The Validity and Reliability Analysis of English National Final Examination for Junior High School

Dwi Jayanti[1], Nida Husna[2], Didin Nuruddin Hidayat[3]

Universitas Islam Negeri (UIN) Syarif Hidayatullah Jakarta[1, 2,3], Indonesia
Correspondence: Dwi Jayanti, UIN Syarif Hidayatullah Jakarta, Indonesia. e-mail: dwijayanti230@gmail.com

Abstract

The aim of this study is to analyze the validity and reliability of the English National Final Examination for Junior High School academic year 2017-2018. This study took one sample of test package as the tests are similar in terms of the content and differ only in the position of test number. This study was undertaken in one of the private schools in South Jakarta and involved two English subject teachers as validators/validity analysts. The data was analyzed by using a descriptive method. Content validity was analyzed by comparing the test material in the syllabus to the test item, while the construct validity was analyzed by comparing the graduation standard to the test item. Furthermore, the reliability was analyzed by using Kuder-Richardson Formula (KR20-KR21). The findings of this study show that the English National Final Examination for Junior High School academic year 2017-2018 fulfilled the criteria of validity and reliability. The content validity shows 92% valid, the construct validity shows 92 % valid, and the reliability shows coefficient 0,784, which means reliable. To conclude, the English National Final Examination for Junior High School academic year 2017-2018 has fulfilled the criteria as a good test.

Keywords: validity, reliability, tests, English national final examination

1. Introduction

Along with the globalization era, communication with other people abroad using a foreign language becomes more crucial. Learning a foreign language, especially English becomes an essential need in order to build this communication and gives more contribution in facing competitive lives. That is one reason, why English is the first foreign language that studied in Indonesia from elementary to university level.

English in Indonesia is a first foreign language. This subject is inserted into the curriculum as a major which function as a tool in developing students' knowledge and skills in many aspects, such as science, technology, social studies, economy, even culture and art. English has four major

Dwi Jayanti, Nida Husna, Didin Nuruddin H   *Vol. 3 No. 2; 2019*
The Validity and Reliability Analysis of English National Final Examination
for Junior High School

skills to be learned: listening, reading, writing and speaking. In teaching learning process to gain the educational curriculum objectives, many attentions addressed to develop these four skills. It involves the quality of teacher, curriculum, syllabus, teaching method, facility, government policy, the quality of test given, etc. One of the specific attentions in this study is related with the quality of test given.

National Examination plays an essential role. Moreover, in Indonesia, National Examination is one of the requirements that student must pass. Passing National Examination as one of the criteria for students to graduate from a certain education level and be able to continue their study to the next education level. National Examination also has important function as the standardized test to measure students' achievement in certain subject. Government, in this case Ministry of National Education holds National Examination once a year and administered nationally. One of National Examination purpose is to measure the competency achievement in a particular subject in a group of science, technology and language subject as mentioned in National Education Ministry regulation (No.59/2011). The regulation of this National Examination related with the topics and items must refer to competence standard, basic competence, graduation standard, indicators and test specification.

In line with the Ministry of National Education regulation (No 59/2011), that stated "National Examination, abbreviated as *Ujian Nasional* (UN, hereafter), is a national standardized test which is administered nationally in order to test students' competencies achievement on a particular subject in a group of science and technology subjects", the ministry must develop and construct the National Examination question based on test specification. These examination tests and the question must refer to the competence standard and basic competence. In competence standard, the four aspects of language are measured well known as listening, reading, writing and speaking skills. These four skills are elaborated by each point in basic competence and indicators. Furthermore, from these basic competencies then creates material and test specification for students, in order to test and assess students' competence. Start from this point test specifications are developed and constructed. All items and topics in English National Examination are explicitly stated in the test specification.

The National Examination committee as a representative of Ministry of National Education responsible to develop and construct the qualified topics and item in order to test students' competencies. Based on another Ministry of National Education regulation (No 57/2015), National Examination is not only for educational decision again, but also used for mapping the quality of programs and/ or the education unit as a consideration to improve the quality of education. Improving the quality of education means that improving the quality of test topics and items. In this context, content validity, construct validity and reliability should be considered as essential things in measuring students' achievement. The criteria of good test validity and reliability are able to measure students' knowledge, ability, and competencies by presenting appropriate test score.

In language testing and assessment, to measure the quality of the test we know the validity test. Brown (2004) stated that validity is the complex criterion of a good test which is test actually is intended to measure. Gronlund (1998), cited in Brown (2004), stated that validity is the extent to which inferences made from assessment results are appropriate, meaningful, and useful in term of the purpose of the statement. Validity has been identified as the most important quality of the test. One type of validity is content validity, based on Fulcher in Language Testing and Assessment stated that "content validity is defined as any attempt to show that the content of the test is a representative sample from the domain that is to be tested" (Fulcher, 2007). In academic reading test, the types of text will be judged by an expert. These may be subject teachers, or language teachers who have many years' experience. The judges are asked to look at texts that have been selected for inclusion on the test and evaluate them for their representativeness within the content area. Secondly, the items used on the test should result in responses to the text from which we can make inferences about the test takers' ability to process the texts in ways expected of students on their academic courses.

