



ORIGINAL ARTICLE

## Analyzing The Content Validity Of Final Exams Of The Course Communicative English Skills

1. Melese Mengesha, 2. Ephrem Gebresilassie

<sup>1</sup>, Department of English Language and Literature, Wollo University, Ethiopia

<sup>2</sup>. Department of Foreign Languages and Literature, Mekelle University, Ethiopia

<b>ARTICLE HISTORY</b> Received: 30.01.2016 Revised 26.03.2016 Accepted 11.05.2016	<b>ABSTRACT</b> <i>Teaching and assessment are principal elements of the entire teaching and learning process. Assessments help teachers to get feedback on the academic progress of students. Besides, it also shows the learners where they need to improve and to plan their own strategies for better future performance. Hence, the researchers assessed the content validity of selected exams for the course Communicative English Skills. The course is delivered in every higher institution such as colleges and universities in Ethiopia. To achieve the objectives, the syllabus, exam papers designed and four English teachers are used as sources of data. The analysis focused on skills, linguistic elements and cognitive components. Hence, it is found that unbalancing the linguistic elements is identified. It is also observed that content relationship is given more attention than cognitive relationship between the syllabus and the exams. In addition to this, inadequacy of exam items is observed during the analysis.</i> <b>Key words:</b> Assessment, coursebook, analysis, content, validity
<b>CITATION OF THIS ARTICLE</b>	Melese Mengesha, Ephrem Gebresilassie. Analyzing The Content Validity Of Final Exams Of The Course Communicative English Skills. Inter. J. Edu. Res. Technol. 7[2] 2016; 15-21 <b>DOI:</b> 10.15515/ijert.0976 4089.7.2.1521

### INTRODUCTION

According to Bachman (1990), Alderson (1991), Weir (1993) and Teshome (1995), assessment has a significant part in a teaching- learning process since it is an integral part of the task. It offers assistance for teachers to follow up the academic achievement of their students. In addition, the information that teachers deduce from exams help them to check a curriculum and improve it.

Educational route is a platform in which one course of action comes after the other, and in each step there is refinement process which is assessment. As a foundation whereupon judging one's academic achievement is made, assessment plays a tremendous role in such a step-by-step improvement of a person's academic career. Hence, the position of assessment in an educational achievement has continued being the most essential instrument (Holmes et al 2000, Elliot et al 1996).

To ensure quality of education and assess carefully the students' academic achievement, critically designed evaluation tools contribute a lot. However, numerous investigations clearly depicted that wrong utilization of measuring instruments has significant effect on learners' educational accomplishment and their overall scholarly advance. Meyers (1988) affirms that regardless of how contemporary data has been dispersed, the academic accomplishment of learners can never be as viable of course as needed without the utilization of well-organized and professional measuring devices.

Assessment should not be carried out just for the purpose of grading learners. However, pedagogically, sound teaching always involves solid assessment of the behavioral change and academic progress of learners. In Ethiopia, where quality of education is a burning issue, developing and using apposite measuring devices is extremely helpful. The benefits of such sort of devices are superb to the extent language competence is concerned. But evaluation has been earned a little consideration. Since quality of

education is the center of discussion nationwide, such kind of research is priceless to speed up and bring radical improvement in the academic endeavor in the nation.

Academicians like Gronlund and Linn (1990) insert that an authentic exam exhibits at least three major qualities such as validity, reliability and practicality. Validity refers to the appropriateness and usefulness of the test to judge the performance of a learner on a certain course or training. It is the degree of extent that a test measures what is supposed to be measured. Reliability, on the other hand, refers to whether a test accurately and consistently evaluates students' performance. Accordingly, there are three ways of checking reliability: split half procedures, test-retest and alternative form reliability.

The third quality of an authentic test is practicality which entails the extent of economical, easy to be administered and scored, as stated by Hughes (1989).

Despite the fact that validity, reliability and practicality play significant role to have effective tests, the present research is concerned with validity, specifically content validity, of the selected exam papers. According to Weir (1993), a test and its items are supposed to measure the achievement of the course objectives. Hence, validity is considered as a significant frame of reference for exam designers to develop effective exams.

Scholars such as Bachman (1990) state that validity is further divided into four subdivisions such as face validity, concurrent validity, predictive validity and content validity. The present research focuses on the content validity of the selected exams. Content validity emphasizes on the exam's ability to adequately representation of all of the contents of a particular coursebook or syllabus in order to have a wash-back effect on the teaching and learning process (Bachman, 1990; and Weir, 1993).

