International Journal of Science Academic Research
Vol. 05, Issue 10, pp.8401-8409, October, 2024
Available online at http://www.scienceijsar.com

ISSN: 2582-6425

Research Article

PREVALENCE OF ESTIMABLE LEARNING ASSESSMENT PRACTICES: QUESTS AND IMPLICATIONS

***Firdissa Jebessa Aga**

Institute of Educational Research, Addis Ababa University, P.O.Box: 1176, Addis Ababa, Ethiopia

Received 26th August 2024; **Accepted** 24th September 2024; **Published online** 22nd October 2024

Abstract

The study examined the prevalence level of estimable learning assessment practices at four Ethiopian universities. Concurrent nested mixed research design was employed in the course of the study. Data were collected through questionnaires from staff members, and students or PhD and MA in Education and Behavioral Studies (CEBS), and Teaching of English as a Foreign Language (TEFL) at four universities; and through interviews from two officials at each of the universities. The results have shown that learning assessment quality enhancement practices and mechanisms were not as expected at the universities. Specifically, creating a detailed blueprint for the whole assessment tasks, using assessment quality circle, using assessment procedures that generate learners' interest, matching the developmental (formative) and judgmental (summative) roles of assessment, designing tasks that assess relevant generic skills along with subject-specific competencies, and placing assessment items in increasing order of difficulty were at a middling level. Furthermore, putting in place a robust internal quality assurance system for learning assessment was negligible at the universities. The results, therefore, imply that the universities along with the Ministry of Education (MOE) should make utmost concerted efforts to put in place a robust system of learning assessment practices that result in effective student learning.

Keywords: Prevalence, Estimable Practices, Ethiopian Universities, Learning Assessment, Quality

INTRODUCTION

This article presents the results of a study conducted on the extent to which estimable learning assessment practices prevailed at four Ethiopian universities, and to drive implications for improvement. The extent to which staff members, and students of MA and PhD at the universities witness some estimable learning assessment practices, and the presence of robust internal quality assurance system for learning assessment have been explored. The study has been founded on the view that estimable learning assessment practices inform proper teaching and enhance effective student learning. Assessment practices can be considered effective when they result in quality learning with actual and potential impact on the overall quality of a given educational program. An effective assessment for Sanga (2016:1) "helps to improve student learning and informs the staff members of their teaching process". Proper understanding of estimable practices of learning assessment by staff members and students materializes the constructive alignment among teaching, learning and assessment in such a way that change in one compels a sympathetic adjustment of the rest (Stiggins, 2007; Sanga, 2016; Firdissa, 2023). That is why different stakeholders who have different purposes and/or are affected by the education system have developed vested interests in the quality learning assessment matters at universities. Particularly, quality learning assessment gives insights and data for government leaders at different levels to direct and/or support assessment practices in accordance with the political and accountability requirements. To be effective in their instructional leadership, leaders' knowledge, skills and attitudes should progress along with assessment requirements and innovations. Cognizant of this fact and along with the fast changing assessment practices and contexts, the Ethiopian Government has put in place different policies and guidelines for employing appropriate learning assessment methods with the hope to ensure effective implementation of its programs. Consequently, Higher education Institutions (HEIs) in Ethiopia have given serious consideration to the quality of learning assessment in line with the desired Learning outcomes (LOs). Staff members in Ethiopian HEIs are, therefore, conversant with learning assessment through Higher Diploma Program (HDP), which is a practice-based training program for teacher educators at HEIs in Ethiopia. Basically, HDP has one year duration whereby staff members attend 2 hours discussion classes for 2 days per week, supplemented with additional classroom observations and secondary school visits for a week or two. Addis Ababa University has tailored the Program to its context by dipping the duration to a maximum of intensive four months by integrating different competencies, removing repetitive issues, and arranging intensive schedules (Firdissa, 2021). Assessment is an integral part in both cases (the National framework and that of the University). Whereas the National framework on HDP has four modules dealing with 'Reflective Teacher Educator,' 'Developing Active Learning,' 'Improving Assessment,' and 'Action Research, Making a Difference' (MoE, 2006); that of the AAU has five modules dealing with Understanding Higher Education, Modularization and Modular Curriculum, Managing Learning and Assessment, Subject Area Teaching, and Action Research and Field-based Learning (AAU, 2014a). Particularly, Module three on *Managing Learning and Assessment* presents the basics of assessment with rationales, principles, methods, importance and grading procedures. Staff members at Ethiopian HEIs are, therefore, mindful of assessment practices and tenets. AAU (2014a) suggests that a variety of assessment methods should be designed to satisfy all LOs. In designing or redesigning modules, it is, therefore, vital to identify and reach a consensus by staff members and academic leadership on appropriate parameters of assessment; and to decide which can be left to individual staff members or subject coordinators. Concerning the general provisions on examinations, AAU (2019: 78, Article 82, No. 821) indicates that:

