

The Content Validity of Achievement Test of SMK
Tarakanita Jakarta: A Descriptive Study



ELISABETH HANUGRAH HANUNG HAPSARI

2215041349

**A thesis submitted in partial fulfillment of the requirement for
the degree of "Sarjana Pendidikan"**

ENGLISH DEPARTMENT

THE FACULTY OF LANGUAGES AND ARTS

STATE UNIVERSITY OF JAKARTA

2008

Abstrak

ELISABETH HANUGERAH HANUNG HAPSARI. *Validitas Isi Tes Pencapaian Akhir di SMK Tarakanita SMK: Studi Deskriptif.* Skripsi. Jurusan Bahasa dan Sastra Inggris. Fakultas dan Seni. Universitas Negeri Jakarta. 2008.

Tujuan dari penelitian ini adalah untuk mendeskripsikan validitas isi dari test sumatif semester satu dan dua dari SMK Tarakanita. Materi penelitian ini adalah silabus kelas satu dari tahun ajaran 2007/ 2008 SMK Tarakanita dan test sumatif semester satu dan dua tahun ajaran 2007/ 2008 di SMK Tarakanita. Indikator-indikator dari tujuan pembelajaran dianalisis terhadap daftar tujuan pembelajaran yang dibuat berdasarkan tujuan pembelajaran yang terdapat dalam silabus kelas satu tahun ajaran 2007/ 2008 SMK Tarakanita. Soal-soal yang terdapat di dalam tes sumatif semester dianalisis terhadap daftar indikator yang dibuat berdasarkan indikator-indikator yang mengindikasikan tujuan pembelajaran di dalam silabus kelas satu tahun ajaran 2007/ 2008 SMK Tarakanita. Hasil dari penelitian ini menunjukkan bahwa tingkat validitas isi dari kedua tes tersebut rendah yang didukung oleh hasil analisa. Pada tes semester pertama persentase dari validitas isinya sebesar 25 % yang menunjukkan bahwa tingkat validitas isi dari tes tersebut rendah. Pada hasil analisa tes semester ke dua, persentasi dari soal-soal yang terdapat dalam tes sumatif semester ke dua yang validitas isinya valid adalah 38% yang menunjukkan bahwa validitas isi dari tes tersebut sedang. Namun didukung oleh hasil analisa dari indikator tujuan pembelajaran pada semester ke dua, yang menunjukkan adanya beberapa indikator yang tidak mengindikasikan tujuan pembelajaran dengan tepat maka disimpulkan bahwa validitas isi dari tes semester ke dua adalah rendah.

Abstract

ELISABETH HANUGERAH HANUNG HAPSARI. *The Content Validity of Achievement Test of SMK Tarakanita Jakarta: A Descriptive Study.* A Thesis. English Department. Languages and Arts Faculty. State University of Jakarta. 2008.

The objective of this study is to describe the content validity of the first and second semester achievement tests of a vocational school, SMK Tarakanita. The materials of this study are the syllabus of the first grade SMK Tarakanita 2007/ 2008 and achievement tests including summative test of the first and second semester used by the first grade teacher during 2007-2008 in SMK Tarakanita. The indicators in the syllabus are analyzed against a checklist designed based on the learning objectives in the syllabus of the first grade SMK Tarakanita 2007/ 2008. The test items of the achievement test are analyzed against checklist based on the indicators in the syllabus that indicate the learning objectives in the curriculum. The study revealed that the first semester achievement test shows poor content validity as supported by the result of the analysis which percentage of the content validity is 26 % that. On the analysis of the second semester achievement test the percentage of the content validity is 38% which shows an average level of content validity. However, due to the inappropriateness of the indicators in the second semester in framing the learning objectives, the second semester test is considered shows poor content validity.

Acknowledgement

Thanks to:

In the name of Jesus Christ, the Cherish and Sustainer of the world. Let all praises be to Jesus, the Savior of this world. The greatest goes to Allah, who gives His blessing and grace so that the writer finishes to conduct this study.

The writer would like to give her best gratitude, respect, and appreciation to the following persons that have helped and support her during the finishing process of this study.

1. Her parents for their unconditional love and prayers.
2. Mrs. Sri Sumarni as the advisor and academic advisor for the guidance and support.
3. Mr. Ifan Iskandar as the Head of English Department.
4. Mrs. Yani and Mrs. Kipti Head Mistress and Vice Head Mistress of SMK Tarakanita who permits me to do the research in SMK Tarakanita.
5. All of friends who support and love me always.

Table of Content

Abstrak.....	i
Abstract.....	ii
Acknowledgment.....	iii
Table of content	iv

Chapter I

Introduction.....	1
1.1. Background.....	1
1.2. Purpose of the study.....	4
1.3. Identification of problem.....	4
1.4. Research Questions.....	5
1.5. Significance of the study.....	6

Chapter II

Literature Review.....	7
2.1. Studies on the Test Validity.....	7
2.2. Types of Test	10
2.3. Achievement Test.....	11
2.4. Validity.....	13
2.5. Content Validity.....	15
2.6. Taxonomy Bloom.....	16

2.7. Theoretical Framework.....	18
2.8. Definition of Term.....	18

Chapter III

Methodology.....	20
3.1. Research Design.....	22
3.1.1. Material of the study	22
3.2. Data Collection Method	23
3.2.1 Document Analysis.....	23
3.2.2. Instrument.....	23
3.2.3. Procedure.....	23

Chapter IV

Data Description, Analysis and Research Findings.....	25
4.1. Data Description.....	25
4. 2. Analysis.....	30
4.3. Research Finding.....	35

Chapter V

Conclusion and Suggestion.....	37
5.1. Conclusion.....	37
5.2. Suggestion.....	37

References

Appendices

Appendix 1 Syllabus of the First Grade of SMK Tarakanita

Appendix 2 Analysis on the Achievement Test of the First Semester of the First Grade of SMK Tarakanita

Appendix 3 Analysis on the Achievement Test of the Second Semester of the First Grade of SMK Tarakanita

Appendix 4 Achievement Test of the First Semester of the First Grade of SMK Tarakanita

Appendix 5 Achievement Test of the First Semester of the First Grade of SMK Tarakanita

Chapter I

Introduction

1.1. Background

This study is meant to describe the content validity of the first and second semester achievement tests of a vocational school, SMK Tarakanita. Concerning about the consistency of the test items and the indicators of the learning objectives in the syllabus is important in order to provide an appropriate means to measure students learning achievement. Being able to provide such good quality test is obliged by government as stated in the Government Regulation no 19/ 2005. It is stated that the evaluation is meant to gain the information of students' achievement of the competence, the basis to make learning progress report, and to improve the learning process.

