Prototype of Indonesian Reading Test For The Foreign Students

Laili Etika Rahmawati, Sarwiji Suwandi, Kundharu Saddhono, Budhi Setiawan
Universitas Muhammadiyah Surakarta, Universitas Sebelas Maret Surakarta
Laili.Rahmawati@ums.ac.id

Abstract-The research aimed to develop the prototype of Indonesian reading test for the foreign students. The research method applied the research and development (R&D). This research is part of the research in third stage that is product development. The product development is based on the graduates competency standards (SKL) of Indonesian language for the foreign speakers (BIPA) which is consisting of seven levels, namely BIPA 1, BIPA 2, BIPA 3, BIPA 4, BIPA 5, BIPA 6, dan BIPA 7. Specifically, the product which is developed is the reading competence. Data collecting technique is conducted by the technique of documentation, interview, and test. The test of data validity is done by the method of expert judgement. Based on the results of the development, the prototype of Indonesian reading test for the foreign students can be described as follows: (1) the reading competency test for the foreign students consists of seven packages of questions which are developed based on basic competencies established by the government; (2) the test was developed in the form of objectives tests for BIPA level 1 through BIPA level 5 and subjectives test for BIPA 6 and BIPA 7 levels; and (3) the assessment criterias are based on the degree of difficulty of the problems.

Keywords-prototype; test; Indonesian language; reading competence; foreign students.

I. INTRODUCTION

There are four competencies that must be mastered by the language users, namely listening, speaking, reading, and writing. Reading is the ability to understand information submitted by other parties through written means. The reading competency test is intended to measure the competence of the test participants to understand the content of the information contained in the text. Therefore, the reading text material which is tested should contain information that is demanding to be understood. In general, decent discourse taken as a reading competence test material is not much different from listening competence test material (Nurgiyantoro, 2016).

Pedagogically good reading is appropriate to the level of development of the soul, interests, needs, or attention. In the world of education, activity and the task of reading is something that can not be bargained. Most of the science acquisition is obtained from the reading activities. The success of one's study will be greatly determined by the ability and willingness in reading. This is supported by the statement that various job demands required adequate reading competence.

The initial phase of this research was conducted by exploratory study of reading test instrument which has been developed by Indonesian language program organizer for foreign students. The reading test instrument developed by the program organizer is used to measure the reading competencies before, during and after the learning process. The result of this exploratory study shows that the reading competence test is developed in the form of multiple choice tests. The discourse used in reading test consists of several topics, namely: (1) social service; (2) assistance of the Saudi King for earthquake victims in West Sumatera; (3) urine-powered cars; (4) alternative fuels from mushrooms; (5) ola morning exercise, (6) toxoplasma, (7) drug dose, and (8) writing talent. The discourse is sourced from the mass media published in 2009 and tested in 2016. Each reading text or the discourse is developed into 3-5 questions. The length of the discourse is about 50-200 words. The total of reading question consists of 20 items. Allocation of reading test time is 90 minutes.

The results of exploratory studies conducted are verified by conducting the interviews. BIPA teachers act as the informants because BIPA teachers are the ones who develop reading test instruments for foreign students of BIPA learners. The interview result shows that the reading competence test is purposely made short (50-200 words each reading) to help the students to easily understand the presented text. The 90 minute time allocation for 20 items is considered as too long because the foreign students are able to complete the reading test at least 45 minutes to finish 20 items of questions.

As a form of reinforcement on the results of interviews with BIPA teachers, foreign students who act as tested parties are also interviewed related to the reading test instruments that being done. From the interviews process, it shows that with short reading text (about 1 to 3 paragraphs) the reading activity can be easier to be understood. According to the students' opinions, it could be known that not all the topics which are tested are interesting and their needs. The interesting topics that attract students include, urine-powered cars, alternative fuels from mushrooms, exercise, and drug doses. Students are interested in these topics because they already have an initial knowledge (schemata) about the discourse given.

From the result of the document analysis, it could be found that the questions' items did not refer to the reading presented, whether the tests' participants can answer correctly were not because they are capable of understanding the reading text, but they only had the initial knowledge about the discourse or because of answering based on speculation. This causes the reading competency test instrument which is developed, considered to be categorized as invalid and unreliable.