Corresponding Author:** ***Firdissa Jebessa Aga
Institute of Educational Research, Addis Ababa University, P.O.Box: 1176, Addis Ababa, Ethiopia.

Student learning shall be assessed in a variety of ways/continuous assessment in the form of tests, assignments, presentations, etc. to determine the final grade earned. This shall account for 50% of the total module/course grade. The remaining 50% shall be allotted for a final exam conducted at the end of module/course delivery. Staff members shall monitor the students' academic performance by keeping track of records.

In the same vein, AAU (2014 a: 35-36) presents the following points concerning assessment of modular curriculum:

1) Performance of learners in a module should be evaluated in relation to the achievement of the modular-objectives (criterion-referenced) rather than on competitive basis (norm-referenced) and normal distributions;
2) the old system of using the normal curve for determining grades should be replaced by initial planning of correspondence between number-grades and letter- grades while determining the latter;
3) Failing grades for a module can be determined by learner performance below 60 percent of the total. it is suggested, however, that each instructor with the consultation of his/her department can modify the suggested grading scale;
4) Assessment of student work should be continuous, valid, and reliable; and
5) There should be a meaningful and effective system of evaluating, revising, up- grading or phasing out academic programs (AAU, 2014 a: 35-36; Firdissa, 2022, 2021, 2023).

Inherent within the aforementioned provisions are the need to ascertain student achievement of the set assessment modalities, master the LOs which are inherent within the modalities, and achieve learning with understanding. As a result of achieving learning with understanding, the learners demonstrate desirable changes in knowledge, skills, and attitude and/or behavior; and engage in high level cognitive thinking and actions.

LITERATURE REVIEW

Conceptions of estimable learning assessment practices and vested interests

Estimable learning assessment practices are all about good practices that result in proper teaching, and effective and quality student learning. For Scott and Webber (2016:6), good assessment practices support "quality teaching and learning" … then it is essential the assessment be planned and structured deliberately" in a way it meets the demands of the multiple stakeholders who have stakes in the overall education system in general and the graduates in particular with acceptable quality and utility. For Sanga (2016), "[t]he quality and utility of assessment rests upon the extent to which the performance measured represents appropriate and meaningful forms of human achievements that are relevant in real-life situations" (p.4). This shows that learning assessment services different purposes, which are, nonetheless, "neither separate nor entirely compatible" (Brown & Knight, 1994:13) with acceptable quality. Quality assessment for Ainslee (2018: cited in Firdissa, 2022) "basically focuses on the targeted areas with complete precision". The same source went on describing that assessment in the education industry should have content validity, reliability, generating interest by the student, and consequential relevance. Redecker and Johannessen (2013) have the view that "[t]he quality of teacher-made tests would improve greatly if they were not administered immediately but given to a few colleagues for review first". This, then increases validity, reliability, interest, and relevance aspect of the test. The validity of the test content for Ainslee (2018:1) deals with the test content to "be highly organized and should come across as clear and simple to the candidates attempting the test. It should not consist of faulty language or spelling defaults. The content should be in accordance with the subject that is being assessed and should not be out of the syllabus or topic". Reliability with reference to assessment signifies that each and every aspect of the assessment has a measurable outcome, and the quality of being accurately measured without the buildup of any flaw to fit for the purposes of multiple groups. By implication, assessment should reflect the simultaneous demands of multiple audiences and/or actor groups for multiple purposes, among others: test takers, students, score users, staff members, the governments, university management, employers, financing bodies, funding stakeholders, and the society at large (Brown & Knight, 1994; Luoma, 2001; OET, 2017; Firdissa, 2022). Individual students, therefore, make use of feedbacks to enable him/her to work on points in need of attention rather than to keep practicing points of strength, as a result of which assessment of effective learning is considered flexible as it considers individual learner needs to make sense of the feedbacks in the context of his/her own experiences (Brown & Knight, 1994; Bryan & Clegg (2006).