Pasal 64

(2) Penilaian sebagaimana dimaksud pada ayat (1) digunakan untuk : menilai pencapaian kompetensi peserta didik, bahan penyusunan laporan kemajuan hasil belajar dan memperbaiki proses pembelajaran.

By the implementation of Kurikulum Tingkat Satuan Pendidikan (KTSP) or School-based Curriculum, the responsibility to designed and deliver evaluation is partly belong to teachers. As stated in the Government Regulation no 19/ 2005, evaluations in elementary and high educations consist of evaluation by teachers, school and government.

Pasal 63

(1) Penilaian pada jenjang pendidikan dasar dan menengah terdiri atas: penilaian hasil belajar oleh pendidik, penilaian hasil belajar oleh satuan pendidikan, penilaian hasil belajar oleh Pemerintah

Stem from this regulation teachers have the responsibility to designed tests that are based on the syllabus which they also design. According to Hughes this approach on developing test is called syllabus-content approach. (Hughes, 2003, p.13). He also states that the disadvantage of this approach is that if the syllabus is badly designed, the result of the test will be very misleading. For that reason, testers that are teachers needs to be very careful in developing the content of the test and the syllabus which are the specification of the content of the test. Brown called such procedures as the procedure evaluating the content validity of a test. (Brown, 2003, p.32).

Despite the importance of teacher to deliver such content valid achievement test, the number of study on language testing is very limited. The lack of study on language test is shown in the number of the research on this topic in English Department of State University of Jakarta. Based on the result of a research on the trends of the researches topic in English Department of State University of Jakarta, the amount of research which topic is language testing is 12 researches out of 503 researches in the department during 2001-2005.

Among the limited researches on language testing, a study on the content validity of the achievement test written by Puspitasari (2006) was done before those government regulations are implemented. The subject of this study is

achievement test for year 2 senior high students develop by MONE Office. This achievement test is developed directly based on the learning objectives written in the curriculum. To investigate the extent of the content validity, the test content was checked against a checklist made based on the test specification which is the learning objectives in the curriculum 1994.

The method used to investigate the content validity in the previous study is no longer valid to analyze achievement tests nowadays. By the implementation of the latest government regulation on national education that gives teachers responsibility to develop the syllabus and achievement test, the achievement test for senior high student is developed using syllabus-content approach. The achievement tests are not developed directly based on the learning objectives but based on the indicators of the learning objectives that are made by teachers and stated in the syllabus. Stated in the socialization of the “Indikator dikembangkan sesuai dengan karakteristik peserta didik, satuan pendidikan dan potensi daerah. Digunakan sebagai dasar penyusun penilaian.” (Diknas 2007). However, direct analysis on the test content validity of the achievement test against the specification of it that is the syllabus, open a risk of misleading result. According to Hughes “The disadvantage of syllabus-content approach is that if the syllabus is badly designed, or the books and other materials are badly chosen, the result of the test can be very misleading.” (Hughes, 2003, p.13). Hence, analysis on the syllabus design is important to be included in the method of analyzing the content validity of the achievement test for senior high students which is developed using syllabus content approach.

Stem from the gap between the importance to develop a content valid achievement test and the limited number of researches on language testing also the weakness of the previous content validity study on achievement test, the researcher is triggered to study the content validity of achievement test for senior high students and include study on the design of the syllabus to see if the achievement test is content valid. The research problems formulated to achieve the objectives of this study are:

1. What indicators in the syllabus indicate the learning objectives with the learning objectives in the curriculum?
2. To what extent the test items of the achievement test are consistent with the indicators in the syllabus which indicate the learning objectives in the curriculum?

Checklist will be the instrument of this study. Each indicator in the syllabus will be analyzed against the checklist based on the learning objectives in the curriculum. The test content will be analyzed against the checklist based on the indicators in the syllabus.

1.2. Purpose of the study

The objective of this study is to consider the consistency of the test items and the learning indicators in the syllabus.

1.3. Identification of problems

The questions that identify the research questions are:

- 1.3.1. To what extent the indicators in the syllabus indicate the learning objectives with the learning objectives in the curriculum?
- 1.3.2. To what extent the test items of the achievement test are consistent with the indicators in the syllabus which indicate with the learning objectives in the curriculum?
- 1.3.3. To what extent the content of the achievement test are consistent with the indicators in the syllabus which don't indicate the learning objectives in the curriculum?
- 1.3.4. To what extent the content of the achievement test are consistent with the indicators in the syllabus which indicate the learning objectives in the curriculum?
- 1.3.5. To what extent the content of the achievement test are consistent with the indicators in the syllabus which don't indicate the learning objectives in the curriculum?

1.4 Research Questions

The research questions of the research answered are:

1. What are the learning indicators stated in the syllabus of the first grade SMK Tarakanita?
2. To what extent the test items of the achievement test are consistent with the indicators in the syllabus?

1.5. Significance of the study

This result of this study will increase teachers' awareness in evaluating their teaching responsibility by considering the content validity of the tests they design.

CHAPTER II

Literature Review

2.1. Studies on Test Validity

Studies on test validation have been done due to its benefit to improve the learning and teaching process in general and the test itself in specific. On 2000, a validity study on test was written by Srijono. The study is aimed to find out to what communicative competence components are assessed in English EBTANAS tests for SMU. It studies on the content of the national final exam of senior high school, included into proficiency test, during 1994-1999 in Indonesia. The data are classified and analyzed based on communicative competence components communicative basic skills. The result of this study shows that the validity of the items in English EBTANAS for SMA is high in testing discourse and grammatical competence and moderate in testing students' communicative competence but low in testing strategic competence.

