Practices of Assessing Graduate Students' Learning Outcomes in Selected Ethiopian Higher Education Institutions

Desalegn Chalchisa
Addis Ababa University, Ethiopia

Abstract

The study explored the practices of assessing graduate students' learning outcomes by instructors of higher education institutions (HEIs) in Ethiopia. Survey method was employed in collecting data from 131 instructors, which were selected with stratified sampling. Questionnaire and interview were used as instruments for collecting data on the instructors' practices of types and frequency of using assessment strategies, types of test items in practice and use of assessment for improving learning. The data from these three variables were analyzed in terms of demographic factors: bands or fields of study, instructors' academic rank and experience in teaching in HEIs. The results showed that there were no significant differences among instructors from different bands or field of study in the type and frequent use of assessment strategies, in their practice of different types of tests items and use of assessment results for improving learning. Also no significant mean differences were observed in teaching experiences in HEIs. However, significant differences were observed among the mean scores of the above three variables by academic rank with better use of assessment strategies, different types of test items by lecturers than professors and associate professors. It was concluded that professors and associate professors should pay attention to the assessment of students learning and further study need to be conducted for identifying why professors and associate professors were paying less attention to student assessment than lecturers.

Introduction

Assessment has long been recognized as maintaining a central position in students' learning (Craddock and Mathias, 2009). Mode of assessment can also have a powerful influence on the learning behaviour of students (Biggs and Tang, 2007) and assessing the performance of students is one of the most important activities of instructors (Trotter, 2006). Offering a variety of assessment methods is often recommended as good practice in response to numerous critiques of the over-reliance on traditional tests. The arguments include the need to use strategies which more appropriately assess different kinds of learning processes, the need to cater for differences in students' learning preferences and styles and the need to enhance learners' psychological approaches to learning (Pellegrino, Chudowsky and Glaser, 2001; Furniss, 2003; Connolly, 2004; Struyven, Dochy, &

Janssens, 2005).). Studies (Miller, Linn & Gronlund, 2009; Norton, 2007; Race, 1998; Ebel & Frisbie, 1991) showed that assessment is one of the most important activities of higher education instructors. Instructors in higher education used different strategies in order to assess students' learning outcomes. Broadly they can be categorized as traditional type of tests (selective type and supply type tests) and performance and problem- based assessments (portfolios, project work, work- based learning, reviews and annotated bibliographies, self -and peer assessment and group work). Each of these strategies has their pros and cons. Many traditional types of tests encourage surface approach to learning where the intention is to memorize and rehearse – this is passive learning (Norton, 2007). On the other hand, performance and problem- based assessments are argued to encourage deep approach to learning where the intention is to make sense of the subject in terms of understanding and prior knowledge –very much an active, transformative and constructivist approach to learning (Norton, 2009; Ebel and Frisbie, 1991). In higher education institution, especially at a graduate level, designing assessment tasks that reward deep rather than a surface approach to learning and creating innovative and authentic types of assessments are encouraged.

According to Cartwright, Weiner and Veneruso (2009), outcome-based assessment in higher education has three stages: (1) defining the most important goals for students to achieve as a result of participating in an academic experience (outcomes); (2) evaluating how well students are actually achieving those goals (assessment); and (3) using the results for improving the academic experience.

The main purpose of learning assessment at classroom level is to ensure that students know what and how well they are learning. One of the reasons for students' assessment of learning outcomes is to provide feedback for students and teachers about learners' progress in order for both be able to improve the efficacy of their work (UNECO, 2006). In addition, it provides feedback to educators, parents, policy makers, and the public about the effectiveness of educational services (Pellegrino, Chudowsky and Glaser, 2001; Stephens and Moskowitz, 2004).

The Ethiopian Education and Training Policy (TGE, 1994) states in article 3.31 that: "Continuous assessment in academic and practical subjects, including aptitude tests will be conducted to ascertain the formation of all round profile of students at all levels" (p.18). In the implementation of this policy, the Federal Democratic Republic of Ethiopia (EDRE) further promulgated Higher Education Proclamation in 2009. Among other things, the policy covers important issues related to students' assessment in Article 22 where institutional quality enhancement through student evaluation and assessment and other means of quality enhancement strategies were stressed (FDRE, 2009, p. 4988-4989). Furthermore, Article 41 (EDRE, 2009, pp.5005-5006) provides mode of students' assessment in HEIs. In particular, Sub-articles 41.4-7 provide details of assessment of students including fairness in assessment and competency- based assessment (Sub-article 41.4), presence of institutionally recognized and well-defined student assessment and examination methods and systems at academic unit levels (Sub-article 41.5), the need

to establish "adequate structures, at the necessary levels of the pertinent academic units, which shall constitute leading authorities at any of the levels on student assessments, examinations and grading of results" (p.5006 Sub-article 41.6) and how to handle student complaints against assessment methods, examination and grading (Sub-article 41.7, p.5006).

Assessment of students' learning outcomes is one of the focuses of Ethiopian universities senate legislations. In order to harmonize the Ethiopian HEIs academic policies including student assessment, currently, the Ministry of Education (MOE) has issued the "Harmonized Academic Policy of Ethiopian Public Higher Education Institutions" (MOE, 2003). This harmonized policy emphasizes assessment as one of the duties of instructors both at undergraduate and graduate programs.

Statement of the Problem

HEIs in Ethiopia are currently implementing the modular approach of students learning and instruction, which requires lists of manageable learning outcomes that serves as the basis for measuring students' achievement (Curriculum Design and Delivery BPR-Implementation Working Group, 2009). The rational for introducing modular instruction in Ethiopian higher education is that it combines advantages of separate instructional innovations, such as performance objectives, self-pacing, active learning and frequent feedback. It is also argued that the modular approach is "competency-based type of curriculum-which stresses identification of professional/vocational skills, job-specific skills and transferable skills a graduate may have after completing the curriculum" (Curriculum Design and Delivery BPR-Implementation Working Group, 2009, p.3). The modular approach (Curriculum Design and Delivery BPR-Implementation Working Group, 2009) suggests three modes of delivery: interactive (where the instructors introduce the module, provide materials, assignments and guidance for self-learning through discussion and lecture methods), independent learning (where students learn independently based on the materials, guidelines and assignments that they had been given) and cooperative learning (where students meet in small groups to undertake group assignments, dialogues/debates, paper presentations, or book reviews). The Curriculum Design and Delivery BPR-Implementation Working Group (2009) suggested that the assessment of students in the modular delivery should be on continuous basis in relation to achievement of the modular-objectives (criterion-referenced) with a passing standard of 50 percent. However, informal feedback from the students show that the assessment of students is not continuous and far from the suggested strategies. Teachers continue to use conventional methods of assessing students' outcome with a claim that their assessment is continuous with the application of different strategies.