The research results of Behfrouz and Nahvi (2013), entitled "The Influence of Problem Characteristics in Reading Ability in the IELTS" describes that one of the abilities that involves the thinking process is the reading skill. Reading is an ability that makes students become familiar with other people's ideas; comparing and contrasting different ideas; test and evaluating the arguments, interpretations, beliefs, or theories; making inferences, predictions, or interpretations; and exploring the implications and consequences. In essence, reading is an ability that encourages students to think critically.

The research data revealed that participants showed the differences' results in the IELTS test because of different forms of questions that given as a treatment. The results showed that the characteristics of the problem had a significant influence in the ability of reading tests on IELTS.

Based on the result of the research, the development of Indonesian language competency test model, especially reading competence, also needs to consider the characteristics of the questions' items. Differences in the characteristics of the questions that can significantly differentiate the ability of test participants could be applied to be able to shows the validity of the test instrument. The test instrument of reading competence which is being developed, attempts to combine the various forms of reading tests so that the portrait of language competence is clearly visible and accurate.

Specifically, language competency assessment can be divided into four categories: (1) discrete; (2) integrative; (3) pragmatics; and (4) communicative (Wahyuni dan Ibrahim, 2012; Nurgiyantoro, 2016).

Discrete test is a test that emphasizes or concern only on one aspect of language at a time. Each item questions is intended to measure only one aspect of language, such as phonology, morphology, syntax, or vocabulary. The discrete test is not only related about the language aspect, but also various language skills. If a language test is specifically intended to measure only one language skill, such as listening, reading, speaking, or writing, without relating to other abilities, the test of language skill is considered as a discrete test. Thus, the ability to listen, speak, read, and write should be tested separately. Receptive and productive skills should be tested in different tests.

An integrative test emerges as a reaction to discrete test theory. In an integrative test, the aspects and language skills are covered simultaneously. Thus an integrative test is a linguistic test that attempts to measure several aspects of language or language skills at a time.

Pragmatic test is an approach in language skills test to measure how well learners use language elements according to the real communication context. Communicative test is based on the view that language usage should be suitable to everyday communication. The communicative approach emphasizes the role of the elements of non-adherence, especially the elements associated with the implementation of good communication.

Weir (1990: 7) stated that *"in testing communicative language ability we are evaluating samples of performance, in certain specific contexts of use, created under particular test constraints, for what they can tell us about a candidate's communicative capacity or language ability."* In that definition, Weir shows several key words in communicative language tests, namely: (1) communicative language skills, (2) use of specific contexts, (3) test constraints, and (4) capacity. Bachman (1990) introduces communicative language skills as a qualification framework to describe language knowledge and language capacity to be implemented in the context of appropriate language communication.

Communicative competence test emphasizes the importance of language function as a communicative function. Communicative competence is a competence to understand and use the language in communication activities in factual and reasonable according to the context of the conversation. Competence includes grammatical, sociolingual, contextual, and strategic competencies.

Grammatical competence is the competence which associated with various linguistic elements. Sociolinguistics competence is a competence which related to the ability to use language in accordance with social conditions of the society. Contextual competence is a competence related to the ability to use the language in accordance with the situation and condition (context) of the talks conducted. Strategic competence is the ability to choose a communication strategy appropriate to the desired effect (Douglas, 2000).

Various aspects of language and communicative functions in order to understand and use the language must be integrated in communicative tests. It means that the language test is known as containing the aspects of language, but it must be integrated in the form of understanding and the use of language in a reasonable and contextual. Linguistic test which is intended to measure the grammatical competence which considered as the basic ability to communicate needs to get their own attention. However, it should not be independent from the communicative language functions. If it is forced to conducted, it will turn into another type of language test that does not measure the language's communicative competence. Thus, the communicative test will form linguistic test, comprehension, and language usage, or language competency tests.

Concretely a communicative test will involve a fourth test of language proficiency and tests on all four aspects must be contextual. The communicative test should be designed in such a way that there is a resemblance between the language ability to be tested and the reality of everyday language use in a particular context (Nurgiyantoro, 2016; Brown, 2004).

The test can be divided into two forms, namely subjective test and objective test. The subjective test form is also called an essay test (Suwandi, 2011; Nurgiyantoro, 2016). The essay test is a form of question that demands the test participants' answers in the form of a description using their own language. The essay test demands to think broadly about existing questions and use what is known regarding the questions to be answered. An essay test gives students the freedom to compose and present their own answers in a relatively limited sphere. That is why an essay test is also called as a subjective test, although the naming is also associated with the assessment activities that are also subjective.