On 2001, another research on test validity was written by Bintang Martha Ulida. The objective of the study is to examine and evaluate ISF Marlins Test for the Shipping Industry using the Bachman and Palmer model and in particular the research will attempt to evaluate statement about the ISF Marlins test's reliability, construct validity, authenticity, interactivenss, and practicality. It focuses on the proficiency test (ISF Marlins Test) used for shipping industry employment in Indonesia. The data is analyzed from the context of usefulness for each criterion. First of all, the Target Language Use (TLU) domains for the test and on board are set. Then, task and test characteristics are set too. After that, the test usefulness is

evaluated which are done by setting the minimum acceptable level, answering the question for logical evaluation and collect the empirical data for each criteria.

The next two years, on 2005, a validation study is written by Gu Weiping and Liu Juan. This study analyzes College English Test Band 4 (CET-4), which is included to proficiency test, in China. The objective of this study was to examine the validity of CET-4 test by comparing its results with that of a performance-based test, to determine objectiveness of CET-4 in reflecting students' communicative competence and to ascertain the wash back effect of CET on college English teaching. The results showed that although there was no significant difference between the two classes in their CET-4 scores, there was significant difference between them in their experimental test scores. The scores of the treatment class were significantly lower than those of the control class, which indicates that CET cannot objectively reflect students' communicative competence, and thus its validity is low.

At the same year, a validation study on proficiency test is written by Akihiro Ito in Japan. This study focuses on the English language test in a Japanese nationwide university entrance examination. The purpose of this study was to investigate reliability and validity of English question in JFSAT , the goals were to determine if the JFSAT – English test is a reliable and valid measure of students' English ability. The type of validity studied here is criterion validity. Criterion validity was estimated by correlating JFSAT – English test with language ability measure (a carefully constructed close test). The result of the study shows that the paper-pencil pronunciation test should be eliminated and a

listening comprehension test might be included as one of the subsets in the JFSAT – English test. The other subtests, however, showed satisfactory validity.

On 2006, a study was done in order to validate the simulated test of College English Test Band 4. The research was written by Yang Miao in China. The purpose of this study is to check if the simulated test of CET 4 served the specific purpose of predicting and diagnosing test takers' achievement of CET 4. Based on Messick's framework of validation, the test content coverage and representativeness were checked, and correlation analysis including inter-consistency reliability, item correlation, factor analysis and item analysis were computed. To measure the content validity of the test paper was reviewed and checked against the test specification to see its content coverage and representativeness.

In the same year, 2006, Puspitasari wrote a study on the content validity of achievement test, the end of term English written test prepared by MONE office for year 2 SMA students. This research studies on the end of term English written test which is included into achievement test that is made based on the learning objectives stated in curriculum in Indonesia. The data were analyzed using the checklist made based on the reading objectives in 1994 curriculum. From the result of the study it is shown that not all of the reading comprehension questions items in the reading section of the End-of- Term English Written Test prepared by MONE office for year 2 SMA students match to the reading objectives stated in the 1994 curriculum.

The method used to analyze the content validity in the previous study is no longer valid to analyze achievement tests nowadays. By the implementation of the latest government regulation on national education that gives teachers responsibility to develop the syllabus and achievement test, the achievement test for senior high student is developed using syllabus-content approach. The achievement tests are not developed directly based on the learning objectives but based on the syllabus made by teacher. According to Hughes “The disadvantage of syllabus-content approach is that if the syllabus is badly designed, or the books and other materials are badly chosen, the result of the test can be very misleading.” (Hughes, 2003, p.32). Hence, analysis on the syllabus design is important to be included in the method of analyzing the content validity of the achievement test.

2.2. Types of Tests

According to Brown, tests are part of prepared administrative procedures to evaluate students’ performances at time in a curriculum. (Brown, H., 2004, p.4). One of the ways to categorize test based on its purpose or decision to be made based on the result of the test. Bachman states that:

“One way of classifying language tests, therefore, is according to the type of decision to be made. Thus we can speak of selection, entrance and readiness tests with regard to admission decision, placement diagnostic test with regard to identifying the appropriate instructional level or specific areas in which instruction needed, and progress, achievement attainment or mastery test with respect to decision about how individuals should proceed through the program or how well they are attaining the program’s objective.” (Bachman, L., 1990, p.70).

Bailey categorizes tests into eight types of test based on the purpose of the test that are admission test, placement test, progress test, aptitude test, proficiency test, diagnostic test, dominance test and achievement test. (Bailey, 1998, p.40). In

simpler version, Hughes categorizes tests based on the purpose of the test into four types that are proficiency test, achievement test, diagnostic test and placement test. (Hughes, 2003, p.11).

Brown states that a proficiency test is a test that aims to test global competence in a language which is not limited to any one course, curriculum, or single skill in the language. (Brown, 2004, p.44). Bailey defines diagnostic test as a test that is used to more closely identify students' particular strength and weaknesses. (Bailey, 1998, p.39). According to Bachman, placement test is a test designed to measure student language ability in order to group students with similar level of language ability, language aptitude, language needs, and professional or academic specialization. (Bachman, 1990, p.58). The type of test that is studied in this research is achievement test.

2.3. Achievement Tests

According to Grondlund, achievement test is a test which is used to get information on how much students have learnt. (Gronlund, N, 1982, p.1). The information derived from an achievement test is potential to be the basic information to decide which learning objectives students need to learn, though the main purpose of achievement test is to get information towards students' achievement. Brown states that:

“Achievement tests can also serve the diagnostic role of indicating what a student needs to continue to work on in the future, but the primary role of an achievement test is to determine whether course objective have been met – and appropriate knowledge and skill acquired – by the end of a period of instruction.” (Brown, H,2004, p.47)

Because the main function is to measure students achievement towards learning objectives, such test should refer to certain well framed learning objectives. Brown states that: “Achievement tests are (or should be) limited to particular material addressed in a curriculum within a particular time frame and are offered after a course has focused on the objectives in questions.” (Brown, H, p.47, 2004). Based on the school based curriculum socialization the first step to develop a test is determined the indicators that will be the indicator of students achievement toward the learning objectives. (Departement Pendidikan Nasional 2007). Based on the theory on the requirements to limit of the achievement test and the socialization of school based curriculum we can infer that a set of indicators should be determined at first in developing achievement test.