Another type of validity test is construct validity. Construct validity is a degree to which a test measures what it should have to be measured, as Messick (1998) stated:

> "... as an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and action based on test scores..." (p. 40).

Another criterion to measure the good test is reliability test. According to Brown, a reliability test is consistent and dependable. The same test given to the same student or matched student on two different occasions will give a similar result (Brown, 2004). Because the point to prove reliability of the test is consistency, so that related to the test, especially UN for Junior High School must give the consistency results in whenever and wherever the English National Final Examination is conducted.

The English National Final Examination should be based on the curriculum. According to the government, the curriculum that should be used by school is curriculum 2013 revision year 2017 (K13). Curriculum 2013 is replacing the Curriculum 2006 or mentioned as *Kurikulum Tingkat Satuan Pendidikan* (KTSP). In Curriculum 2013 for Junior High School level the basic of the test is reading test with the various level of difficulty question that requires students to use their High Order Thinking Skill (HOTS). Curriculum 2013 has four aspects: knowledge, skill, behavior (social) and attitude (spiritual). Related to this study, this kind of curriculum is being data of analysis.

The significant study of this research for the teachers is they are able to know whether the English National Final Examination that is conducted by government is valid and reliable or not. For the Junior High School, they can evaluate and prepare the suitable material in order to equip their graduate with high order thinking skill, also to improve the knowledge especially about the validity and the reliability of the test. Furthermore, for the government, this study will enrich

their information related to the theory and practical in conducting qualified test that has high validity and reliability, as an improvement in the future.

2. Method

This study was applied the descriptive research approach. In this study, descriptive data are utilized as the main sources, with supplementary numerical data to measure the score of reliability. Numerical data were collected, analyzed, and interpreted. Descriptive statistic allows the researcher to describe data and examines the reliability that provides information about condition, situations, and events that occur in the present (Brown, 2011).

The researcher describes the English National Final Examinations in terms of validity and reliability. The data were collected from the National Examination question sheets year 2018, Answer sheet resulted from the English National Final Examination year 2018, and syllabus of Junior High School. This study involved 2 English subject teachers as the content validity analysts. For numerical data, the researchers got the students' scores of English National Final Examination year 2018 from one of the private Junior High Schools in South Jakarta.

Since the researchers used descriptive research approach, the data were analyzed by comparing the items of the test to the syllabus and graduation standard to find the validity both of content and construct. Content validity compares the material test in the syllabus to the content of items in English National Examination year 2018, while the construct validity compares the graduation standard to the content of items in English National Examination year 2018 (Sugiyanto, 2017). The result shows the level of validity in percentage. While the answer sheet and students' score of English National Examination were analyzed to find out the reliability. The score is measured and calculated by Kuder-Richardson Formula (KR-20/KR-21) and Microsoft excel.

3. Findings

3.1. Validity of the Test

The validity test both content and construct validity were analyzed by measuring in tables, using Microsoft excel. The percentage was calculated using the following formula:

$$Percentage(PQ) = \frac{nm}{N} x100\%$$

Notes:
Nm   : Number of item(s)
N   : Total question items
Each kind of the validity was analyzed as follows:

Content Validity

Analysis of the content validity of English National Final Examination was done by comparing the standard competence in the syllabus to the content of each item of the test, then

calculate the percentage of the learning material in the content of each item included in the test (Geribka, 2016).

From the analysis, it is found that the English National Final Examination of Junior High School academic year 2017-2018 after calculated in percentage represent 92% materials in the syllabus curriculum 2013. Based on the result above, it means the English National Final Examination is categorized valid in term of content validity. Furthermore, as findings from this analysis, the rest of 8 % material of test that out of the syllabus of curriculum 2013 still refer to the previous curriculum 2006 known as KTSP (*Kurikulum Tingkat Satuan Pendidikan*). The material that does not match with the syllabus of curriculum 2013 is about invitation. Based on the expert judgment on this case is English teacher, it can be happened because the 8% is a subset of syllabus in curriculum 2013 and syllabus in curriculum 2006.

From the analysis, it is also found that the materials included in the test are: 16 % grade VII, 50 % grade VIII consist of 42% materials in curriculum 2013 and 8% materials in curriculum 2006, and 34 % grade IX. For conclude, the materials were mostly taken from grade VIII covers 50% or half of all the materials.