The assessments of students learning outcomes consist of a sequence of activities from planning to marking. Out of these activities, writing appropriate test items, providing feedback to improve the students learning as well as instruction takes the major portion of

instructors' time and demands knowledge and skills of students' assessment (Nitko, 1996; Linn & Gronlund, 2000). The quality of such assessments provides solid evidences for the achievement of the learning outcomes as well as the use of sound assessment practices (Colbert, Wyatt-Smith, & Klenowski, 2011).

Practices of assessment strategies influence the quality of teaching and learning. The types of strategies practiced by instructors and how frequently they practice them have an impact on the quality of students learning (Kelleghan & Greany, 2001; Hale &Astolfi, 2011). Studies also show that the type of test items practiced by instructors should match with the learning outcomes. Teaching in the graduate programs demands by its nature higher level of thinking such as analysis, synthesis and evaluation. The learning outcomes should be stated at these higher levels of thinking. The type of test items that match with these levels should be carefully selected and implemented.

Another important issue is the use of assessments is that they should guide students learning and teachers' activities. In this regard studies suggest that there should be a balance between formative and summative assessments so that they serve their intended purposes (Miller, Gronlund & Linn, 2009; Yorke, 2003; Mehrens & Lehmann, 1991). However, most often instructors' assessment strategies in Ethiopian higher education institutions (HEIs) were criticized as dominated by summative evaluation without due attention to formative evaluation (Daniel, 2004).

So far, as to the knowledge of the researcher, no study was conducted on the assessment strategies, type of test items practiced by graduate instructors and the use of assessment for improving student learning in the Ethiopian context. This study was designed to address these issues in the context of Ethiopian HEIs.

Objectives

The general objective of the study was to critically analyze the practices of instructors in assessing students' learning outcomes in selected HEIs in Ethiopia. The specific objectives of the study were to:

1. determine the assessment strategies that instructors practiced in assessing the learning outcome of students in HEIs;
2. identify the type of test items instructors use in assessing graduate students' learning outcomes in HEIs;
3. explore how instructors use the results of assessment of graduate students learning outcomes to improves students learning and instruction;
4. analyze the alignment of the methods of students' assessment with the learning outcomes; and
5. evaluate the appropriateness of the methods employed to assess students in HEIs and thereby to suggest the effective methods for better students' learning.

Basic Research Questions

The basic research questions that guide this study were:

1. What types of assessment strategies did instructors practice in assessing the learning outcome of students in HEIs?
2. What were the types of test items instructors use at the graduate level in assessing the learning outcome of graduate students in HEIs?
3. For what purposes did instructors use the result of assessment of graduate students learning outcomes in HEIs?
4. Were the methods of assessment of graduate student learning outcomes in HEIs aligned with the learning outcomes?
5. Were the methods of assessment of graduate student learning outcomes in HEIs appropriate?

Significance of the Study

Assessment of students' learning outcomes is one of the most controversial issues in higher education today. Most of the studies on the assessment of students' learning outcomes in Ethiopia were conducted at the primary and secondary levels focusing on the status and challenges of continuous assessment. This study is unique as it was conducted on the instructors' practices of assessment of students learning outcomes at the higher education institutions level. The study would be useful for instructors of Ethiopian HEIs providing feedback on the assessment strategies, types of test items and use of assessment in results for improving graduate students learning in Ethiopian HEIs. It would be also useful in influencing the quality of the teaching-learning process in HEIs, particularly for giving attention to improving the quality of students' learning.

The study also contributes for the leadership of the HEIs in Ethiopia by providing feedback on the instructors' practices of the assessment of students learning outcomes. It can also provide an important input for the researchers who wish to conduct research on assessment of students in HEIs.

Review of Literature

A number of writers (Kennedy, 1999; James, Mcinnis, & Devlin, 2002; ETS, 2003; Furniss, 2003; Stayhorn, 2006; UNESCO, 2006; Greaney & Kellaghan, 2008; Clarke, 2012) defined assessment in a more or less the same way. According to Kellaghan (2001), assessment is "the process of obtaining information to make educational decisions about students, to give feedback to the students about his or her progress and strengths and weaknesses, to judge instructional effectiveness and curricular adequacy and to inform policy" (p.19). Clarke (2012) also classified the types of assessment of learning outcomes into three: large scale survey assessment, examinations and classroom assessment. On

the other hand, Stephens and Moskowitz (2004) classified types of assessment into four: individual low stakes (classroom assessment), individual high stakes (examinations), system low stakes (assessment for monitoring and benchmarking) and system high stakes (assessment for accountability).

Classroom assessment includes a variety of activities, tools and procedures for collecting and interpreting written, oral, and other forms of evidence on student teaching (Clarke, 2012). A number of studies (for instance, Pellegrino, 2001; Kazin & Payne, 2009; Hernandez, 2012) showed a strong link between effective classroom assessment activities and better students' learning outcomes.

According to Jackson, Wisdom and Shaw (2003), "Intended learning outcomes are what students will know and be able to do as a result of engaging in the learning process. They represent statements of achievement expressed from the learners' perspective" (p.1). Similarly, Kellaghan and Greany (2001) stated that learning outcome is a statement of what a learner is expected to know, understand and be able to do at the end of a period of learning. Learning outcomes are linked to the relevant level, and they should generally be assessable and be written in terms of how the learning is represented. In this regard, Kennedy (1999) noted that learning outcomes should be matched with the method of assessment to the different kinds of learning outcomes in order to appropriately test their achievement.

At HEIs level, students' learning should focus on knowledge gained, skills and abilities acquired and demonstrated, and attitudes or values changed (Cartwright, Weiner, & Veneruso, 2009). These are the challenging to assess and may require knowledge and skills of the development and use of assessment strategies and the collection of reliable and valid data (Struyven, Dochy, & Janssens, 2005; Young & Ontario, 2005; Struyven, Dochy, & Janssens, 2002; Cartwright, Weiner & Veneruso, 2009).