The subjective test allows test participants to demonstrate their ability to apply knowledge, analyze, link, and evaluate new information (questions) that should be done. This test requires test participants to be able to relate the facts and concepts to organize them into logical coherence, and then pour the results of that thought in the form of written expression using their own words. This test kit can assess different types of competence, such as expressing opinions, logical thinking, and concluding. The weakness of this type of test instrument, among other in which is the limited material coverage in question.

The objective test is also called a short answer test. As the name implies, a short answer test requires the test takers to give a short answer, even just selecting certain codes that represent the alternative answers that have been provided. The answer to the objective test is definitely, there is only one possible correct answer. When the answer is definite, the result of the work corrected by anyone will produce a score more or less the same.

In general, the objective test can be divided into two types, namely choosing answers and supplying answers. The objective test of selecting answers can be differentiated into several types, namely: (1) multiple choices; (2) true-false/yes-no questions; (3) matching; and (4) cause and effect. Tests supplying answers can be distinguished on: (1) stuffing (the essay) or completing; (2) short answers; and (3) description.

The type of answer-picking test is an assessment tool that only assesses the low-thinking skills. Test participants who do not know the correct answer of the answer-picking test are possible to answer by guessing. Test participants tend not to attempt to understand the material, but memorize the questions and answers. In addition this type of test is less able to provide enough information to be used as feedback in diagnosing the test participants' competences.

Both the subjective and objective test form, each has their weaknesses and advantages. The subjective test has the following advantages: (1) it is appropriate to judge thinking processes that involve high-level cognitive activity and not merely remember and understand facts or concepts; (2) forcing test participants to bring their answers into appropriate language according to their own style; (3) it is not a gambling in answering; (4) easy to organize and not spend much time. The weaknesses of subjective test forms: (1) the level of validity and reliability is low; (2) the answer is not easy to be standardized; (3) time allotment to check the work is relatively long. There are several attempts that can be made to reduce the weaknesses of the subjective test, namely (1) the indicators and materials selected for testing should be the main ingredients that represent indicators and materials which are not tested; (2) there is a limit for the correct answer; (3) it is necessary to develop certain criteria that serve as guidelines.

The objective test has several advantages, including: (1) indicators and materials tested are thoroughly; (2) allows only one correct answer; (3) easily to be corrected. The weaknesses of objective test are as follows: (1) the preparation of objective test takes a relatively long time; (2) the testing taken by participants may be gambling in answering; and (3) it will cost a lot to double or copy it. Efforts that can be done to reduce the weaknesses of objective tests are: (1) the preparation of questions items based on the grid; (2) the test instrument may be used for more than once; (3) the use of subjective and objective tests may be combined (Nurgiyantoro, 2016).

The weaknesses of the objective test were shown in the study of Alibakhshi and Ali (2011) on the external validity of the TOEFL in the doctoral entrance examination in Iran which explains that external validity (in general) is one aspect of the construct validity related to the inference that used as the basis for the assessment of the test participants. External validity is considered as important in high-level tests, such as an English-language test for academic/ specific purposes used to evaluate the competence of English skills of the test takers in general in which to investigate the validity of the TOEFL test in general that held at universities in Iran in determining the doctoral candidates. Research result indicates that there is a significant difference between the mean values of the participants in the TOEFL test and their mean value in academic language ability. Furthermore, TOEFL scores could not predict the test participants' score significantly in using academic language in the situations when using target language. Therefore, the TOEFL test developer should take the general ability issue as a consideration when planning the TOEFL test model.

The study of Alibakhshi and Ali is very relevant to this research. The results of theirs study can be taken into consideration in the validity and reliability test when the test instrument developed. The TOEFL's inability to predict the test particicpants' scores significantly is a finding to be considered in this study. External factors that influence the test results need to be controlled so that the test instrument can measure what should be measured (valid).

II. METHODS

The method used in this research is research and development (R & D) method. Gall, Gall, dan Borg (2007: 589) stated that *"research and development is an industry-based development model in which the findings of research are used to design new products and procedures, which then are systematically field-tested, evaluated, and refined until they meet specified criteria of effectiveness, quality, or similar standars"*.