Related to Curriculum 2013, that emphasizes the students' knowledge and skills, the whole material consists of short functional text: greeting card (6%), short message (10%), announcement (12%), advertisement (6%) and invitation (8%). For the genre or type of text: Descriptive text (16%), Narrative text (14%), Recount text (14%), Report text (14%) and Procedure text (0%). The material that is found out of the syllabus in curriculum 2013 revision 2017 is an invitation (8%), it is lately known as a subset of the curriculum in English National Final Examination academic year 2017-2018. All in all the short functional text covers 42 % from all the materials and genre or type of text cover 58% of all materials.

Construct Validity

Analysis of the construct validity of English National Final Examination was done by comparing the graduation standard to the content of each item of the test, then calculate the percentage of the learning material in the content of each item included in the test (Sugianto, 2016).

From the analysis, it is found that 46 items number of test represent the materials in the graduation standard. In other words, the English National Final Examination contains 92% of the materials in the graduation standard. It means that the English National Final Examination in term of construct validity is valid.

3.2. Reliability of the Test

Analysis of reliability of the test was measured by using the Kuder Richardson's formula (KR 20/21). The reason of choosing this formula is because the English Final Examination for Junior High School was conducted one time and use Multiple Choice Questions (MCQs). (MCQs) is a kind of dichotomy scoring. The formula as follow:

$$r_{11=} \left(\frac{n}{n-1}\right)\left(\frac{S^2 - \sum pq}{S^2}\right)$$

Sudjiono (2011) interpreted that the reliability coefficient is as follows:

rii ≥ 0.70 : reliable

rii < 0.70 : unreliable

Before calculating the reliability of the test, the total variance (S2) of the test was calculated by the formula:

$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N}$$

$$S^2 = \frac{29.091 - \frac{(755)2}{20}}{20}$$

$$= \frac{29.091 - \frac{(570.025)}{20}}{20}$$

$$= \frac{29.091 - 28.501,25}{20}$$

= 29.4875

The result of calculating the value of total variance (S^2) is 29.4875.

After getting the value of total variance (S2), the reliability of the English Final National Examination for Junior High School academic year 2017-2018 is calculated, and the result is:

$$r_{ii} = \frac{n}{n-1}\left(\frac{S^2 - \sum pq}{S^2}\right)$$

$$= \frac{50}{49} \ (\frac{29,4875 - 6,8075)}{29,4875}$$

$$= (1,020) \ (\frac{22,68}{29,4875})$$

$$= (1,020) \ (\ 0,7691 \ )$$

= 0, 78479

From the analysis of reliability test, the value of reliability is 0,784; it means that the English National Final Examination for Junior High School academic year 2017-2018 is reliable.

Table 1: Coefficient Interval Reliability Test

0.00 – 0.200	Very low
0.20 – 0.400	Low
0.40 – 0.600	Medium
0.60 – 0.800	High
0.80 – 1.00	Very High

4. Discussion

The content and construct validity analysis show a positive result. Nevertheless, more attention should be paid to empirical evidence that there is clearly a domain in the kind of text programmed listed in syllabus of English for Junior High School but do not appear in the final examination test. The procedure text programmed is listed in the syllabus to measure one part of the students' reading competencies in grade IX of Junior High School, but do not appear in the final examination test. From this evidence, the question should be why the procedure text programmed must be listed in the syllabus but, in fact is not inserted into final examination test as one of the examination contents. This evidence is not in line with the Ministry of National Education regulation (No 59/2011), that stated "National Examination, abbreviated as *Ujian Nasional* (UN), is a national standardized test which is administered nationally in order to test students' competencies achievement on a particular subject in a group of science and technology subjects" (Kemendikbud, 2011). The ministry must develop and construct the National Examination question based on test specification. It reflects the redundancy and useless items in the test specification mentioned in the syllabus.

Furthermore, from the content and construct validity analysis, 46 items (92 %) of the final exam match with the syllabus of curriculum 2013 and the rest 4 items (8%) do not match with the syllabus of curriculum 2013; this mismatch content is about invitation material. This material is regarding to the previous syllabus in curriculum 2006 known as KTSP (*Kurikulum Tingkat Satuan Pendidikan*). To conclude, the 8% of final exam content that does not match with the latest syllabus is a subset of syllabus in curriculum 2013 and syllabus in curriculum 2006. The question should be why this material still mentioned in final examination even though it is not listed in the latest syllabus that is used anymore. This implies that the test results do not have any capacity to measure the junior high students' English competencies in curriculum 2013 (K13) comprehensively.