Assessment plays important roles in providing information for decision making in education. Describing the level of student learning, identifying and diagnosing learning problems planning further teaching/learning and motivating students are among the roles of classroom assessment (Kellaghan & Greany, 2001). Bearing this in mind, learning outcomes have become the components of module or course design in HEIs.

Nitko (1996), Linn and Gronlund (2000) and UNESCO (2006) identified a number of assessment approaches. The most pertinent to the assessment of students' learning outcomes are the criterion-versus norm-referenced approach and the summative versus formative approaches. According to UNESCO (2006) norm-referenced assessment can be appropriate when the objective is to make comparisons between learners. Criterion-referenced assessment is most appropriate for assessing more complex learning and performance outcomes. The use of criterion-referenced approach provides information on a learner's level of achievement.

The practices of norm-referenced or criterion- referenced approaches vary from institution to institutions. Some universities (Addis Ababa University and Jimma University) clearly specified in their legislation the approach of students' assessment. For

instance, Jimma University (2011) stated that "all assessments must be carried out on the principle of outcome - based learning (must know and be able to do)" (p.8), which is a criterion-referenced approach. On the other hand Mekelle University (n.d.) assessment policy suggests the use of norm and criterion-referenced testing approaches as appropriate.

Another approaches to assessment are formative and summative approaches (York, 2002, Yorke, 2003; Craddock & Mathias, 2009; UNESCO, 2006; Hernandez, 2012). Formative assessment is concerned with how judgments about the quality of student responses (performance, pieces, or works) can be used to shape and improve the students' competence by short-circuiting the randomness and inefficiency of trial-and-error learning (Sadler, 1989, p.120). On the other hand, summative assessment is a judgment, which encapsulates evidences up to a given point usually at the end of a program, a course or end of unit of instruction (Black & Wiliam, 1998; Knight, 2002; Yorke, 2003; Taras, 2005; UNESCO, 2006; Craddock & Mathias, 2009; Hernandez, 2012).

Assessment strategy is any means utilized to gather data for assessment and evaluation purposes (Hale and Astolfi, 2011). The assessment strategies suggested by Hale and Astolfi (2011) in their student learning strategies survey were (a) take-home examinations, (b) in-class closed book examinations, (c) in-class open-book examinations, (d) objective examinations, subjective examinations, (e) unit or mid-term examination (f) cumulative final examination, (g) non-cumulative final examination, (h) individual student examinations, (i) group examinations, (j) individual or group portfolios of student work, (k) individual term paper or project and (l) group term paper or project. UNESCO (2006) also described similar lists of assessment strategies for assessing knowledge, skills, attitudes and behavior of students.

Fisseha (2010) stressed the need to use assessment strategies such as performance assessment, portfolios, authentic assessment and student self- and peer assessment together with feedback and comments, which are based on constructivist theories of learning and teaching. He further suggested that teachers should be responsible for providing feedback that students need in order to re-learn and refine learning goals. Moreover, Yohannes (2006) identified assessment strategies and their functions as follows.

> Knowledge and understanding is often assessed through written examination. Independent project work or essays are typically used to demonstrate capability in a range of intellectual skills linked to subject-specific knowledge, understanding and practical skills. Transferable skills, such as written communication skills can be assessed through essays or report writing; oral communication skills can be assessed through a range of activities including group or individual presentations in seminars. Team working skills can be assessed through collaborative projects (pp.8-9).

The use of effective assessment strategies by Ethiopian HEIs instructors is questionable (Daniel, 2004). According to him:

> Assignment in the form of projects, and term papers are limited in quite many programs. The types of examinations given seem to encourage students to

memorize facts. The evaluation practice is generally summative. There is little or no habit of planning a test particularly with the aim of improving teaching (p.78).

Daniel (2004) further noted that the primary purpose of assessment of students' learning outcomes is to improve student learning. In practice, however, assessments in many programs of HEIs are poor in terms of giving feedback and in motivating further learning.

Assessment of students' learning outcomes demands from instructors at least three activities, which this study explored. First, the use of variety of strategies such as tests, assignments, individual and cooperative projects, portfolios, and other pertinent strategies to the assessment of learning outcomes expected to be achieved as a result of learning. Second, the test items also should be related to the learning outcomes to meet the needs of students and the validity of the assessments. Third, the result of assessment should be used to improve students' learning.

Methods

Design of the study

The approach of this study is both qualitative and quantitative. The concurrent QUAN and qual design were used to collect data from instructors. The survey method with the use of questionnaire was the dominant method of data collection. Qualitative data were collected using interview in order to substantiate the quantitative data.

Sources of Data

Primary and secondary sources were used for the study. The primary sources were collected from instructors of higher education institutions and the secondary sources were the assessment policies and legislations of the sample universities.

Participants

The participants of the study were instructors from higher education institutions. Six universities namely: Adama, Addis Ababa, Bahir Dar, Jimma, Hawassa and Mekelle universities were included in the study. These universities were selected purposely because of their long experiences and the possibility of getting data in the bands/fields of study. Accordingly, they were selected to represent Engineering and Technology, Social Science and Humanities, Natural and Computational Sciences, Medicine and Health Sciences, Agriculture and Life Sciences, and Business and Economics bands/fields respectively. The bands/fields are common to all universities at undergraduate level, but at the graduate level the focus of each university differ on their experiences since their establishments. For instance, Adama University was established as a technical college and was upgraded to a university that focuses on engineering and technology. The same is true for almost all the universities included in this study except for Addis Ababa University,

which currently focuses on graduate programs in all the bands. Table 1 shows the population and sample of instructors by field of study from the sample universities. The data for total population of instructors in each band/field of studies were obtained from the Ministry of Education Annual Abstract (2012).

Table 1: Population and sample sizes

University	Band/Field of study	Population size		Sample size	
		F	**M**	**F**	**M**
Adama	Engineering and Technology	15	356	2	31
Addis Ababa	Social Sciences and Humanities	50	272	6	38
Bahir Dar	Natural and Computational Sciences	122	982	1	15
Hawassa	Agricultural and Life Sciences	15	124	2	11
Jimma	Medicine and Health Sciences	45	297	4	26
Mekelle	Business and Economics	22	139	2	12
Total		**269**	**2170**	**17**	**133**

Table 1 shows the population and sample size by university and fields of study. A total of 150 instructors (17 females and 133 males) were selected to participate in this study from six universities. The population includes all instructors who were teaching both at graduate and under graduate programs. It should be made clear that an instructor could teach only at the undergraduate program or both at the graduate and undergraduate programs. In this study, those instructors who teach at the graduate level were included. The sample size of female instructors teaching at the graduate level is very small since the enrolment of female students in Ethiopian HEIs was also very small, which as a consequence affects the number of female instructors at Ethiopian HEIs.