Clarifying the definition above, Sukmadinata (2012) stated that research and development is a process or steps to develop a new product or refine an existing product, which can be accounted for. These products are not necessarily in the form of hardware shapes, such as books, modules, classroom or laboratory teaching aids, but can also be software, such as computer programming for data processing, classroom learning, libraries or laboratories, or models - educational model, learning, training, guidance, evaluation, management, etc. More briefly, Sugiyono (2013) defined that research and development as a research method used to produce a particular product and test the effectiveness of the product.

The strategy used in the research of this development phase is the Criterion-referenced Language Test Development (CRLTD) Workshop which aims to develop the test that refers to the guidelines of criterion reference (PAK). CRLTD not only refers to who will be tested, but also the goal of the language to be achieved. Goals should be specific, for professional or academic purposes (Douglas, 2000).

There are several steps that must be done in constructing the test, according to Hughes (1996) there are seven steps to be taken in developing test instruments, namely: (1) determine the type of test which is going to be developed (achievement test, proficiency test, diagnostic test, or placement test); (2) determine the purpose of test development; (3) determine the ability of the test participants which are going to be tested; (4) determine the details of the results which are going to be measured; (5) determine the accuracy of the test; (6) determine the impact of the test; and (7) describe the limitations and the weaknesses of the test.

Douglas (2000) has different procedures for developing language tests for specific purposes. There are eight steps that must be taken to develop the test for specific purposes, namely: (1) describing the purposes of the test which are going to be developed; (2) describing the situation of the target language and the details of the tasks which are going to be performed; (3) describing the characteristics of language users/test participants; (4) determining the construction which is going to be measured; (5) describing the content of the test; (6) describing the assessment criteria; (7) providing specific examples of task/item that will be considered as the basis of generalization; and (8) identifying the quality of test instruments developed by testing validity, reliability, and visibility.

Research subjects in this research of development stage are students, stakeholders, researchers, and experts. Students act as test participants whose test results serve as the basis for data analysis. Stakeholders is the organizer of the implementation of the Indonesian language program for foreign students who are tasked to test the students and provide input on the implementation of research. Researchers are responsible for conducting analysis and synthesis of development outcomes. The expert serves as an expert in charge of providing inputs on reading competence test instruments for foreign students developed and/or providing improvement advice in the development of test instruments.

The output of this development stage is an Indonesian language competency test instrument for foreign students which is developed based on the test objectives and description of the target language criteria as well as details of tasks to be done based on the construct validity.

III. FINDING AND DISCUSSION

The reading competency test for foreign students is developed based on the general description and basic competencies contained in the Graduate Competency Standards (SKL) of Indonesian Language Courses and Training for Foreign Speakers (BIPA) published by the Directorate General of Early Childhood Education and Public Education, Ministry of Education and Culture of the Republic of Indonesia. General description and basic competence of reading BIPA 1, BIPA 2, BIPA 3, BIPA 4, BIPA 5, BIPA 6, and BIPA 7 are shown in table 1 below.

Table 1. General Description and Basic Competency of Reading BIPA 1 to BIPA 7

Level	General Description	Basic Competence
BIPA 1	Able to understand and use the expression of self introduction and the fulfilment of routine daily concrete needs by simple communication with others cooperatively.	1.1 Able to read the short text loudly and simply related to self information, such as: name, address, occupation, country of origin, family, etc. 1.2 Able to find information related to direction, location, around environment, and daily activities, such as advertisement, map, announcement, menu, and schedule. 1.3 Able to find informations from the short and simple text, such as personal mail.
BIPA 2	Able to express feelings simply, describe the environment, and communicate daily and routine needs.	2.7 Able to read procedural text contained in public places, e.g. in banks, ATMs, immigration offices, hospitals, campuses, libraries. 2.8 Able to perform simple and routine tasks in accordance with the text content of guidelines, instructions, procedures, etc. 2.9 Able to abstract the text description of feelings, self-background, environment, and matters related to their needs, such as education, health, social interaction, and entertainment.
BIPA 3	Able to express briefly and coherently the experiences, hopes, goals, and plans with accompanying reasons in the context of life and daily work tasks.	3.9 Able to understand text that uses language with everyday words or work-related. 3.10 Able to understand texts that use language with everyday words or related to the common things found while working, studying and relaxing. 3.11 Able to understand the description contained in a non-formal/personal letter relating to an event, feeling, or expectation.
BIPA 4	Able to report on observations of events and express ideas on the topic of the students' field, both concrete and abstract, smoothly without any obstacles that interfere the understanding of partners as listeners.	4.7 Able to understand the contents of articles and reports relating to the latest issues. 4.8 Able to re-express moral messages in fairy tales or folklore. 4.9 Able to understand short story with simple language.
BIPA 5	Able to understand long and complicated texts and able to express ideas from a variety of perspectives on various topics spontaneously and fairly smooth	5.5 Able to understand long and complex factual and literary texts with different writing styles. 5.6 Able to understand long articles