Based on the government regulation on national education no. 19th/ 2005 every institution that runs education should do lesson planning, learning process, learning assessment, and learning process monitoring for the sake of the efficient and effective learning process.

Pasal 17

(2) Setiap satuan pendidikan melakukan perencanaan proses pembelajaran, pelaksanaan proses pembelajaran, penilaian hasil pembelajaran dan pengawasan proses pembelajaran untuk terlaksananya proses pembelajaran yang efektif dan efisien.

It is also explained that lesson planning include syllabus and instruction that contain at least, learning objectives, material, teaching method, learning source and learning evaluation.

Pasal 20

Perencanaan proses pembelajaran meliputi silabus dan rencana pelaksanaan pembelajaran yang memuat sekurang-kurangnya tujuan pembelajaran materi ajar, metode pengajaran, sumber belajar dan penilaian hasil belajar.

From both government regulations it is obvious that school in which teachers are responsible to make the syllabus in which based on the socialization of school based curriculum contains the indicators of the learning objectives. From Brown's theory of achievement test and the government regulation above we can conclude that achievement test refer to a set of indicators that is developed by teachers.

According to Hughes the approach of developing achievement test based on the indicators that are stated in the syllabus is called syllabus-content approach. Hughes states that:

"A final achievement test which content are based directly on a detailed course syllabus or on the books and other materials used refer to syllabus-content approach. The disadvantage of syllabus-content approach is that if the syllabus is badly designed, or the books and other materials are badly chosen, the result of the test can be very misleading. Successful performance on the test may not truly indicate successful achievement of course objective." (Hughes, A, 2003, p.13)

In other words the guarantee of an achievement test which content is developed using syllabus-content approach partly depends on the design of the syllabus.

2.4. Validity

One of the important criteria of an effective test is validity. According to Linn and Miller validity is consideration weather a test is suitable with its use and result interpretation. (Linn, R. L., Miller, M. D., 2005. p. 68). Hughes also

suggests that validity is a matter of giving a consistent measure of the accuracy of a test in measuring the ability meant to be measured which is an indication that a test is ideal for its purpose. (Hughes, A, p.8-9, 2003). Commonly the validity of a test is measured by measuring several evidences. Bachman states that the widely accepted forms of validity evidence are grouped into three general types that are: content relevance, criterion relatedness, and meaningfulness of construct.

Criterion related evidence is used to see the validity of a test compared to another test. Brown states that criterion related evidence of the validity of a test is used to see if the test result of a test is not diverge from the result of a test that is used to measured the same ability which validity is not questionable. (Brown, 2004, p.24). According to Genesee and Uphsur criterion-relatedness evidence is the quality of a test in relation to the quality of another qualified test. (Genesee, F. Uphsur, J, 1996, p.66).

Construct validity is the validity of a test in relation to the theory of the abilities measure in the test. Bachman states that construct validity concerns to the consistency of the test takers' performance with the performance of ability that tester predict based on the theory of the ability (Bachman, L. 1990, p.154). According to Chapelle construct validity refers to the validity that is based on the testers' judgment towards abilities being measured due to the interpretation of the score obtained from the tests result. (Chapelle, C.,1998. P.50).

Other types of evidence that are also the evidences of validity though are not widely accepted are consequential validity and face validity. According to

Brown consequential validity is the validity of tests that is related to any effect of the test toward the test takers and social consequences of the test result interpretation. (Brown, 2004, p.26). As cited by Brown, Mousavi states that

“Face validity refers to the degree to which a test looks right, and appears to measure the knowledge or abilities it claims to measure, based on the subjective judgment of the examinees who take it, the administrative personnel who decide on its use, and other psychometrically unsophisticated observers.”

(Mousavi, 2002, p.244)

2.5. Content Validity

Hughes states that a test is said to have content validity if it sample the language skill, structure, etc that are meant to be measured. (Hughes, A., 2003, p.26). According to Anastasi, content validity is an examination of the test content whether it covers a representative sample of the behavior domain to be measured. (as cited in Weir, C., 1990, p.25).

Hughes states that “A comparison of test specification and test content is the basis for judgments as to content validity.” (Hughes, A., 2003, p.27). Nevertheless, the absence of test specification can be replaced by syllabus since the content of the achievement test is based on the indicators of the learning objectives which are stated in the syllabus.

According to Anastasi, a set of useful guidelines for establishing content validity are:

- 1. The behavior domain to be tested must be systematically analyzed to make certain that all major aspects are covered by test items in the correct proportion.*
- 2. The domain under consideration should be fully described in advance, rather than being defined after the test has been prepared.*

3. *Content validity depends on the relevance of the individual's test responses to the behavior area under consideration, rather on the apparent relevance of the item content. (as cited in Weir, C. 1990, p.25)*

In the newer year, Brown proposed similar procedure of test content validation weather the classroom objectives identified and appropriately framed and lesson objectives are represented in the form of test specification. (Brown, p.32, 2004). From both theories above we can infer that the first procedure to evaluate a test content validity is to identify the extent of the indicators in the syllabus indicate the learning objectives. The next step then is to identify the extent test content assess the indicators in the syllabus which indicate the learning objectives.

2.6.Taxonomy Bloom

According to the socialization of school based curriculum, the indicators of the achievement of the learning objectives/ competences are the signs or the indicators of the learning objectives or the basic competence which is characterized by the changing of behavior that is measurable including attitude, knowledge, and skill. The indicators of the learning objectives are developed based on students' characteristics, schools and regional potency. Indicators are used as the basis to develop evaluation. (Diknas 2007). It s also stated that in order to develop the indicators, each of the competences should be defined into more than two indicators. The verb that is used in the indicator is lower or equal with the verb that is used in the basic competence or standard competence. (Sosialisasi KTSP, Diknas, 2007)

Pengembangan Indikator

“Indikator menggunakan kata kerja operasional yang dapat diukur dan/ atau diobservasi. Tingkat kerja dalam indicator lebih rendah atau setara dengan kete kerja dalam KD maupun SK.” (Diknas, p.132, 2007)

Benjamin Bloom headed a group of educational psychologists who developed a classification of levels of intellectual behavior important in learning. Bloom identified six levels within the cognitive domain, from the simple recall or recognition of facts, as the lowest level, through more complex and abstract mental levels, to the highest order which is classified as evaluation. Verb examples that represent intellectual activity on each level are listed here.