Instruments

Two instruments were used in this study: instructors' questionnaire and interview guides. The questionnaires were adapted from Hale and Astolfi (2011) of Strategies Used for Assessing Student Learning Survey. The first of the questionnaire requires the respondents to specify their personal characteristic and the second part of the questionnaire was about the assessment practices of instructors with three major subsections. In part two, the first subsection measures types and frequency of assessment strategies, the second part measures types of test items practiced and the third part measures use of results of assessment. Each subsection was a five point scale with Never (1), Rarely (2), Often (3), Frequently (4), and Very Frequently (5).

Procedures

The instruments were adapted from Hale and Astolfi (2011) with modifications to the context of the practices of students' assessment in HEIs in Ethiopia. The interview guide

was prepared to enable the collection of data that complement the data collected through questionnaire. Important input from PhD students of the department of Curriculum and Teachers Professional Development were solicited to ensure the content validity of the questionnaire and interview guides. Furthermore, the questionnaire was pilot-tested on instructors that were not participated in the final study. The result of the pilot test showed that the Cranach Alpha reliability index for the whole instrument was 0.86 and for the three subsections of the questionnaire: types and frequency of assessment strategies used by instructors, types of test items used by instructors and use of results of assessment were 0.79, 0.64, and 0.84 respectively. Item-total correlations were used for selecting those items with correlation coefficients of 0.15 and above, which is an acceptable point-biserial value for the selection of an item to be included in a questionnaire (Ferguson & Takane, 1989). Hence, items with correlation coefficients of greater than or equal to 0.15 were selected for final use.

The data was collected by PhD students and instructors from the college of education with half day training on the instruments and data collection procedures. Three sub-scores were calculated from the three sub-scales: scores for the types of strategies and frequency subscales, types of test items used by instructors subscale and use of the result of assessment subscale. These sub-scores were used as dependent variables in conjunction with the independent variables: fields of study/bands, instructors' academic rank and work experiences as instructor in HEIs. The raw scores obtained for each subscale from all sample universities were converted to T-scores (a scaled/standard score with mean of 50 and standard deviation of 10) for the purpose of analysis. The mean and the standard deviation of the T-Scores vary by universities and field of study from the preset mean of 50 and standard deviation of 10 and the T-Scores were computed for the sum of scores from all universities rather than for individual universities. Therefore, the result of significant tests for comparing the mean scores of universities or field of study is the same whether raw scores or T-Scores are used. But T-Scores have the advantage of putting the scores on the same scale making comparisons of the mean scores meaningful.

Method of data analysis

The data coded and entered into SPSS. Percentages, mean scores and analysis of variance were used as appropriate. Percentages and mean scores were used to find out types and frequency of assessment strategies, types of test items practiced, and use of the results of assessment. One way analysis of variance (ANOVA) was used to explore mean score differences among instructors with different academic ranks and experiences and among bands/fields of study with respect to types and frequency of assessment strategies, types of test items and use of assessment result to improve students' learning. Qualitative data were obtained through interview, which was organized according to the major themes included in the quantitative analysis. The responses on interview guides were analyzed in terms of the appropriateness of the assessment strategies and types of test items used to assess graduate students learning outcomes. The appropriateness criteria for the

assessment strategies are variety, balance and the use of assessment results for improving learning.

Results

In this section, the quantitative data were organized to answer the basic questions. The analysis was carried out for three dependent variables: types and frequency of assessment strategies, types of tests items used by instructors and use of assessment results in improving student learning with demographic characteristics that the researcher believed to be pertinent. Qualitative data were used to substantiate the findings of quantitative data.

Types and frequency of assessment strategies

In order to assess the instructors' practices of assessing students' learning outcomes, eleven items were used to assess the types and frequency of assessment strategies. Examples of the item include individual projects, group projects, term paper, student portfolios, short essay, long essays, etc. Table 2 presents the summary of the most frequent ratings and means ratings of each item on types and frequency of assessment strategies.

Table 2: Types of assessment strategies employed by instructors

Item	Most frequent rating	%	$\overline{X}$
Computational problems	Rarely	23.0	2.73
Individual project(s)	Frequently	33.3	3.57
Group project(s)	Frequently	25.2	4.33
Short essay	Frequently	29.8	2.88
Long essay	Very frequently	21.0	4.77
Individual presentations	Frequently	29.8	3.30
Group presentations	Frequently	25.2	3.41
Student lab books	Never	41.2	2.08
Portfolios	Never	60.5	1.38
Individual term paper	Frequently	40.5	3.60
Group term paper	Very frequently	30.5	4.25

The data from Table 2 suggest that instructors at the Ethiopian HEIs were using frequently individual projects, group projects, individual presentations, group presentations, short essay, group term papers, and individual term papers in assessing graduate students learning outcomes. However, very frequently instructors use long essay and group term papers. Instructors rated student lab and portfolios most frequently as "Never" indicating that they were least used assessment strategies.

The qualitative data shows there is variation in the use of different types of strategies

based on the instructors' decision. In this regard, instructor "F" said:

> The assessments of the different courses depend on the objectives of the courses. Some instructors use tests and final examinations while others use a combination of different tools. However, there are some instructors who use examinations and take them out of 100 and grade students using student's relative position. It all depends on the instructors' decision.

The responses of instructors on the types and frequency of assessment strategies subscale were scored and their raw scores were converted to T-score. Descriptive statistics of types and use of assessment strategies were presented by university (band/field of studies) in Table 3. ANOVA was used to analyze the mean differences in instructors' ratings of the types and frequency of assessment strategies by band of study (Table 3).