Brown as cited in Tirussew et al (1992:16) also defined the word as:

*...an inference to be made about a student's degree of attainment of the universe of situations and or subject matter domain. The test behavior serves as a sample and the important question is whether the test constitutes a representative sample to the content domain of interest.*

This indicates that content validity focuses on the extent of an exam to cover the contents incorporated in the syllabus (curriculum) of a given course. In order to measure content validity, it is necessary to quantify the frequencies of practice items in the syllabus (coursebook) and exam items in the exam papers. Alemu (1983), quoting Oller (1979), states that analyzing language contents can be quantified by counting certain language elements in the discourse.

Generally, assessments of a right content validity support the teaching and learning process to a great extent by encouraging learners to plan their own strategies for better future performance. Accordingly, the feedback from the result of such assessments helps the teachers, students and material designers to view progresses in the academic endeavor and the effectiveness of a program and curriculum.

The above discussion implies that teachers who develop exam items should give great concern to establish the content validity of their exams because contents of exams should reflect what has been incorporated in the syllabus of the given course and the activities practiced in classroom. Local researchers such as Tibebe (1992), Kifle (1995), Asmare (2008) and Alemu (1983) have conducted a research on validity aspect of tests. Tibebe (1992), for example, studied the predictive validity of ESLCE (Ethiopian Secondary School Leaving Certificate Examination) English exams of the academic year 1987/88 compared to two locally designed proficiency tests while Kifle (1995) studied the content validity of grade ten English language tests with reference to English for New Ethiopia textbook. Asmare (2008) has also investigated the content validity of English language tests at Hawasa College of Health Science. Alemu (1983) on the other hand has conducted research on the national examinations for grades six and eight where one of his specific objectives was content validity of the exams.

This study focuses on analyzing the content validity of the final exams for the course Communicative English Skills in higher institutions in Ethiopia for the academic year 2014, 2015 and 2016. This is on the ground that content validity is the most important part of validity because it gives helpful insight to judge whether an assessment adequately samples the course objectives (Hughes, 1989). Again, focusing on final or summative exam is logical in the sense that such type of exams are expected to cover all the contents of the course (syllabus) as they are given at the end of a semester or term. Moreover, a common final exam is prepared by all teachers for all students of a particular institute.

The studies of the aforementioned scholars focused on the validity (Content and Predictive) of tests and national exams both for high school and primary school students. Their findings in general imply that there is weak content validity in the tests they analyzed. However, they did not investigate the case of final exams at university. Hence, this paper deals with this gap.

In language education, the greater a test's content validity, the more likely it is to be accurate measure of what it is supposed to measure. If an exam doesn't consider the language components incorporated in the

coursebook and if the items are poorly constructed, it will be irrelevant and unfair to measure the achievement of the testees. Additionally, the purpose of assessment in language teaching and its significance for language teachers was not given due attention for many decades. The consequence of this reality is designing poor quality assessment devices. However, the introduction of communicative language teaching has helped teachers to be considerate in preparing effective assessments because the modern education requires to pay due attention to content validity of exams and teaching. Besides, the researcher has found that researches concerning content validity of summative exams, are extremely thin at Wollo University. Furthermore, Girma (1997:46) proved that content validity of exams has not been given consideration in the Ethiopian situation. This may imply that exams are negatively influencing the teaching and learning process. Bi-implication, this may have been influencing the quality of education in general. Assessment can also assist the language teacher by revealing the academic achievement of his/her students as the results offer him/her lots of information on the effectiveness of his/her teaching. Therefore, designing valid exams in terms of content is unquestionably triggering issue in the Ethiopian context.

In language instruction, there are two major issues namely the content of the language and the targeted instructional objectives (outcomes) to be considered in assessment. However, practically content or knowledge is majorly assessed than proficiency in skills. With regard to this, Heaton (1988:7) states:

*Linguists are examining the whole complex system of language skills and patterns of linguistic behavior. Indeed, language skills are so complex and so closely related to the total context in which they are used as well as to many non-linguistic skills (gestures eye – movements, etc) that it may often seem impossible to separate them for the purpose of any kind of assessment.*

The communicative era requires teachers to integrate the four major skills to realize genuine classroom practices. On the basis of this reality, the assessment items that teachers design should be relevant to enable student use the language in the real social situation communicatively. This implies that designing exams need to give attention on assessing the application of these skills in the real life context. That is why Heaton (1988:8) notes ‘...it is usually extremely difficult to separate one skill from another....’