- **Knowledge:** *arrange, define, duplicate, label, list, memorize, name, order, recognize, relate, recall, repeat, reproduce state.*
- **Comprehension:** *classify, describe, discuss, explain, express, identify, indicate, locate, recognize, report, restate, review, select, translate,*
- **Application:** *apply, choose, demonstrate, dramatize, employ, illustrate, interpret, operate, practice, schedule, sketch, solve, use, write.*
- **Analysis:** *analyze, appraise, calculate, categorize, compare, contrast, criticize, differentiate, discriminate, distinguish, examine, experiment, question, test.*
- **Synthesis:** *arrange, assemble, collect, compose, construct, create, design, develop, formulate, manage, organize, plan, prepare, propose, set up, write.*
- **Evaluation:** *appraise, argue, assess, attach, choose compare, defend estimate, judge, predict, rate, core, select, support, value, evaluate.*

Tim PEKERTI – AA PPSP LPP from Sebelas Maret University (2007) design operational verbs that can be used in developing the basic learning objectives in the curriculum that are the basic competence. According to Tim PEKERTI – AA PPSP LPP (2007) the operational verb listed in the book are intended to help teachers to develop the standard competence and the basic competence. The operational verbs designed by Tim PEKERTI – AA PPSP LPP (2007) are not contrasting with the operational verb listed by Bloom. It adds more

verbs under the six classification of cognitive domain competence. In other words, the operational verbs proposed by Blooms are covered in here with additional operational verbs.

2.7. Theoretical Framework

Stem from the gap between the importance to develop a content valid achievement test and the limited number of researches on language testing also the weakness of the previous content validity study on achievement test, the researcher is triggered to study the content validity of achievement test for senior high students and include study on the design of the syllabus to see if the achievement test is content valid. According to Brown the procedure of test content validation weather the classroom objectives identified and appropriately framed and lesson objectives are represented in the form of test specification. (Brown, p.32, 2004). Department of Education states that the indicators of the learning objectives are developed based on students' characteristics, schools and regional potency. Indicators are used as the basis to develop evaluation. (Diknas 2007). It s also stated that in order to develop the indicators, each of the competences should be defined into more than two indicators. The verb that is used in the indicator is lower or equal with the verb that is used in the basic competence or standard competence. (Sosialisasi KTSP, Diknas, 2007)

2.8. Definition of term

The achievement tests in this study are first and second semester summative test designed by SMK Tarakanita teachers to measure students' achievement in the first grade.

The content validity in this study is the extent of the consistency between stem and distracters in each of the test items with the learning indicators in the syllabus.

CHAPTER III

Methodology

The methodology used in this study is descriptive study. This methodology has been employed in studies on the validity of language assessment. Huriah (2003) employed this methodology in her research on the analysis of an achievement test of English based on Bloom's taxonomy of cognitive domain. Puspitasari (2006) also employed the same methodology in her research on the content validity of the reading section of the end of term English written test prepared by MONE office for year 2 SMA students. Knupfer and Mc Lallen (2001) state that descriptive research is critical to educational research, because educational events cannot be reduced to a controlled laboratory environment.

The objective of this study that is to describe the degree of the content validity and the significance of this study that is to enlighten English teachers on the content validity study matches the one of the three purposes of descriptive study. Based on Knupfer and Mc Lellan (2001) the three main purposes of descriptive research are to describe, explain and validate findings. The description often illuminates knowledge that we might not otherwise notice or even encounter. (Knupfer & Mc Lellan 2001).

The research questions of this study suit one of the research questions that are commonly addressed in a descriptive study. According to Knupfer and Mc Lellan (2001) the research question of a study will position the analysis into one of two areas:

1. That which describes the data according to a particular organization (descriptive).
2. That which draws inferences about cause and effect (inferential)

The research problems of the study will be answered by describing the data which are the indicators in the syllabus into its organization that is indicating the learning objectives in the curriculum and the test items which assess the indicators that indicate the learning objectives. From the explanation above we can conclude then that the research question of this study match with the research question of descriptive study methodology answer.

Moreover, the answers of the research questions are presented in the form the percentage of the data with single variable that are the percentage of the indicators indicate the learning objectives and the test items which assess the indicators indicating the learning objectives. Knupfer and Mc Lallen (2001) mention that the answer of a descriptive study might simply report the percentage summary on a single variable.

Theories on the instruments of the descriptive study might require one of instruments that include observation, survey, questionnaires and interview. (Knupfer & Mc Lallen 2001). However some several studies that use descriptive study employ checklist as the instrument of collect the data. A study on the content validity of the reading section of the end of term English written test prepared by MONE office for year 2 SMA students written by Puspitasari (2005) employs a checklist that is made based on the reading objectives of the 1994 curriculum. The same instrument also employed in a descriptive study written by

Huriah (2003). The checklist was developed based on Bloom's taxonomy. So that even though the instrument employed to collect the data is different with the instrument employed in a descriptive study based on Knupfer and Mc Lallen (2001), the instrument of this study match with the instrument employed to collect the data in the previous studies which use descriptive methodology.

3.1 Research Design

The research questions in this study are:

1. What are the learning indicators stated in the syllabus of the first grade SMK Tarakanita?
2. To what extent the test items of the achievement test are consistent with the indicators in the syllabus?

The answer of the first question is important to be used to analyze the consistency of the stem and distracters of the test items with the learning indicators in the syllabus.

3.1.1 Material of the study:

The subjects of this study are the syllabus of the first grade SMK Tarakanita 2007/ 2008 and achievement tests including summative test of the first and second semester used by the first grade teacher during 2007-2008 in SMK Tarakanita.

3.2 Data Collection Methods:

3.2.1 Document Analysis

The documents used are first grade syllabus 2007/ 2008, first semester and second semester achievement tests which are used by teachers. The analysis is focused on the indicators of the learning objectives in the syllabus and the tests items. The analysis on the syllabus is based on the operational verbs in Bloom's Taxonomy.