The different parties look for different tenets of the students' achievements. Employers look for pragmatic saleable knowledge, skills and behaviors and what they be able to do in accordance with their mission. University leadership wants to ascertain that students achieved the set LOs, and accordingly want to get adequate information about how much students have gained from their studies. Equally, teachers want to ensure that students have achieved learning with understanding. They can see evidence of students' progress during the learning process and provide timely feedback so that the learners take proper action in the moment. Parents, and students themselves need to know the level of their confidences to apply their competencies to out of school lives. Parents are, therefore, informed about what and how their children learned during the school days. In the long term, all the stakeholders make use of the information obtained from through assessment to support continuous improvement and innovations in learning (Brown & Knight, 1994; OET, 2017; Firdissa, 2022). Ainslee (2018) also explains that *generating interest by the student* deals with "the reason why tests should be objective in nature. Subjective tests are lengthy in nature not even generating interest of the staff members, leave alone the students. So, assessments should be explicit and creative which does not give a sense of boredom to the candidates".

Consequential relevance also deals with the reason for conducting an assessment, which requires a lot of time, dedication, and resources. This is because nobody would want so much hard work to go in vain. By implication, the assessment result should be so exact that it can be used as a tool to compare and analyze the data for future reference of the candidate's performance (Ainslee, 2018).

Along with the vested interests in learning assessment quality, validity, reliability, interest, and Consequential relevance, the American Association for Higher Education have devised nine principles of good practice for assessing student learning:

1. The assessment of student learning begins with educational values,
2. Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time,
3. Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes,
4. Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes,
5. Assessment works best when it is ongoing, not episodic,
6. Assessment fosters wider improvement when representatives from across the educational community are involved,
7. Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about,
8. Assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change,
9. Through assessment, educators meet responsibilities to students and to the public (Astin, Banta, Cross *et al.* nd).

The pragmatic conceptions of estimable learning assessment practices and the vested interests of the different parties have come with the increasing demand for accountability that can be achieved by ascertaining the level of student learning vis-à-vis the setout intended LOs, and the changing landscape of work and life demanding adaptable competencies.

Constraints in learning Assessment

The constraints in learning assessment have come along with the changes in learning assessment and consequent happenstances. Equally, assessment is "one of the most contentious and politicized dimensions within societies" (Scott & Webber, 2016). Our world is witnessing rapidly changing practices in assessment and "in the context within which assessment operates" (Bryan, & Clegg, 2006). The changes are partly internal to HEIs as a response to issues confronting them, and partly due to external pressures. Among others, increased class size, changing curricula, the need to support students better, declining resources, assignments and study time, modularization and assessment, plagiarism, computer-aided assessment, declining student retention, specifications and assessment of new kinds of learning outcomes, and problems of innovation in one way or another affect assessment of learning at universities (Bryan, & Clegg, 2006; Özturgut, 2011). Equally, substantive paradigm shifts from descriptions of programs in terms of years of study to learning outcomes; increasing cohort size and the shrinking unit of resource, the changing of student body, and policy climate have driven the changes in assessment at HEIs. Policy climate has noticeable influence on assessment in relation to quality assurance, enhancement, and accountability climate, which differs from nation to nation.

Crook *et al.* (2006, in Bloxham and Boyd, 2007), nonetheless, have the view that "equitable and consistent procedures are not sufficient to deliver good-quality assessment practice". The authors have the view that procedural changes and efficiencies determine individual practices. Bryan and Clegg (2006) also have the view that:

Worries about declining standards have resulted in institutions being cautious about approving changes to assessment, and extremely cautious about innovating in assessment in ways with which external examiners might be unfamiliar, of which they might not approve or to which students might object. The dominant culture is conservative and defensive rather than bold. It is often more difficult and more time consuming to gain approval for changes in assessment than for changes to any other aspect of courses (P.20). As a result, learning assessment practice lags well behind its equivalent in the school sector, relying largely on a limited range of tried (but not always tested) methods. It is dealt with in an ad hoc way and the situation is not mitigated by the 'amateur' status of many academics regarding assessment (Bloxham & Boyd, 2007, citing Murphy, 2006, Swann & Ecclestone, 1999, & Ramsden 2003). Acknowledging Price (2005), Bloxham and Boyd (2007) also assert that staff members "learn the craft of assessment informally through being assessed [themselves] and through being part of a community of practice, but lack scholarship regarding assessment". As the same source indicates, most staff members "have survived this approach to professional learning reasonably unscathed but it is not a recipe for enhancement; it provides no reliable route for ensuring that research on assessment reaches those doing the assessing".