Additionally, assessing learners’ language competence requires integrating different cognitive components such as contents, skills and linguistic elements because these domains are expected to be attained at the completion of a course (Nigussie, 2002). All in all, the content of an assessment is determined by the contents of the syllabus, texts, instructional materials and classroom activities.

Even though the effectiveness of an assessment can be influenced by numerous reasons, the major one is the capability of the items to be adequate samples for the language contents of the coursebook and classroom tasks. Therefore, the research, in particular, was planned to answer the accompanying questions:

1. Are the utilized exams items really authentic enough to assess language knowledge and skills defined by the objectives of the coursebook in an appropriate proportion?
2. Do the exam items match with the expected learning outcomes of the coursebook?
3. To what extent does their content coverage reflect the contents in the coursebook?
4. What are the qualities and shortcomings of the selected exams?
5. What are the potential factors for the existence of such weaknesses in the exams?

### **Objectives of the Study**

This research has an objective of analyzing the content validity of the assessments of the course Communicative English Skills (EnLa 1011). Specifically, this research focuses on:

- Analyzing the authenticity of the exams.
- Identifying the relationship between the exam items and the learning outcomes of the course.
- Revealing the adequacy of the assessment to represent the contents of the coursebook.
- Identifying the strength and weakness of the assessments.
- Identifying the factors affecting the quality of the assessments.

## **MATERIALS AND METHODOLOGY**

### **Design of the Study**

Since content validity is rational or logical validity, and since the research is highly dependent on the logical understanding and view of the sample exams, the research is more concerned with logical analysis. Besides, the research made a comparison and contrast between the contents of the selected exams and the coursebook of the course Communicative English Skills (EnLa 1011).

### Sources of Data

The researcher used diverse sources such as the syllabus of the course, sample exam papers prepared for first year students in the three consecutive years – 2014, 2015 and 2016 and four English language teachers who teach the course in Kombolcha Institute of Technology as sources of the needed data for the completion of this research work. The teachers were selected using convenient sampling technique from the eight teachers who teach English in the campus. Unstructured interview was employed to gather data from the sample teachers. However, the exam papers were randomly selected.

### Tools of Data Collection

The researcher has employed document (exam) analysis and unstructured interview to collect the needed data.

The researcher took Communicative English Skills final exam papers administered in the last three consecutive years in the institute. The exams are used as sources of primary data because the exams are the main focus of the study. So as to analyze the different facets of the exams subtly, objectivity, arrangement, diversity, discrimination power, appropriateness, and representativeness of the sample exams, were seriously explored. The problems exhibited in the exams were traced based on such various pedagogical parameters. During the analysis, the relationship between the skills coursebooks and the extent of considering them in the exam papers was emphasized. The degree of the relationship between the content of the exams papers and the syllabus was computed using Pearson correlation by comparing the four skills, linguistic elements and cognitive domains.

Next to the content analysis, the researcher conducted unstructured interview with the sample teachers. The researcher chose unstructured interview as they will be more careful and reserved in the formal interview.

## RESULTS AND DISCUSSIONS

The analysis of the exams focused on the skills, linguistic elements and cognitive components that are encompassed in the preparation of the exams. The syllabus for the course Communicative English Skills (EnLa 1011) has been taken from the newly harmonized curriculum which is revised and designed in September 2013 by experts from the whole universities found in Ethiopia for higher education institutions. The main aim of the syllabus is to enable the learners use the language spontaneously and flexibly in academic and social environment communicatively through listening, speaking, reading and writing.

The result of the analysis of the final exams shows that reading occupies the major area covering 28.6% in the syllabus while speaking, vocabulary, listening, grammar and writing cover 20%, 14.3%, 14.3%, 11.4% and 11.4% respectively (See Table 1). This implies that reading is given major emphasis than the other skills.

The analysis of the distribution of skills and linguistic elements in the final exam also show that reading has received the major attention covering 63% and 30.8% of the items in the final exam of the academic years 2014 and 2015. Hence, reading is ranked first in the 2014. On the other hand, grammar is given more attention and ranked first comprising 30.8% and 33.3% in the final exams of the academic years 2015 and 2016 respectively (See Table 2). This data clearly depicts that grammar is considerably overemphasized in the final exams while it is ranked fourth in the coursebook. In contrary, the figure indicates that listening and writing skills are totally neglected in all the final exams. Besides, even though speaking comprises 20%, which is second in the coursebook, it is remarkably neglected covering only 11.2% of contents of the exams on average. The vocabulary part, which covers 14.3% of the syllabus, is a bit overemphasized in the exams covering 21.2% of the sample exams. On the other hand, reading is given almost equivalent attention (33.3%) with that of the coverage in the syllabus (28.6%). The result of the interview also yielded similar outcome. Many interviewees indicated that grammar covers the largest part in the exams while listening is totally ignored.