3.2.2. Instruments:

Checklist designed based on the learning objectives in the Standar Isi Kuriulum Tingkat Satuan Pendidikan.

Checklist designed based on the indicators in the syllabus made by English teachers of SMK Tarakanita.

3.2.3. Procedure:

The indicators in the syllabus are analyzed against the checklist based on the learning objectives in the curriculum.

The content of the achievement test are analyzed against checklist based on the indicators in the syllabus that are relevant with the learning objectives in the curriculum.

Indication and content validity percentage:

$$\frac{X}{\sum X} \times 100\%$$

Legend:

X : the number of the test task which are relevant with syllabus

$\sum X$: the total number of the test task

CHAPTER IV

DATA DESCRIPTION, ANALYSIS AND RESEARCH FINDINGS

4.1. Data Description

Data description towards the first research question is a syllabus of the first grade of SMK Tarakanita 2007/ 2008, achievement test of the first semester of SMK Tarakanita and achievement test of the second semester of SMK Tarakaita. The data analyzed in the syllabus are the operational verbs of the indicators of the learning objectives. The data analyzed from the achievement tests are the tests items.

4.2. Analysis

Syllabus Analysis

KOMPETENSI DASAR	INDIKATOR	
1. <u>Memahami</u> ungkapan-ungkapan dasar pada interaksi sosial untuk kepentingan kehidupan	Siswa dapat: 1. <u>Mengucapkan</u> salam (greetings) pada saat bertemu dan berpisah sesuai waktu dan tingkat formalitas dengan lafal yang berterima say	☆ ☆
	2. <u>Memperkenalkan</u> diri sendiri dan orang lain dengan tepat	☆

	3. <u>Menggunakan</u> ungkapan terima kasih dan responnya dengan tepat	☆
	4. <u>Melakukan</u> percakapan yang menyangkut ungkapan penyesalan dan permintaan maaf serta responnya secara tepat	☆
1.2 <u>Menyebutkan</u> benda-benda, orang, ciri-ciri, waktu, hari, bulan, dan tahun	5. <u>Menyebutkan</u> kata/ungkapan yang tepat untuk mendeskripsikan benda berdasarkan warna, bentuk, asal (<i>origin</i>), ukuran, bahan, jumlah dan kualitas.	☆☆☆
	6. <u>Menyebutkan</u> kata/ungkapan yang tepat untuk mendeskripsikan orang berdasarkan profesi, kebangsaan, ciri-ciri fisik, dan ciri-ciri non fisik.	☆☆☆
	7. <u>Menyebutkan</u> waktu (<i>time of the day</i>), hari, tanggal, bulan, dan tahun dengan tepat.	☆☆☆

<p>1.3. <u>Mendeskripsi</u> <u>kan</u> benda- benda, orang, ciri-ciri, waktu, hari, bulan, dan tahun</p>	<p>8. <u>Menggunakan</u> bilangan (<i>cardinal / ordinal</i>) dengan tepat dalam berbagai konteks.</p>	<p>☆☆☆</p>
	<p>9. <u>Mendeskripsikan</u> benda yang terkait dengan warna, bentuk, asal (<i>origin</i>), ukuran, bahan, jumlah dan kualitas dengan tepat.</p>	<p>☆☆☆</p>
	<p>10. Mendeskripsikan orang yang terkait dengan profesi, kebangsaan, ciri- ciri fisik, dan ciri-ciri non fisik dengan tepat.</p>	<p>☆☆☆</p>
	<p>11. Mendeskripsikan suatu kejadian berdasarkan waktu (<i>time of the day</i>), hari, tanggal, bulan, dan tahun.</p>	<p>☆☆☆</p>
<p>1.4 <u>Memahami</u> ungkapan- ungkapan permohonan dan permintaan serta penawaran untuk kepentingan kehidupan</p>	<p>12. <u>Menyampaikan</u> ungkapan penyesalan dan permintaan maaf serta pemberian responnya dengan tepat.</p>	<p>☆☆☆</p>
	<p>13. <u>Menyampaikan</u> ungkapan simpati serta pemberian respon terhadapnya dengan tepat</p>	<p>☆☆☆</p>
	<p>14. <u>Mengungkapkan</u> perasaan dalam berbagai konteks dengan tepat</p>	<p>☆☆☆</p>

	15. <u>Menggunakan</u> ungkapan yang menyangkut permintaan dan pemberian ijin dengan tepat	☆
	16. <u>Menggunakan</u> ungkapan yang menyangkut perintah dan permintaan secara tepat.	☆
	17. <u>Menggunakan</u> ungkapan yang menyangkut penawaran barang dan jasa secara tepat	☆
1.5 <u>Menjelaskan</u> secara sederhana kegiatan yang sedang terjadi	18. Menceritakan peristiwa yang sedang terjadi sesuai waktu dan tempat kejadian.	☆☆
	19. <u>Mengungkapkan</u> pernyataan dengan menggunakan "there is/are" sesuai waktu dan tempat kejadian.	☆☆☆
	20. <u>Membuat</u> pertanyaan tentang peristiwa yang sedang terjadi sesuai waktu dan tempat kejadian	☆
	21. <u>Mengungkapkan</u> perasaan / pendapat tentang peristiwa yang sedang terjadi dengan tepat	☆☆☆
1.6 <u>Memahami</u> memo dan menu sederhana, jadwal perjalanan kendaraan umum, dan rambu-rambu lalu lintas	22. <u>Menulis</u> memo dengan benar.	☆