Table 3: Mean and standard deviation of Types and frequent use of assessment strategies

University	Bands/fields of study	N	$\overline{X}$	s
Adama	Engineering and Technology	16	52.73	12.93
Addis Ababa	Social Sciences and Humanities	34	49.78	9.00
Bahir Dar	Natural and Computational Sciences	20	46.45	8.17
Hawassa	Agricultural and Life Sciences	17	47.34	11.02
Jimma	Medicine and Health Sciences	23	52.55	9.02
Mekelle	Business and Economics	18	51.19	10.42
Grand mean		**128**	**50.00**	**10.00**

Table 3 indicates that the mean scores of instructors' ratings for Adama Bahir, Jimma and Mekelle universities were higher than the grand mean, a mean of the combined scores of the sample universities. This indicates 50 percent of the instructors from these bands rated the types and frequent use of assessment strategies higher than the grand mean. This may suggest that Adama and Mekelle universities use better testing strategies than other universities. Table 4 below presents the summary of ANOVA for mean differences in the mean rating of the types and frequency of strategies practiced by instructors in the assessment of students learning outcomes by sample universities instructors.

Table 4: Summary of ANOVA for strategies of assessment of learning outcomes

Source of variation	SS	df	MS	F	p
Between Groups	668.20	5	133.64	1.36	0.25
Within Groups	12031.80	122	98.62		
Total	**12700.00**	**127.00**			

The overall F-test (Table 4) for the mean rating of instructors on the types and frequency of assessment of learning outcomes was not significant ($F_{(5, 122)} = 1.36$, p = 0.25). This shows that instructors from the sample universities practice similar assessment

strategies in assessing graduate students learning outcomes. In agreement with the quantitative data, the qualitative data shows instructors use variety of instruments in assessing students' learning outcomes. One of the interviewed instructors, instructor-A, said that:

> In my courses I use variety of instruments including review of scientific articles, individual presentations and term papers, which contribute 50 percent for the final grading of students. The remaining 50 percent for grading students is resulted from mid- and final examinations and I decide the final grading on norm- referenced approach.

Both the quantitative and qualitative data from instructors on types and frequency of assessment strategies show that not all instructors used variety of instruments in assessing students' learning outcomes. Some instructors viewed that the final grade of students is based on one shot final examinations. This indicates the assessment strategies of sample universities lack balance.

Types of test items practiced

Table 5 shows the type of test items used by instructors in assessing students learning outcomes. The Table presents the most frequent rating of each type of test items and their corresponding mean ratings.

Table 5: Types and frequency of test items for assessing learning outcomes

Item	Most frequent rating	%	$\overline{X}$
Multiple choice	Never	31.9	2.72
True-False	Never	41.6	2.40
Matching	Never	37.3	2.47
Short Answer	Frequently	26.0	4.06
Short/limited response Essay	Frequently	31.1	3.57
Long/extended response Essay	Frequently	32.2	3.65
Computational problems	Never	38.2	2.17
Oral tests	Never	46.6	2.14
Individual performance checklists	Never	27.4	2.61
Group performance checklists	Never	30.6	2.48

* The rating scales were: 1= never, 2= rarely, 3= often, 4 = frequently, 5= very frequently

As indicated in Table 5, instructors used short answer items and short and long essays in their assessment of students' learning outcomes. The traditional selection types of test items were never practiced. The use computational problems, oral test, and individual and group performances checklists most frequently were also rarely practiced although these types of test items are useful for assessing higher order cognitive and affective learning outcomes.

Qualitative data from instructors showed that most commonly they use essay tests in assessing students' learning outcomes. Selective types of tests (true/false, matching and multiple choice items) and short answer items were reported to be used less frequently. Instructor "G", for instance, said: "in Chemistry courses essay type items such as long essay and work out questions are common. There are no objective type questions as it used to be in the undergraduate program". On the other hand, an instructor "H" mentioned: "I use most commonly essay tests essay tests (most common type) and sometimes short answer and objective types such as true/false, multiple choice and matching".Table 6 shows descriptive statistics of types of tests items used by instructors for measuring student learning outcomes.

Table 6: Mean and standard deviation of types of test items by band/field of study

University	Bands/fields of study	N	$\overline{X}$	s
Adama	Engineering and Technology	13	52.46	7.94
Addis Ababa	Social Sciences and Humanities	34	47.46	13.20
Bahir Dar	Natural and Computational Sciences	20	44.98	6.43
Hawassa	Agricultural and Life Sciences	17	53.06	8.19
Jimma	Medicine and Health Sciences	23	54.50	6.81
Mekelle	Business and Economics	18	49.97	9.66
Grand mean		**125**	**50.00**	**10.00**

On the types test items for assessing students learning outcomes, the descriptive scores of Medicine and Health Sciences and agricultural and Life Sciences bands were $\overline{X}$ = 54.50 with standard deviation s = 6.81 and $\overline{X}$ = 53.06 with standard deviation s = 8.19 respectively. Both these bands and the Engineering and Technology band ($\overline{X}$ = 52.46 and s = 7.94) were rated the use of different types of test items higher than the grand mean, a mean when the ratings scores of all universities combined. The other bands or fields of study were rated the use different types of test items lower than the grand mean. This indicates 50 percent of the instructors from these bands rated the use of different types of test items higher than the grand mean. In Table 7 ANOVA was used to analyze the mean differences in the mean ratings of types of test items practiced in assessing students learning by instructors from the bands/fields of study.

Table 7: Summary of ANOVA for types of test items for assessment of learning outcomes

Source of Variation	SS	df	MS	F	p
Between Groups	671.49	5	134.30	1.35	0.25
Within Groups	11809.97	119	99.24		
Total	**12481.456**	**124**			

As observed from Table 7, ANOVA has resulted in mean differences in the use of different types of test items, which were not statistically significant ($F_{(5, 119)} = 1.35$, p= 0.25). This finding suggests that there was no significant difference among instructors from different bands in the use of types of test items. Most instructors decide on the type of test items used for assessing learning outcomes. In the absence of assessment policies and guidelines, instructors seem to use those test items that are easy to deal with and takes less time in writing and scoring.

Use of test items to improve students' learning

The use of assessment results in improving students learning is one of the major purposes of students' assessment of learning outcomes. In this regard, a five-point scale containing six items was prepared and administered to instructors. Table 8 shows the most frequent rating for each item in this scale.

Table 8: Use of assessment results for instruction

Item	Most frequent rating	%	$\overline{X}$
Revise what has been taught	Frequently	25.6	3.25
Provide remedial instruction	Rarely	30.2	2.21
Provide enrichment instruction	Rarely	39.3	2.41
Improve methods of teaching	Frequently	25.6	3.35
Motivate students learning	Frequently	25.6	3.23
Assign grades	Frequently	46.5	3.62

* The rating scales were: 1= never, 2= rarely, 3= often, 4 = frequently, 5= very frequently

As shown in Table 8 except the item "Provide remedial and enrichment instructions" all items were rated most frequently as "Frequent".   This shows that instructors used the results of assessment of students learning outcomes for revising, improving the methods of teaching, motivating and grading students.