Hence, balancing the linguistic elements is feeble in the final exams of the course since grammar is overemphasized while listening and writing skills are completely neglected in the exams.

Concerning the distribution of cognitive components in the syllabus, it is found that the syllabus emphasizes on developing majorly (39%) the application skills of the students while due emphasis is also given to comprehension (24.4%) and building the knowledge (19.5%) of the students. However, least emphasis (7.3 and 2.4%) is given to analysis and evaluation schemes. This figure indicates that the course is designed majorly to develop the application and comprehension skills of the learners. On the other hand, analysis, evaluation and synthesis are least emphasized cognitive domains in the course syllabus for the course Communicative English Skills (See Table 3).

The researchers have also traced the similar consideration of the cognitive components in the final exam papers like the syllabus. As a result, it is found that recalling (knowledge) skills and comprehending skills of the students has been given due attention in the final exam. This implies that there is contradiction between the contents of the syllabus and that of the final exams of the course. More significantly, the exams didn't contain any item that measures the evaluation and synthesizing skills of the learners in the final exams of all the academic years. Additionally, analysis is totally neglected in the final exams of the year 2015 (See Table 4). Hence, the majority of the items in the final exams measure the learners recalling skills.

The analysis and the ideas of the interviewees indicate that content relationship is given more attention than cognitive relationship between the syllabus of the course and the exams. As a result, it can be deduced that there is weak relationship between the course syllabus and the exams.

In order to make it more tangible, the researcher used Pearson correlation relation to show how the exam items are strong enough to represent the content covered in the syllabus.

As a result, it has been found that the relationship between the syllabus and the final exams is delicate. The correlations in the content (skills and linguistic components) elements of the years 2014, 2015 and 2016 are 0.6856, 0.4784 and 0.1579 respectively. This correlation indicates that the items in the final exams are not as strong as required to measure the competency of the learners. Hence, there is low content correlation between the syllabus and the final exams. According to this figure, only 68.56%, 47.8% and 15.8% of the contents of the exams of the academic years 2014, 2015 and 2016 respectively are similar with that of the syllabus. This implies that 32.44%, 52.2% and 84.2% of the items (questions) included in the final exams of the years 2014, 2015 and 2016 respectively are different from the contents of the syllabus (see Table 5).

Towards the difficulty level and adequacy of the items of the exams, most of the interviewees believe that the questions in the final exams are moderate in terms of difficulty but inadequate. The researchers have also found that the number of items of the exams of 2014 and 2015 are inadequate to represent the components of the syllabus. This indicates that the exams have weak discrimination power of distinguishing clever students from the weaker ones. It is also identified that there is redundancy of exam items in the final exams.

When the interviewees were asked about the factors for the deficiency of the exams, they forwarded lack of time to think thoroughly to plan and design smart exams because of being busy on doing some routine tasks, fluctuation of class and exam schedules, repeatedly change of academic calendar of the university, readiness of the students to face well-crafted and challenging exams, and shortage of accessing input such as computer, printing and photo copying machine as main threats.

Table 1: The distribution of Skills and Linguistic elements in the coursebook

No	Skills and linguistic elements	Frequency	Distribution (%)
1	Reading	10	28.6
2	Speaking	7	20
3	Vocabulary	5	14.3
4	Listening	5	14.3
5	Grammar	4	11.4
6	Writing	4	11.4
Total		35	100

Table 2: The distribution of Skills and Linguistic elements in the sample final exams in 2014, 2015 and 2016

No	Content	2014		2015		2016		Average	
		No of items	%	No of items	%	No of items	%	No	%
1	Reading	17	63	8	30.8	8	17.8	11	33.3
2	Vocabulary	2	7.4	6	23.1	11	24.4	7	21.2
3	Grammar	8	29.6	8	30.8	19	42.2	11	33.3
4	Speaking	0	0	4	15.4	7	15.6	3.7	11.2
5	Listening	0	0	0	0	0	0	0	0
6	Writing	0	0	0	0	0	0	0	0
Total		27	100	26	100	45	100	33	100

Table 3 Distribution of cognitive components in the syllabus

No	Cognitive components	Frequency	Distribution (in %)
1	Knowledge	8	19.5
2	Comprehension	10	24.4
3	Application	16	39
4	Analysis	3	7.3
5	Evaluation	1	2.4
6	Synthesis	3	7.3
Total		41	100