In principle, occurrence of estimable practices of learning assessment at universities can serve as a catalyst to maximizing the potential benefits of assessment to inform teaching and improve learning. There, however, is scarecity of focus on critical matters for the quality of student learning, mainly in Ethiopia. As to my knowledge, estimable practices of learning assessment have never been topics of research in Ethiopian HEIs. Equally, "…there has been little investigation into the effect of classroom-based assessment on instructional and learning practices" (Muñoz & Álvarez, 2009). Assessment in HEIs has, therefore, remained under-conceptualized, in spite of the recent growth in interest towards enhancing the quality of graduates and being sensitive to answerability (accountability) (Bryan, & Clegg, 2006; Gerritsen-van Leeuwenkamp, Brinke, & Kester, 2017). Moreover, there is also a lack of established practices to regularly review newly prepared teacher made-test as assessment tools. The case has been exacerbated with the fluidity, contestability and complexity of quality conceptions in general and learning assessment quality in particular owing to the fact that quality with its indicators is determined by a wider set of criteria (Firdissa, 2009).

RESEARCH METHODOLOGY

The Research Design and participant selection procedures

Concurrent nested mixed research design was used in the course of the study. Both quantitative and qualitative data were collected at the same time, although the former was given more weight over the latter. Both sorts of the daata were collected from four Ethiopian public universities. The data of this paper is part of a larger study from which two papers have been published one from a pilot result and the other from the main result on different themes. For the sake of anonymity, the universities have been labelled as Unv4, Unv2, Unv3, and Unv4 standing respectively for university 1, university 2, university 3, and university 4. The selection was made using a lottery method from the universities functioning prior to 2015 in Ethiopia. That is, the names of each university were written and put in a container. Randomly picking the numbers four times, the four universities have been selected for the study.

From the selected universities, staff members (teaching and research), and students of PhD and MA at the CEBS and TEFL were selected to participate in filling questionnaires; and AVPs, and (associate)deans were selected to participate in interviews. All the participants at the universities were purposely selected on the basis of seniority, age, and availability of the required cohorts of students in Masters, and PhD programs.

Data Gathering Tools and Collection Procedures

Two data gathering tools were utilized in the course of the research work: two types- closed and open-ended questionnaires (one for the staff members, one for the students), and a semi structured interview schedule that was presented to eight subjects (two each at the four universities). Both of the questionnaires were dispatched to the subjects by hand delivery. The interviews were administered through face-to-face deliberations using a pre-prepared questions and the answers were tape recorded and transcribed later. Coding of the respondents and interviewees was made after the event as soon as the data were collected. In this case, coding was made in relation to the following.

a) The identities of the respondents of the open-ended questions of the questionnaires were coded as TR (TR1, TR2, TR3…TR109) for staff members, and SR (SR1, SR2, SR3…SR267) for students.
b) 1IUnv4, 2IUnv4; 1IUnv2, 2IUnv2; 1IUnv3, 2IUnv3; 1IUnv4, and 2IUnv4 signifying that there were 2 interviewees at each universities.

Methods of Data Analyses

The quantitative data were generated from the questionnaires, and the qualitative ones were generated from interviews, and open questions of the questionnaires. Though both the quantitative and qualitative data were collected at the same time, the quantitative data were analyzed first and the qualitative ones were used to supplement that of the quantitative ones. By way of discussions, the qualitative and quantitative results have been mixed leading to conclusions and implications.

RESULTS AND DISCUSSION

Whereas 400 copies of the questionnaires were dispatched at the four universities, 376 copies (94, 117, 112, and 53 respectively from Unv1, Unv2, Unv3, and Unv4) were properly filled and returned. The return rate was 94%. This section, therefore, presents the respondents' biodata on sex, respondent groups within the universities, respondent group per colleges, qualifications, ranks, students' program level and years of study, and years of experiences at their respective universities. This has been followed by presentation of the results on the prevalence of estimable learning assessment practices at the universities.