	almost without obstacles, except in professional and scientific fields.		and technical instructions even if the text is not related to the field.
BIPA 6	Able to understand long, complex, or implicit texts and express ideas in clear, structured, systematic, and detailed language according to social, professional, and although there is an obstacle for the complex academic situations.	6.5	Able to interpret text related to the social, academic and professional sphere.
		6.6	Able to analyze text related to social, academic, and professional domains with structured language and systematic pattern of organizational texts.
BIPA 7	Able to understand information almost at all fields easily and express ideas spontaneously, fluently, precisely by distinguishing nuances of meaning, as well as reconstructing arguments and data in a coherent presentation.	7.9	Able to read all forms of written language without difficulty, including abstract text of complex structures and languages, such as guidance books, scientific articles, and literary works.
		7.10	Able to use facts to make the exact conclusion of various scientific texts and literary works.
		7.11	Be able to Reading critically various texts (interpreting, responding and then evaluating).
		7.12	Able to Integrate (synthesize) almost identical ideas or problems from various texts of scientific papers and literary works.

Referring to the general description table and basic competence of BIPA 1, BIPA 2, BIPA 3, BIPA 4, BIPA 5, BIPA 6 and BIPA 7, the construction of reading competency test for foreign students is also developed in seven levels.

Reading Competency Test of BIPA Level 1

BIPA Level 1 reading competency test is developed based on seven types of text that fit the bill described in BIPA 1 basic competence. Construction of BIPA 1 reading competence test can be seen in table 2 below.

Table 2. Reading Competency Test Construction of BIPA 1

Section	Topic of the Text	Task Form	Number of Task Items	Score per Task Item	Maximum Score
A	Self Introduction	Reading aloud	50 vocabularies	0,1	5
B	Map	Objective 'True' / 'Fals'	5	1	5
C	Job Vacancy	Filling the uncomplete sentence with match making technique	5	1	5
D	Annoucement	Multiple choice with four alternate answers	5	1	5
E	Food Menu in the Restaurant	Objective 'yes'/ 'no'	5	1	5
F	Schedule	Completing the uncomplete text	5	1	5
G	Electronic Mail	Matching	5	1	5

Reading Competence Test of BIPA level 2

The BIPA level 2 reading competence test was developed into 35 items' questions from seven different types of discourse. Each item if answered correctly by the test takers will get a score of 2. Then the maximum score that can be achieved is 70. The reading competence test construction of BIPA 2 can be seen in table 3 below.

Table 3. Reading Competence Test Construction of BIPA Level 2

Section	Discourse Title	Type of Question
A	How to Charge Mobile Phone Pulse in ATM	Multiple choices with four possible answers
B	The Procedures in Passport-Making	Objective Answer by 'yes'/ 'no"
C	Train Check-In and Boarding Instruction	Objective Answer 'true'/ 'false'
D	The Baron Beach	Paragraph Identification
E	Bakwan (Javanesse traditional food)	Completing the uncomplete text by matching technique
F	Greeting Tradition	Matching technique
G	Eating Procedures of Indonesian People	Objective Answer 'true'/ 'false'

Reading Competence Test of BIPA level 3

The BIPA level 3 reading competence test was developed into 35 items of questions from seven different types of discourse. Each item if answered correctly by the test takers, they will get a score of 3. Then the maximum score that can be achieved is 105. Reading competence test construction of BIPA level 3 can be seen in table 4 below.