	23. <u>Menjelaskan</u> memo yang sudah ada dengan tepat.	☆☆☆
	24. <u>Menulis</u> menu dan menjelaskannya dengan tepat.	☆
	25. <u>Menjelaskan</u> tanda-tanda dan lambang (misalnya: rambu lalu lintas) dengan benar.	☆☆☆
	26. <u>Menyusun</u> berbagai macam jadwal (<i>time table</i>) dan menjelaskannya dengan benar.	☆☆
	27. <u>Menggunakan</u> bentuk kata sifat dan keterangan secara tepat untuk membandingkan sesuatu.	☆
1.7 <u>Memahami</u> kata-kata dan istilah asing serta kalimat sederhana berdasarkan rumus	28. <u>Menggunakan</u> berbagai ungkapan untuk menyatakan pilihan (<i>preferences</i>) dengan tepat	☆
	29. <u>Menggunakan</u> ungkapan untuk menyatakan pengandaian bentuk 1 (<i>conditional type I</i>) dengan tepat.	☆
	30. <u>Menggunakan</u> berbagai ungkapan untuk menyatakan kemampuan (<i>capabilities</i>) dengan tepat.	☆
	31. <u>Menggunakan</u> ungkapan untuk meminta dan memberi arah dan lokasi (<i>direction</i>) dengan tepat	☆
1.8. <u>Menulis</u> undangan sederhana	32. <u>Merangkai</u> sejumlah kata yang mengandung unsure undangan sederhana.	☆

	33. <u>Menyusun</u> kalimat dengan benar untuk membentuk undangan.	★
	34. “ <u>Menyusun</u> ” kalimat dengan benar untuk membentuk undangan	★ ★
	35. <u>Menulis</u> undangan sederhana (misalnya: undangan ulang tahun) dengan benar	★ ★ ★

Data Analysis

Based on the data above, not all of the indicators in the syllabus are appropriately indicate the learning objectives. The indicators with one star are the indicators that indicate higher competences than the curriculum requires in the learning objectives. The indicators with two stars are the indicators that indicate lower competences. The indicators with three stars are the indicators that indicate the competences in the learning objectives appropriately. The indicators that indicate lower or higher competences will be another consideration to measure the content validity of test other than measuring the consistency of the stem and distracters with the indicators in the syllabus.

First Semester Achievement Test Analysis

Learning Indicators	Test Items
Siswa dapat:	-
<u>1. Mengucapkan</u> salam (greetings) pada saat bertemu dan berpisah sesuai waktu dan tingkat formalitas dengan lafal yang berterima	
<u>2. Memperkenalkan</u> diri sendiri dan orang lain	41, 55, 73

dengan tepat	
3. <u>Menggunakan</u> ungkapan terima kasih dan responnya dengan tepat	-
4. <u>Melakukan</u> percakapan yang menyangkut ungkapan penyesalan dan permintaan maaf serta responnya secara tepat	-
5. <u>Menyebutkan</u> kata/ungkapan yang tepat untuk mendeskripsikan benda berdasarkan warna, bentuk, asal (<i>origin</i>), ukuran, bahan, jumlah dan kualitas.	74, 93, 116, 119, 126, 175, 178, 179
6. <u>Menyebutkan</u> kata/ungkapan yang tepat untuk mendeskripsikan orang berdasarkan profesi, kebangsaan, ciri-ciri fisik, dan ciri-ciri non fisik.	55, 73, 76, 81, 87, 94, 95, 103, 112, 114, 181, 182
7. <u>Menyebutkan</u> waktu (<i>time of the day</i>), hari, tanggal, bulan, dan tahun dengan tepat.	40, 189, 174 51, 60, 120 24
8. <u>Menggunakan</u> bilangan (<i>cardinal / ordinal</i>) dengan tepat dalam berbagai konteks.	79, 173, 190, 191, 192
9. <u>Mendeskripsikan</u> benda yang terkait dengan warna, bentuk, asal (<i>origin</i>), ukuran, bahan, jumlah dan kualitas dengan tepat.	12, 63, 70, 82, 113, 139, 166, 168
10. <u>Mendeskripsikan</u> orang yang terkait dengan profesi, kebangsaan, ciri-ciri fisik, dan ciri-ciri non fisik dengan tepat.	89, 100, 122, 176, 196, 199, 200

11. <u>Mendesripsikan</u> suatu kejadian berdasarkan waktu (<i>time of the day</i>), hari, tanggal, bulan, dan tahun.	169,
12. <u>Menyampaikan</u> ungkapan penyesalan dan permintaan maaf serta pemberian responnya dengan tepat.	-
13. <u>Menyampaikan</u> ungkapan simpati serta pemberian respon terhadapnya dengan tepat	-
14. <u>Mengungkapkan</u> perasaan dalam berbagai konteks dengan tepat	143
15. <u>Menggunakan</u> ungkapan yang menyangkut permintaan dan pemberian ijin dengan tepat	-
16. <u>Menggunakan</u> ungkapan yang menyangkut perintah dan permintaan secara tepat.	32
17. <u>Menggunakan</u> ungkapan yang menyangkut penawaran barang dan jasa secara tepat	25, 43

Data Analysis

Based on the table above, 54 test items out of 200 items are consistent with the learning indicators in the syllabus. Some of the test items are not consistent with

any indicator Very view items are consistent with certain indicators. Certain test items are consistent with the indicators that inappropriately frame the learning objectives.

Second Semester Achievement Test Analysis

Learning Indicators	Test Item	Indicator Consistent
18. Menceritakan peristiwa yang sedang terjadi sesuai waktu dan tempat kejadian.	15, 16, 22, 28, 33	x
19. <u>Mengungkapkan</u> pernyataan dengan menggunakan "there is/are" sesuai waktu dan tempat kejadian.	-	x
20. <u>Membuat</u> pertanyaan tentang peristiwa yang sedang terjadi sesuai waktu dan tempat kejadian	-	11
21. <u>Mengungkapkan</u> perasaan / pendapat tentang peristiwa yang sedang terjadi dengan tepat	3, 11,	x
22. <u>Menulis</u> memo dengan benar.	-	x
23. <u>Menjelaskan</u> memo yang sudah ada dengan tepat.	42,43,58	
24. <u>Menulis</u> menu dan menjelaskannya dengan tepat.	-	
25. <u>Menjelaskan</u> tanda-tanda dan lambang (misalnya: rambu lalu	13	

lintas) dengan benar.		
<u>26. Menyusun</u> berbagai macam jadwal (<i>time table</i>) dan menjelaskannya dengan benar.	5	
<u>27. Menggunakan</u> bentuk kata sifat dan keterangan secara tepat untuk membandingkan sesuatu.		
<u>28. Menggunakan</u> berbagai ungkapan untuk menyatakan pilihan (<i>preferences</i>) dengan tepat	14, 2	
<u>29. Menggunakan</u> ungkapan untuk menyatakan pengandaian bentuk 1 (<i>conditional type I</i>) dengan tepat.	-	
<u>30. Menggunakan</u> berbagai ungkapan untuk menyatakan kemampuan (<i>capabilities</i>) dengan tepat.	4	
<u>31. Menggunakan</u> ungkapan untuk meminta dan memberi arah dan lokasi (<i>direction</i>) dengan tepat.	47, 48, 49, 50, 51, 57	
<u>32. Merangkai</u> sejumlah kata menjadi kalimat yang mengandung unsure udangan sederhana.	-	
<u>33. “Menyusun”</u> kalimat dengan benar untuk membentuk undangan	-	
<u>34. Menulis</u> undangan sederhana (misalnya: undangan ulang tahun) dengan benar	-	