Biodata of the Respondents

Whereas 36 (9.6%) is a missing system, 299 (88%) of the respondents were males and just 41 (123%) were females showing male dominance. For the fact that the data sources were selected using purposive and availability sampling, no conscious efforts were made to get representative female subjects. The case, nonetheless, could signal the prevailing females' underrepresentation in teaching as well as in research posts at HEIs in Ethiopia. The result on the respondent groups within the universities has been presented in Table 1.

Table 1. Overall Respondent groups within the Universities

University	Students		Instructors		Total	
	Count	%	Count	%	Count	%age
U1	53	56	41	44	94	25
U2	98	84	19	16	117	31
U3	91	81	21	19	112	30
U4	25	47	28	53	53	14
Total	267	71	109	29	376	100

Table 1 shows that 267 (71%) and 109 (29%) of the respondents were students and staff members respectively. When it comes to disciplines, whereas 187 were from the CEBS, the remaining (i.e. 189) were from TEFL. The details can be seen from Table 2.

Table 2. Respondent Groups per colleges within the universities

College	University	Students		Staff members		Total	
		Count	%	Count	%	Count	%
CEBS	U1	21	60.0	14	40.0	35	18.7
	U2	64	94.1	4	5.9	68	36.4
	U3	47	73.4	17	26.6	64	34.2
	U4	10	50.0	10	50.0	20	10.7
	Total	142	75.9	45	24.1	187	49.7
TEFL	U1	32	54.2	27	45.8	59	31.2
	U2	34	69.4	15	30.6	49	25.9
	U3	44	91.7	4	8.3	48	25.4
	U4	15	45.5	18	54.5	33	17.5
	Total	125	66.1	64	33.9	189	50.3
Overall sum		267		109	29	376	100

It can be depicted from Table 2 that almost equal respondents participated from CEBS (187), and from TEFL (189). Also, the result on the educational qualification of the staff respondents has shown that the majority (63%) of them had doctorate degrees, followed by master's degree holders (37%). The staff respondents were also requested to indicate their respective ranks. The results have been presented in Table 3.

Table 3. The respondents' academic rank

		Frequency	Valid Percent
Valid	Associate Professor	13	12
	Assistant Professor	51	46
	Senior Lecturer	8	7
	Lecturer	36	35
	Total	109	100.0

Table 3 shows that the majority (46%) of the respondents had the rank of assistant professorship, followed by 35%, 12%, and 7% lecturer- ship, associate professorship, and senior lecturer respectively. When it comes to student respondents, just 241 indicated their program level and years of study as can be seen from Table 4.

Table 4. Student Respondents' Program Level and Years of Study Cross Tabulation

		Program Level of Study		Total
		MA	PhD	
Years of Study	1st	121	22	143
	2nd	12	28	40
	3rd	9	14	23
	4th	12	15	27
	others	3	5	8
Total		157	84	241

Table 4 shows that the majority (157) of the student respondents were MA and just 84 were PhD students. When it comes to their years of study, the majority (143) were 1st year, whereas 40, 27, and 23 were respectively at their 2nd, 4th, and 3rd years of study. Requested to indicate their years of experiences at their respective universities, 327 reacted whereas 49 was a missing system, as can be seen from Table 5.

Table 5. Teaching/research experience in years

Years of Experience	Students	Staff Members	Total	%	count	%
Under 3	74	31	7	8	81	25
3-6	36	15	15	17	51	16
7-10	27	11	14	16	41	13
above 10	100	42	54	60	154	47
Total	237	100	90	100	327	100

Table 5 shows that the majority (154) of the respondents had above 10 years of teaching and research experiences. The Table also shows that 81 had under 3 years of teaching and research experience. A further look at the data shows that from those who had under 3 years teaching and research experiences, 74 were students and only 7 were staff members. Furthermore, of those who had above 10 years teaching and research experiences, 100 were students and just 54 were staff members.