Table 4. Reading Competence Test Construction of BIPA Level 3

Section	Discourse Title	Type of Questions
A	Cinema electronic	Multiple Choice with four alternative reading text
B	Tips for Designing a Healthy, Humane, and Functional House	Paragraph identification
C	History of the Day's Name	Objective answer 'true'/ 'false'
D	Personal Letter	Completing the uncomplete text by matchmaking technique
E	Paragraph Development	Paragraph Identification
F	Reader's Letter 1 entitled Establish Social Justice for Disabled People Reader's Letter 2 entitled Provide Friendly Facilities for Disabled People	Statement identification based on the text
G	Elements That Can Make Malicious Cosmetics	Objective 'yes' / 'no'

Reading Competence Test of BIPA level 4

The reading competence test of BIPA level 4 was developed into 35 items' questions from five different types of discourse. Each item if answered correctly by the test takers, he/she will get a score of 4. Then the maximum score that can be achieved is 140. Construction of reading competence test of BIPA level 4 can be seen in table 5 below.

Table 5. Reading Competence Test Construction of BIPA level 4

Section	Discourse	Type of Question
A	Danau Toba Folklore	Objective 'yes'/ 'no'
B	Batu Menangis Legend	Objective 'suitable'/ 'not suitable'
C	Short story of 'Tukang Jahit'	Statement identification based on the text.
D	Dhyptery	Mulptiple Choice with five possible answers.
E	Entrepreneurial of Innovative Waste Management in the Worldwide.	Paragraph Identification.

Reading Competence Test of BIPA level 5

The reading competence test of BIPA level 5 was developed into 40 items' questions from five different types of discourse. Each item if answered correctly by the test takers, he/she will get a score of 5. Then the maximum score that can be achieved is 200. Construction of reading competence test of BIPA level 5 can be seen in table 6 below.

Table 6. The Contruction of Reading Competence Test of BIPA Level 5

Section	Discourse	Type of Question
A	Plagiarism	Multiple choices with five possible answers and the statements identification based on discourse that being read.
B	Reading the Literary Work entitled 'Taghut'	Objective 'yes' / 'no'
C	Definition, Benefits, and the Important Elements of Digital Literacy.	Multiple Choices with five alternative answers, matching, and objective 'yes' / 'no'.
D	Presidential Instruction on Drugs and Food.	Statements identification based on the text that has been read.

Reading Competence Test of BIPA Level 6

The test of reading competence in BIPA level 6 was developed in the form of subjective tests. The test instrument is presented in the form of a discourse entitled 'Education, Social, and Cultural of Japanese People' with a text length of about 2000 words. The test participants are asked to read the text and then interpret, analyze, and present it again in a structured language and have a systematic organization pattern.

Reading Competence Test of BIPA level 7

The test of reading competence in BIPA level 7 is developed in the form of subjective tests. The test participants were asked to compile two scientific articles with the same theme as critical reading techniques. Test participants are asked to interpret, synthesize, and evaluate.

The construction of a reading competence test instrument for foreign students that has been presented in the sub-section of the findings indicates that there is a correlation with the studies that have been conducted by previous researchers. This study is relevant to the research that has been done by Behfrouz and Nahvi (2013) which concluded that the characteristics of reading problems in IELTS had a significant impact on the implementation of the reading test. This is in accordance with the theory presented by Nurgiyantoro (2016) which states that one way that can be done to reduce the weakness of the type of objective test is to combine between types of objective tests with subjective tests. This is appeared in the instrument development of the reading competence test of BIPA 1, BIPA 2, BIPA 3, BIPA 4, and BIPA 5 which are developed in the form of objective tests with various combinations of forms, namely multiple choice, field entry, matching, paragraph identification, and question identification based on the text that has been reads. BIPA 6 and BIPA 7 were presented in the form of subjective tests as the realization of the general description and basic competence presented in the graduate competency standards which established by the government.

The development of a reading competence test instrument for the foreign students is one of the most important things in learning Indonesian as a foreign language. This is relevant to a research conducted by Ashraf and Zolghafari (2018) on the assessment of English teacher language literacy as a foreign language. The literacy level of the teachers has a significant impact on student achievement. The development of a reading competence test instrument for foreign students is one of the real efforts undertaken to provide a means to BIPA teachers and BIPA learners to be able to measure the competence of reading. The reading competence reflects the level of one's literacy. Construction is an important component in the development of reading test instruments (Siong, 2004).