The responses to the interview, instructors revealed that assessment of students' learning outcomes were used for the intended purpose to improve instructions as can be inferred from the following response of one of the interviewed instructor- instructor "B".

I use students' results as means to help students' learning. It helped me to improve my methods of presentation. I feel that the assessment I use has made students learn better and this contributes its share to improve the quality of education.

In agreement with the above opinion another instructor –instructor "C" described how instructors use the results of students' assessment of learning outcome in improving students learning as follows.

What I have learned from my teaching experience is that when students

are given different tasks and values for what they have done, it made them learns more and better so that it in turn contributed to the improvements of students' learning. This has been observed when students have been involved in their final year project work. They showed improved ability and skills in identification of project titles, review of scientific articles, power point preparation, presentations, use of time, and quality of performance in the field.

Complying with the guidelines of continuous assessment demands the provision of remedial and enrichment instructions, which their practices were rated most frequently as "Rarely". Table 9 presents the descriptive statistics for the use of the result of assessment of learning outcomes for improving students learning and instruction.

Table 9: Mean and standard deviation of use of assessment results by band/field of study

University	Bands/fields of study	N	$\overline{X}$	s
Adama	Engineering and Technology	12	50.19	13.55
Addis Ababa	Social Sciences and Humanities	34	50.15	9.10
Bahir Dar	Natural and Computational Sciences	20	46.78	8.26
Hawassa	Agricultural and Life Sciences	17	47.69	11.14
Jimma	Medicine and Health Sciences	23	52.95	9.12
Mekelle	Business and Economics	18	51.57	10.54
Grand mean		**124**	**50.00**	**10.00**

The data in Table 9 show that the mean rating for the use of the results of assessment for improving learning and instruction was higher than the grand mean for the instructors from the bands of Medicine and Health Sciences ($\overline{X}$ = 52.95 and s = 9.12), Business and Economics ($\overline{X}$ = 51.57 and s = 10.54), Engineering and Technology ($\overline{X}$ = 50.19 and s = 13.55) and Social Sciences and Humanities ($\overline{X}$ = 51.57 and s = 10.54. This indicates 50 percent of the instructors from these bands rated the use of results of assessment for improving instruction higher than the grand mean. Blow Table 10 presents ANOVA for the statistical differences in these mean ratings.

Table 10: Summary of ANOVA for use of the results of assessment of learning outcomes

Source of Variation	SS	df	MS	F	p
Between Groups	544.06	5	108.81	1.09	0.37
Within Groups	11755.94	118	99.63		
Total	**12300.00**	**123**			

As depicted in Table 10, the results showed that there were no overall significant differences in the mean ratings of instructors on the use of the results of assessment of from the six bands. The instructors from different bands used result of assessment of

students learning outcomes for the similar purposes.

Assessment Practices by academic rank and teaching experience

Table 11presents the descriptive statistics for the types and frequent use of assessment strategies, types of test items and use of assessment results by academic rank.

Table11: Mean and standard deviation of assessment of students learning outcomes by academic rank

Variables	Academic Rank	N	$\overline{X}$	s
Types and frequent use of assessment strategies	Lecturer	37	53.47	8.31
	Assistant Professor	60	49.77	8.39
	Associate Professor and above	31	46.27	13.27
	Grand mean	**128**	**50.00**	**10.00**
Types of test items	Lecturer	36	53.33	6.54
	Assistant Professor	58	49.07	11.03
	Associate Professor and above	30	47.52	10.60
	Grand mean	**124**	**50.00**	**10.00**
Use of assessment results for improving learning	Lecturer	35	50.21	9.54
	Assistant Professor	58	49.90	9.68
	Associate Professor and above	30	49.88	11.53
	Grand mean	**123**	**50.00**	**10.00**

With respect to type and frequent use of assessment strategies the mean score of instructors at the rank of lecturer was the highest ($\overline{X}$ = 53.47, s = 8.31) as shown in Table 11. Conversely, the mean score of associate professors and above ($\overline{X}$ = 46.27, s = 13.27) was the least in the use of assessment strategies to measure learning outcomes at the graduate level.

Similar results were obtained with type of test items used by instructors. The mean scores of lecturers were the highest ($\overline{X}$ = 53.33, s = 6.54) and that of the associate professors and above was the least ($\overline{X}$ = 47.52, s = 10.60). The data showed that differences among instructors with different ranks on the use of assessment for improving students learning, although finding is consistent with the other two variables. In general, the current study found that as academic rank of instructors increases their mean scores of assessment practices decreases.

Table 12: Summary of ANOVA for assessment of learning outcomes by academic rank

Variables	Source of Variation	SS	df	MS	F	p
Types and frequent use of assessment strategies	Between Groups	879.23	2	439.62	4.61	0.01
	Within Groups	11919.57	125	95.36		
	Total	**12798.80**	**127**			
Types of test items	Between Groups	634.50	2	317.25	3.28	0.04
	Within Groups	11694.01	121	96.65		
	Total	**12328.51**	**123**			

F-test was conducted to check whether or not there is significant differences exist between the mean scores of instructors on the type and frequent use of assessment strategies by academic rank. The result was statistically significant, which would mean that instructors with a rank of lecturers most frequently employed variety of assessment strategies ($F_{(2,125)} = 4.61$, $p = 0.01$) and type of test items ($F_{(2,121)} = 3.28$, $p = 0.04$) than instructors with the rank of associate professors and above. The use of Tukey- test as further post-hoc analysis detected significance differences between the mean scores on frequent and types of strategies of assessment of lecturers and associate professors and above to be significant. Similarly, the mean scores on type of test items of lectures were significantly greater than that of associate professors. One possible explanation for such differences is that as the rank of instructor becomes high, involvement in administrative commitments and university affairs increases, which possibly affect the instructor's involvement in the teaching and learning in general and student assessment in particular. The other reason could be instructors at the lower academic rank are more motivated for academic growth than instructors at the higher academic rank. Due to these reasons lecturers use various assessment strategies than professors.