Data Analysis

Based on the table above, 23 test items out of 60 items are consistent with the learning indicators in the syllabus. Some of the test items are not consistent with any indicator. Very few items are consistent with certain indicators. Certain test items are consistent with the indicators that inappropriately frame the learning objectives.

4.3 Research Finding

What indicators in the syllabus indicate the learning objectives with the learning objectives in the curriculum?

Not all of the indicators in the syllabus frame or indicate the learning objectives appropriately. Some of the indicators indicate the lower or the higher competences than the required in the learning objectives. The percentage of the indicators in the syllabus of the first grade SMK Tarakanita on 2007/ 2008 that indicate the learning objectives in the curriculum is 41,17 %. 14 indicators from 34 indicators designed to indicate the learning objectives indicate the learning objectives in the curriculum. This finding serves as another consideration to judge the content validity of the tests items, other than to judge the consistency of the test items with the indicators. The test items which are consistent with the indicators that do not indicate or frame the learning objectives cannot be considered as content valid.

To what extent the test items of the achievement test are consistent with the indicators in the syllabus?

The percentage of the test items in the first semester achievement test 2007/ 2008 of first grade of SMK Tarakanita that are consistent with the indicators is 26%. 54 items out of 200 items in the test are consistent with the indicators that indicate the learning objectives. The percentage of the test items in the second semester achievement test 2007/ 2008 of first grade of SMK Tarakanita that are consistent with the indicators in the syllabus is 38%. 23 items out of 60 items in the test are consistent with the indicators.

Based on the analysis, it is founded the first semester test shows poor content validity but the second semester test show average content validity. In the first semester test not all of the test items are consistent with the learning indicators. Also, the distribution of the test items is not equal. Moreover, certain items are consistent with the indicators that do not frame the learning objectives appropriately. In the second semester test, even though more of the test items are consistent with the indicators but certain items are consistent with the indicators that do not appropriately frame the learning objectives. For that consideration the content validity of the second semester test can be concluded as low with respect of the development of the learning indicators.

CHAPTER V

CONCLUSION AND SUGGESTION

This chapter presents the conclusion from data analysis of the content validity of achievement test.

5.1. Conclusion

The requirements of a content valid test are first consistency between the test items and the indicators. Consistent items with the indicators will make sure that the tests really measure what it is meant to measure. The second requirement is that the distribution of the test items should be equal. The number of the test items should be equal among the indicators. Certain indicators may need more items but the difference of the items number should not be significant among the indicators.

The development of the learning indicators could also serve as consideration to measure the content validity of the test. Even though the percentage of a test shows a average or high content validity but if the test is developed based on the learning indicators that do not appropriately frame the learning objectives, the content validity of the test is questionable.

5.2. Suggestion

The first thing that a test designer has to consider is the learning indicators of the learning objectives that are the basis to develop a test. A test designer has to make sure that all of the indicators appropriately frame the learning objectives.

The operational verbs used to develop the learning indicators should indicate the right competences required in the learning objectives. The second consideration is that test designers have to make sure that all of the learning indicators are measured in the test. The distribution of the test items is also important that every indicators is measured with an equal numbers of test items.

It is also suggested to hold similar research for every test designer before they deliver the tests in order to measure the content validity of the achievement test. Knowing the content validity level of the tests before they are delivered gives change for the test designers to fix them when the tests show poor content validity. That way will guarantee the tests to gain representative information towards students' achievement.

In conducting a further research on the related topic it is expected to widen the scope of the study into the other types of validity, such as criterion, construct, consequences and face validity. To research all of the types of the evidences of validity will give more complete picture towards the validity of the test yet the effectiveness of the tests.

References

- Akihiro Ito. (2005). A Validation Study on The English language Test in a Japanese Nationwide University Entrance Examination.
- Bachman, L.F. (1990). Fundamental Consideration in Language Testing. Oxford: Oxford University Press.
- Bailey, K.M. (1998). Learning About Language Assessment. United States of America: Heinle & Heinle.
- Brown, H. (2004). Language Assessment Principles and Classroom Practices. New York: Pearson Education Inc.
- Government regulation no. 19/ 2005 on national education standard.
- Gronlund, N. (1982). Constructing Achievement Tests. New Jersey: Prentice-Hall, Inc.
- Hughes, A. (2003). Testing for Language Teachers. Cambridge: Cambridge University Press.
- Huriah. (2003). An Analysis of an Achievement Test of English Based on Bloom's Taxonomy of Cognitive Domain.
- Peraturan Pemerintah Republik Indonesia 2005 tentang Standar Nasional Pendidikan
- Puspitasari, R. (2006). The Content Validity of the Reading Section of The End of Term English Written Test Prepared by MONE Office For Year 2 SMA students.
- Socialization of school based curriculum. Diknas. 2007.
- Srijono, D. (2000). Evaluation of English Ebtanas Validity.

Ulida, B.M. (2001).The Concept of Usefulness in Test Evaluation .A study of the ISF Marlins Test for the Shipping Industry.

Weiping, G., Liu Juan. (2005). Test Analysis of College Students' Communicative Competence in English.

Weir, C. (1990). Communicative Language Testing. Hertfordshire: Prentice Hall International (UK) Ltd.

<http://www.coedu.usf.edu/itphdsem/411.pdf>

(<http://www.officeport.com/edu/blooms.htm>)