Lastly, besides the validity and reliability aspects, there are many aspects that actually should be analyzed comprehensively to prove that the English National Final Examination has fulfilled the characteristic of a good test. Those aspects are item difficulty, item discriminations, analysis of distracter, and validity of each item of the question (Fulcher, 2007; Brown, 2004). Regarding the time limitation, on this occasion the researchers only had a chance to analyze the validity and reliability aspect. It will be worth considered if the next researchers analyze the other

Dwi Jayanti, Nida Husna, Didin Nuruddin H   *Vol. 3 No. 2; 2019*
The Validity and Reliability Analysis of English National Final Examination
for Junior High School

aspects to make the final examination comprehensively proved to be proper as a test instrument of the educational program.

5. Conclusion

According to the result of analysis, it is concluded that the test has fulfilled many criteria. First, from validity of content, the test is valid because found the 46 items (92%) of the English National Final Examination in the academic year 2017-2018 related with the syllabus. Second, from the construct validity found that the test is valid. The value of construct validity is 92 %. Third, from the reliability of the test, the English National Final Examination academic year 2017-2018 is reliable. It is found that the value of reliability is 0,784; it means the test has high reliability.

From this result of analysis, the problem of this study is solved. The problem of this study is to revealed whether the English National Final Examination of Junior High School academic year 2017-2018 is valid and reliable or not. The validity aspect, both content and construct, is analyzed by comparing the test items to the syllabus and graduation standard. Content validity is the suitability of materials in the syllabus to the contents of item of English national Final Examination and construct validity is the suitability of indicators in graduation standard to the content of item of English national Final Examination. Furthermore, the reliability is analyzed by data of answer sheet which is in form of score. From this score, the reliability is calculated and measured by Kuder-Richardson Formula (KR-20/21). From the analysis, it proves that the English National Final Examination for Junior High School academic year 2017-2018 has fulfilled the characteristic of a good test in terms of validity and reliability.

The present study proposes some suggestions. The test developers are expected to give more attention to the characteristic of a good test in constructing the test. In this case, the government as a policy maker is also expected to improve the quality of the school in remote areas, especially the quality of syllabus development, the facilities, also the easy access to the schools; there are many limited aspects needed improvement so as to give positive effects to the test results. The improvement of English Final Examination test quality will affect the quality of education in Indonesia. For the teachers who have an obligation to make tests, it is suggested to keep developing the syllabus and material become better and related to the students' need in linguistic. Furthermore, they can make a good test and has high quality. The teachers' ability in constructing the good test will give effect to the English Final Examination achieving result. For the researchers, especially for linguistic and education area this study can give the critical information about the quality of a test related to the principles of assessment, especially content and construct validity also the reliability. The future researchers can use this study as an additional source in conducting a research related to the principles of assessment and language test. This study is also useful for the future researchers who conduct the same research with different object and also still related to other aspect of validity and other assessment principles such as item discrimination, analysis of distractor, difficult item analysis and others.

References

Brown, J. D. (2004). *Language assessment: Principles and classroom practices.* San Fransisco, California: Longman.

Brown, J. D. (2011). Quantitative research in second language studies. In Eli Hinkel (Ed.), Handbook of Research in Second Language Teaching and Learning (Vol. II, pp. 190-206). London and New York: Routledge/Falmer Taylor & Francis E-Library.

Education Ministry Regulation No 59 Tahun 2011 Tentang Kriteria Kelulusan Peserta Didik dari Satuan Pendidikan dan Penyelenggara Ujian Sekolah /Madrasah dan Ujian National.

Education Ministry Regulation No 79 Tahun 2015 Tentang Kriteria Kelulusan Peserta Didik dari Satuan Pendidikan dan Penyelenggara Ujian Sekolah /Madrasah dan Ujian National.

Fiktorius, T (2014). *A validation study on National English Examination of Junior High School in Indonesia. Thesis.* Pontianak: Tanjung Pura University.

Fulcher. (2007). *Language testing and assessment.* London and New York: Routledge.

Geribka, M. (2016). *A study on the content validity and authenticity of English National Examination 2015 for junior high school.* Thesis. Yogyakarta: English Language Education Study Program, Sanata Dharma University.

Gronlund, N. (1998). *Measurement and evaluation in teaching* (4th ed.). New York: Macmillan Publishing.

Messick, S. (1998). Test validity: A matter of consequences. *Social Indicator Research, 45*(1-3), 35-44.

Nuryanto, M. (2018). *The validity analysis of reading test items on national standard School Final Examination for 12th Grade of MAN 1 Semarang.* Thesis. Graduate Program Magister of Language Study. Surakarta: Universitas Muhammadiyah Surakarta.

Sudjiono. (2011). *Pengantar Evaluasi Pendidikan.* Jakarta: Raja Grafindo Persada.

Sugianto, A. (2016). An Analysis of English National Final Examination for Junior High School in Term of Validity and Reliability. *Journal on English as a Foreign Language p-ISSN 2088-1657 Volume 6, Number 1, 12.*

The copies of English National Examination of Year 2018 for Junior High School