Table 13 presents the descriptive statistics for the three assessment scales namely: types and frequent use of assessment strategies, types of test items and use of assessment results for improving learning by instructors' teaching experiences in HEIs.

Table 13: Mean and standard deviation of assessment of students learning outcomes by teaching experience

Variable	Teaching experience in years	N	$\overline{X}$	s
Types and frequent use of assessment strategies	10 and below	61	50.65	8.23
	11-19	27	51.29	9.77
	20-29	27	47.91	12.52
	30 and above	9	46.08	9.18
	Grand Mean	**124**	**50.00**	**10.00**
Types of test items	10 and below	60	51.42	10.11
	11-19	26	50.32	10.08
	20-29	25	46.90	8.80
	30 and above	9	47.28	10.76
	Grand Mean	**120**	**50.00**	**10.00**
Use of assessment results for improving learning	10 and below	60	49.88	8.90
	11-19	26	50.44	11.04
	20-29	24	50.52	11.32
	30 and above	9	45.46	10.49
	Grand Mean	**119**	**50.00**	**10.00**

On the types and frequent use of different assessment strategies, mean scores of instructors with teaching experiences of below 19 years were greater than that of below 20 years as well as the grand mean ($\overline{X}$ =50.00 and s = 10.00). Similarly, the mean scores of instructors with 19 years and below experiences in teaching in higher HEIs on type of test items is greater than those instructors with teaching experiences of 20 and above and the grand mean. This indicates 50 percent of the instructors with teaching experiences 19 years and lower rated types and frequent use of assessment strategies and types of test items higher than the grand mean. With respect to use of assessment results for improving learning the mean scores of instructors with teaching experiences between 11-29 years were greater than the grand mean whereas mean scores of instructors with teaching experiences 10 years and below and 30 years and above were less than the grand mean.

Table 14: Summary of ANOVA for assessment of learning outcomes by teaching experience

Variables		SS	df	MS	F	p
Types and frequent use of assessment strategies	Between Groups	324.84	3	108.28	1.15	0.33
	Within Groups	11300.15	120	94.17		
	Total	**11624.99**	**123**			
Types of test items	Between Groups	429.44	3	143.15	1.46	0.23
	Within Groups	11353.56	116	97.88		
	Total	**11783.00**	**119**			
Use of assessment results for improving learning	Between Groups	193.03	3	64.34	0.64	0.59
	Within Groups	11548.92	115	100.43		
	Total	**11741.95**	**118**			

The F-tests for the mean scores of instructors with different experiences showed no significant differences on the types and frequent use of assessment strategies, types of test items and use of assessment results for improving the students learning.

Conclusion

The study used the survey method to collect data from instructors about the types and frequent use of assessment strategies, types of test items practiced and used of assessment results in improving students learning. Questionnaire and interviews were used to collect data that complement each other on the practices assessment in HEIs. The data were analyzed and reported using appropriate quantitative statistics and qualitative data analysis protocols.

The study showed that instructors emphasized long essay and group related assessment strategies in assessing students learning outcomes with less emphasis for computational problems, student lab work and portfolio. Individual and group projects, individual term and group term papers were the appropriate strategy that instructors reported practicing. These strategies are in alignment with the modular course delivery approach that is now a trend in Ethiopian universities. However, group projects and term papers sometimes become the product of an individual or few individuals in the group rather than that of the group as a whole. In this respect, strengthening individually produced assessment practices pays off better return as a measure of students' assessment of learning outcomes than group related assessment practices.

Currently, the emphasis given for the summative evaluation and formative evaluation in the harmonized higher education policy and included in the senate legislations of universities in Ethiopia is 50 percent each. Improvement in students learning and as a result better quality of education requires more weight should be given to formative assessment (Pellegrino, Chudowsky & Glaser, 2001; Daniel, 2004; Craddock & Mathias, 2009; Abera, 2012). The weight of 60 percent allocation for formative assessment and 40

percent allocation for summative assessment suggested in Procedures of modularization and block teaching by Addis Ababa University (Curriculum Design and Delivery BPR-Implementation Working Group (2009) sounds an appropriate strategy of assessment at graduate level for enhancing better student learning.

Conventional types of test items such as true/false, multiple choice and matching items were most frequently reported as never used by instructors. This is also positive aspect that is appropriate at graduate level, as in general such items tend to measure lower level learning outcomes. On the other hand, it was found that instructors rated the use computational problems, oral test, and individual and group performances checklists most frequently as "never" used, although these types of test items are useful for assessing higher order cognitive and affective learning outcomes, which a number of researchers have suggested their appropriateness at the graduate level (Yohannes, 2006; Craddock & Mathias, 2009; Fisseha, 2010; Hernandez, 2012).

With respect to the use of assessment for improving learning, it was learnt that instructors use the results of assessment of students learning outcomes for revising, improving methods of teaching, motivating and grading students. On the other hand, the use of assessment for remediation and enrichment is very much limited indicating the need for strengthening the use of assessment in these areas.

Finally, instructors with the rank of lecturer were found to practice variety of assessment strategies and test items than instructors with rank of associate professors and professors. As teaching is the most priority of any instructor of HEIs, associate professors and professors should have to play exemplary roles in the assessing students learning outcomes. The availability of assessment policies is very essential in order to guide the assessment activities instructors. Therefore, universities are required to have their own student assessment policies. As proper assessment of students requires the knowledge and skills of student assessment, universities must provide training for instructors on the assessment of students learning outcomes.

References

Abera, A. G. (2012). Continuous assessment (CA) vis à vis the attainment of major educational domains of physical education in Ethiopia. *International Journal of social science & Interdisciplinary Research*, 1(11), 17-27.

Biggs, J. B., & Tang, C. (2007). *Teaching for quality learning at university*. (3rd ed.). Buckingham: Society for Research into Higher Education and Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.

Cartwright, R., Weiner, K., & Veneruso, S. (2009). *Student Learning Outcomes Assessment Handbook*. Maryland: Montgomery College.

Clarke, M. (2012). *Measuring learning: How Effective Student Assessment Systems Can Help Achieve Learning for ALL*. Washington, D.C.: The World Bank.

Clarke, M. (2012). *What mater most for student assessment systems: A framework paper*. Washington, D.C. The World Bank.

