessment (C-DA). *Language Teaching Research*, *17*(3), 323–342. https://doi.org/10.1177/1362168813482935

Pourhosein Gilakjani, A., & Sabouri, N. B. (2016). How Can Students Improve Their Reading Comprehension Skill? *Journal of Studies in Education*, *6*(2), 229. https://doi.org/10.5296/jse.v6i2.9201

Sabatini, J. P., O'reilly, T., Halderman, L., & Bruce, K. (2014). Broadening the Scope of Reading Comprehension Using Scenario-Based Assessments: Preliminary Findings and Challenges. *L'Année Psychologique*, *114*(04), 693–723. https://doi.org/10.4074/S0003503314004059

Sainsbury, M., Harrison, C., & Watts, A. (2006). *Assessing Reading from Theories to Classrooms*. Berkshire: NFER.

San, K. M. (2019). The Thinking Levels Demanded in Reading Activities in the Coursebook Global A2+. *International Journal of Education and Research*, *7*(5), 23-36.

Sciarra, D. (2004). *School Counseling: Foundations and Contemporary Issues.* Belmont, California: Thomson/Brooks/Cole.

Shabani, K. (2012). Dynamic Assessment of L2 Learners' Reading Comprehension Processes: A Vygotskian Perspective. *Procedia - Social and Behavioral Sciences*, *32*, 321–328. https://doi.org/10.1016/j.sbspro.2012.01.047

Shawn K. Wightman, & R. Craig Roney. (2013). The Effects of Story Performance on Fifth-Grade Students' Comprehension of Narrative Texts. *Storytelling, Self, Society*, *9*(1), 20. https://doi.org/10.13110/storselfsoci.9.1.0020

Sofa. (2016). *Developing A Model of An Online Reading Summative Test For College Students of Reading Comprehension* (Thesis). Universitas Negeri Malang, Malang.

Sulistyo, G. H. (2016). *EFL Learning Assessment at Schools an Introductionto ITS Basic Concepts and Principles*. Malang: CV. Bintang Sejahtera.

Thouësny, S. (n.d.). *Assessing Second Language Learners' Written Texts: An Interventionist and Interactionist Approach to Dynamic Assessment*. 6.

Wilson, K. B., Tete-Mensah, I., & Boateng, K. A. (2014). *Information and Communication Technology Use in Higher Education: Perspectives from Students*. 12.