Table 4 distribution of cognitive elements in the Communicative English final exams

No	Cognitive components	Distribution (in %)							
		2014		2015		2016		Average	
		Freq.	%	Freq.	%	Freq.	%	Freq.	%
1	knowledge	7	28	14	46.67	12	24	11	31.4
2	Comprehension	5	20	10	33.33	16	32	10.33	29.5
3	Application	7	28	6	20	14	28	9	25.7
4	Analysis	6	24	0	0	8	16	4.67	13.3
5	Evaluation	0	0	0	0	0	0	0	0
6	Synthesis	0	0	0	0	0	0	0	0
Total		25	100	30	100	50	100	35	100

Table 5: Correlation relation between contents of the exams and the course guidebook

No	Academic year	content correlation	cognitive correlation
1	2014	0.6856	0.6983
2	2015	0.4784	0.6058
3	2016	0.1579	0.8257

### CONCLUSIONS AND RECOMMENDATIONS

As it has been discussed above, the analysis was based on the input gained from the subject teachers and the analysis made by the researchers. Pedagogically, sound teaching always involves solid assessment of the behavioral change and academic progress of learners. It is also obvious that the position of assessment in an educational achievement has continued being the most essential instrument. Hence, the academic accomplishment of learners can never be as viable of course as needed without the utilization of well-organized and professional measuring devices.

To realize quality education, using quality assessment tools is indispensable. An exam is supposed to cover the contents included in the syllabus, activities practiced and discussed in the classroom. The linguistic elements (skills) include the four macro skills and the micro skills such as grammar and vocabulary. On the other hand, the cognitive components which are devised by Benjamin Bloom include knowledge, comprehension, application, analysis evaluation and synthesis. The analysis of the exams reveals that there is loose relationship between the contents of the syllabus and the final exams in terms of cognitive components. Thought application is given the priority in the course syllabus, the final exams prioritized knowledge and comprehension. Moreover, evaluation and synthesis are totally neglected in the final exam. Besides, grammar is overemphasized in the final exams in contrast to its position in the course syllabus. Furthermore, writing skills and listening skills are not considered in the entire final exams.

As a result, it is time to pay attention on the content validity of the exams of Communicative English Skills in higher institutions. It is important to assure the positive correlation between the contents of the syllabus and the contents of assessments that teachers design and utilize to assess the academic progress of their students. In addition to this, practices and activities done in a classroom should be considered in order to make the assessment valid and to the required standard.

### ACKNOWLEDGEMENT

We really want to express our heartfelt gratitude to the teachers who were willing to participate in the interview. In addition, we would like to thank Mekelle University and Wollo University for the financial and technical support for the completion of this research.

**REFERENCES**

1. Alemu, T. 1983 "Assessment of Grades Six And Eight English National Examinations" Unpublished M.A. Thesis. Addis Ababa: Addis Ababa University.
2. Bachman, L.F. 1990. *Fundamental Consideration in Language Testing*. Oxford: OUP.
3. Bloom, B. (ed.) 1956. *Taxonomies of education objectives*. New York: Mckay.
4. Elliot, N. et al 1996. *Educational Psychology*. Boston: McGraw-Hill companies.
5. Girma, L. 1997. "Representativeness and Adequacy of Physics and Mathematics ESLCE items" *Institute of Educational Journal*. Vol. IV. Addis Ababa University.
6. Heaton, J.B. 1988. *Writing English Language Test*. New York: Longman.
7. Hughes, A. 1989. *Testing For Language Teachers*. Cambridge. Cambridge University Press.
8. Kifle, K. 1995. "An Assessment of Content Validity of High School English Language Tests." Unpublished M.A. Thesis. Addis Ababa University.
9. Nigussie, T. 2002. "The Content Validity of the EGSEC English Examination" Unpublished M.A. Thesis. Addis Ababa University.
10. Nuru, M. 1992. "Level of Questions: A Description of Textbook and Examination Questions in Higher Secondary Schools" Unpublished M.A. Thesis. Addis Ababa University.
11. Oller, J.W. 1979. *Language Tests at School*. London: Longman.
12. Teshome, D. 1995. "The Construction and Validation of Tests in English for Tertiary Education." Unpublished Ph.D. Dissertation. Addis Ababa University.
13. Tirusseaw T. et al (1992) "A Compressive Approach to the Ethiopian School Leaving Certificate Examination." **Institute of Educational research, Addis Ababa University.**
14. Weir C.J. 1990. *Communicative Language Testing*, London: Prentice Hall internal Ltd.