The construction of a reading competence test for foreign students is developed based on general descriptions and basic competencies established by the government. It aims to minimize the discrepancy in the presentation of test instruments that serve as a measure of the reading competence of foreign students. The research conducted by Alfallay (2018) concluded that the determination of test specifications based on clear references is something that should be done by the test developer. This is very influential on the successful of the test implementation. Afallay's research indicates that unpreparedness and non-conformity between the assessment rubric and the test instructions leads to the realization where the test implementation would be not in line with the expectations.

One way to construct a reading competence test instrument for foreign students in order to make the test results in accordance with expectations other than by developing the item in accordance with the general description and basic competencies set by the government is to develop the reading test instrument by combining the type of questions. The combination of the type of questions makes test takers have different strategies to achieve the maximum score. This is relevant to a research conducted by Li and Kaur (2014) about the relationship between reading strategies and the performance of reading tests. The study shows that the Survey of Reading Strategy (SORS) classifies the reading strategy into three categories, namely global reading strategy (GLOB), problem solving strategy (PROB), and support reading strategy (SUP). All of the three strategies can be applied in reading tests according to the discourse specifications and types of questions presented.

IV. CONCLUSION

Based on the results of the research that has been done, it can be concluded that the construction of reading competency test for foreign students is developed based on the general description and basic competencies established by the government of Republic of Indonesia. The reading competency test for foreign students is developed into seven levels, namely BIPA 1, BIPA2, BIPA 3, BIPA 4, BIPA 5, BIPA 6, and BIPA 7. The seven levels of BIPA have different criteria, namely the discourse types, type of questions, the number of questions, and scoring techniques. The higher of the level developed, the score obtained by test participants is also higher. The reading competence test for foreign students at BIPA level 1 to BIPA level 5 is developed in the form of objective test instruments, while BIPA level 6 and BIPA level 7 are developed in the form of subjective test instruments. The type combination of question at each level developed, has aims to maintain the validity and reliability of developed test instruments.

References

Alfallay, I.S. (2018). Test Specifications and Blueprints: Reality and Expectation. *International Journal of Instruction*, *11*(1), 195-210.

Alibakhshi, G. & Ali, H. G. (2011). External Validity of TOEFL Section of Doctoral Entrance Examination in Iran: a Mixed Design Study. *Theory and Practice in Language Studies*, *1* (10), 1304-1310.

Ashraf, H. & Zolfaghari, S. (2018). EFL Teachers' Literacy and Their Reflective Teaching. *International Journal of Instruction, 11*(1), 425-436.

Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Behfrouz, B. & Nahvi, E. (2013). The Effect of Task Characteristics on IELTS Reading Performance. *Open Journal of Modern Linguistics,* 3 (1), 30-39.

Brown, H. D. (2004). *Language Assesment, Principles and Classroom Practice*. San Fransisco: Longman.

Douglas, D. (2000). *Assessing Language for Specific Purposes*. United Kingdom: Cambridege University Press.

Gall, Meredith D.; Gall, Joyce P.; dan Borg, Walter R. (2007). *Educational Research*. Boston: Pearson Education, Inc.

Hughes, A. (1996). *Testing for Language Teachers*. Melbourne: Cambridge University Press.

Li, L. & Kaur, S. (2014). Textbok Reading Strategies and Its Relationship to Reading Test Performance. *Gema Online ,* 14(3), 1-18.

Nurgiyantoro, B. (2016). *Penilaian Pembelajaran Bahasa Berbasis Kompetensi*. Yogyakarta: BPFE.

Siong, L. K. (2004). Exploring the Connection between the testing of reading and literacy: The Case of the MUET. *Gema Online,* 4(1), 1-10.

Sugiyono. (2013). *Metode Penelitian Pendidikan pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

Sukmadinata, N. S. (2012). *Metode Penelitian Pendidikan.* Bandung: Remaja Rosdakarya.

Suwandi, S. (2011). *Model-Model Asesmen dalam Pembelajaran.* Solo: Yuma Pustaka.

Wahyuni, S. & Ibrahim, A. S. (2012). *Asesmen Pembelajaran Bahasa*. Bandung: Refika Aditama.

Weir, C. (1990). *Communicative Language Testing*. New York: Prentice Hall.