Colbert, P., Wyatt-Smith, C., & Klenowski, V. (2011 forthcoming). A systems level approach to building sustainable assessment cultures: Moderation, quality task design and dependability of judgment *Policy Future. Retrieved from http://oucea.education.ox.ac.uk/wordpress/wp-content/uploads/2011/07/Colbert_Wyatt-Smith_Klenowski-FINAL.pdf on May 15, 2013*.

Connoley, R. (2004). *Criterion Referenced Assessment*. Deakin University.

Craddock, D., & Mathias, H. (2009). Assessment options in higher education. *Assessment & Evaluation in Higher Education*, 34(2), 127-140.

Curriculum Design and Delivery BPR-Implementation Working Group. (2009). *Procedures of modularization and block teaching: Master's Program of the Addis Ababa University*. Addis Ababa University.

Daniel, D. (2004). Observations and Reflections of the Higher Education Teachers on the Quality of Teaching and Learning in Higher Education in Ethiopia. *The Ethiopian Journal of Higher Education*, 1(1), 63-81.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of Educational Measurement* (5th ed.). Englewood Cliffs: Prentice-Hall, Inc.

ETS. (2003). *Linking Classroom Assessment with Student Learning*. New Jersey: Educational Testing Service.

FDRE. (2009). Higher Education Proclamation: Proclamation No. 650/2009. *Federal Negarit Gazeta*, pp. 4976-5044.

Ferguson, G., & Takane, Y. (1989). *Statistical Analysis in Psychology and Education* (6th ed.). New York: McGraw-Hill.

Fisseha, M. (2010). Review article: The roles of assessment in curriculum practice and enhancement of learning. *Ethiopian Journal of Education and Sciences*, 5(2), 102-114.

Furniss, E. (2003). *Assessing Learning Achievement*. New York: UNICEF.

Greaney, V., & Kellaghan, T. (2008). *Assessing National Achievement Levels in Education*. Washington D. C. The World Bank.

Hale, C. D., & Astolfi, D. (2011). *Measuring Learning & Performance: A Primer* (2nd ed.). Florida: Saint Leo University.

Hernandez, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education*, 64, 489-502.

Jackson, N., Wisdom, J., & Shaw, M. (2003). Using learning outcomes to design a course and assess learning. *The Generic Centre: Guide for Busy Academics*. York: Higher Education Academy. Retrieved on June 5, 2013 from

http://www.heacademy.ac.uk/assets/hlst/documents/guides_to_current_practice/learning_outcomes.pdf

James, R., Mcinnis, C., & Devlin, M. (2002). *Assessing Learning in Australian Universities*. Melbourne: Centre for the Study of Higher Education.

Jimma University. (2011). *Academic quality assurance guiding principles and standards*. Jimma: Jimma University.

Kazin, C., & Payne, D. (2009). *Ensuring Educational Quality Means Assessing Learning*. Educational Testing Service.

Kellaghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education*. Paris: UNESCO.

Kennedy, D. (1999). *Linking Learning Outcomes, Teaching and Learning Activities and Assessment Activities*. Ireland: University College Cork.

Knight, P. T. (2002). Summative Assessment in Higher Education: practices in disarray. *Studies in Higher Education*, 27(3), 275-286.

Linn, R. L., & Grolund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Mehrens, W. A., & Lehman, I. J. (1991). *Measurement and Evaluation in Education and Psychology* (4th ed.). Orlando, FL.: Holt, Rinehart and Winston, Inc.

Mekelle University. (n.d.). Assessment policy of Mekelle University. Mekelle, Ethiopia.

Miller M. D., Linn, R. L., & Gronlund N. E. (2009). *Measurement and assessment in teaching* (10th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Ministry of Education. (2003). *Harmonized Academic Policy of Ethiopian Public Higher Education Institutions*. Addis Ababa.

Ministry of Education. (2012). *Education Statistics Annual Abstract 2004 E.C (2011/2012)*. Addis Ababa: Ministry of Education.

Nitko, A. J. (1996). *Educational assessment of students* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Norton, L. (2007). Using assessment to promote quality learning in higher education. In A. Campbell. & L. Norton, (Eds.) *Learning, teaching and assessing in higher education: developing reflective practice*. Exeter: Learning Matters Ltd.

Norton, L. (2009). Assessing student learning, in, H. Fry. In S. Ketteridge. & S. Marshall (Eds.). *A Handbook for Teaching and Learning in Higher Education. Enhancing Academic Practice (3rd ed.)*. Abingdon: Routledge.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what Students Know The Science and Design of Educational Assessment*. Washington D.C.: National Academy Press.

Race, P. (1998). *The Lecturer's Toolkit* (2nd ed.). London: Kogan Page Ltd

Sadler, D. R. (1989). Formative assessment and the design of instructional systems, *Instructional Science*, 18, 145-165.

Stayhorn, T. L. (2006). *Frameworks for Assessing Learning and Development Outcomes*. Council for the Advancement of Standards in Higher Education.

Stephens, M., & Moskowitz, J. (2004). *Measuring Learning Outcomes in Developing Countries: A Primer*. Washington D.C.: USAID.

Struyven, K., Dochy, F., & Janssens, S. (2002). *Students' perceptions about assessment in higher education: a review*. Paper presented at the Joint Northumbria / Earli SIG Assessment and Evaluation Conference: Learning communities and assessment cultures. Newcastle, UK: University of Northumbria.

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and

assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, 30(4), 331-347.

Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478.

TGE. (1994). *Education and Training Policy*. Addis Ababa: Berhanina Selam Printing Press.

Trotter, E. (2006). Student perceptions of continuous summative assessment. Assessment & *Evaluation in Higher Education*, 31: 505-21.

UNESCO. (2006). *Assessment of learning outcomes*. Geneva: International Bureau of Education UNESCO.

University of Wolverhampton. (2006). University *Postgraduate Assessment Handbook for Staff*. United Kingdom.

Yohannes, W. (2006). *Outcomes Approach to Learning: A Good Practice for Ethiopian Higher Education*. Paper presented on the proceedings of the Conference "Future Direction of Higher Education in Ethiopia" organized by the Institute of Educational Research, Addis Ababa University, Ethiopia.

York University. (2002). *The Teaching Assessment and Evaluation Guide*. Retrieved date and year from www.yorku.ca/secretariat/senate/committees/scotl/.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45: 477-501.

Young, S. F., & Ontario, K. (2005). Teaching learning and assessment in higher education: Using ICE to improve student learning. *Proceedings of the Improving Student Learning Symposium*. London: Imperial College.