Presentation of data on the prevalence of estimable learning assessment practices

Quantitative and Qualitative data have been presented and analyzed in this subsection. Quantitatively, eight close-ended questions were presented to the respondents to gauge the prevalence of some estimable learning assessment practices at universities in Ethiopia in general and at their respective universities in particular. The Cronbach's Alpha reliability of the closed items is .90. Cronbach's Alpha if item deleted for all range from .877 to .908 (see Appendix 1). The respondents were instructed to rate their responses regarding the extent to which some listed assessment practices were prevalent at their universities by circling "1" for "very little", "2" for "a little", "3" for "medium", "4" for "greatly", and "5" for "very greatly".

The results have been presented in Tables 6. It can be depicted from Table 6 that the overall average means for the extent of using explicit/clear assessment tools, designing tasks that assess relevant generic skills along with subject-specific competencies, employing assessment procedures that generate learners' interest, placing assessment items in increasing order of difficulty, and balancing the developmental and judgmental roles of assessment were respectively 3.55, 3.49, 3.48, 3.7, and 3.46. This shows that the extent to which staff members at the universities were practicing estimable learning assessment remained in between medium and greatly. On the other hand, creating a detailed blueprint for the whole assessment tasks, and using assessment quality circle respectively had average means 3.23, and 3.30 denoting that staff members at the universities practiced estimable learning assessment to a medium extent. When seen per se, the extent of employing assessment procedures that generate learners' interest, using assessment quality circle, balancing the developmental (ongoing) and judgmental roles of assessment, designing tasks that assess relevant generic skills along with subject-specific competencies, and placing assessment items in increasing order of difficulty were respectively rated as medium by 151, 140, 135, 133, and 131 of the respondents. Furthermore, requested to indicate the level of their agreement on the extent to which their universities had robust internal quality assurance for learning assessment, 296 of the subjects properly responded, whereas 80 was a missing system. A frequency analysis of the data supplemented with a descriptive average mean result has been presented in Table 7.

It can be seen from Table 7 that the robustness of internal quality assurance for learning assessment had an overall average mean of 3.38, designating that the majority of respondents rated the level of the robustness of internal quality assurance for learning assessment at the universities as neither agree nor disagree. Seen per se from the Table, the majority (134) had shown their agreement, whereas 83 and 45 indicated their rating as neither agree nor disagree, and disagree. This shows that the robustness level of internal quality assurance for learning assessment was in between agree and neither agree nor disagree.

Table 6. Item-Total Statistics on the prevalence of estimable learning assessment practices

No	The extent of:	Count	Very Little	Little	Medium	Great	Very Great	Total	No Response	$\overline{X}$
1	Employing assessment procedures that generate learners' interest	No.	6	33	151	119	48	357	19	3.48
		%	1.6	8.8	40.2	31.6	12.8	94.9	5.1	
2	Using explicit/clear assessment tools	No.	8	40	112	137	58	355	21	3.55
		%	2.1	10.6	29.8	36.4	15.4	94.4	5.6	
3	Placing assessment items in increasing order of difficulty	No.	6	43	131	126	48	354	22	3.47
		%	1.6	11.4	34.8	33.5	12.8	94.1	5.9	
4	Designing tasks that assess relevant generic skills along with subject-specific competencies	No.	4	43	133	129	48	357	19	3.49
		%	1.1	11.4	35.4	34.3	12.8	94.9	5.1	
5	Using assessment quality circle	No.	10	61	140	100	44	355	21	3.30
		%	2.7	16.2	37.2	26.6	11.7	94.4	5.6	
6	Creating a detailed blueprint for the whole assessment tasks	No.	17	74	114	102	44	351	25	3.23
		%	4.5	19.7	30.3	27.1	11.7	93.4	6.6	
7	Balancing the developmental (ongoing formative)- and judgmental (periodic summative)- roles of assessment	No.	9	47	135	103	62	356	20	3.46
		%	2.4	12.5	35.9	27.4	16.5	94.7	5.3	

Table 7. Robustness of internal quality assurance for learning assessment

		Frequency	Percent	Valid Percent	$\overline{X}$
Valid	Strongly Disagree	11	3	4	
	Disagree	45	12	15	
	Neither nor	83	22	28	
	Agree	134	36	45	
	Strongly Agree	23	6	8	3.38
	Total	296	79	100	
Missing	System	80	21		
Total		376	100.0	100	

Qualitative data were also collected from the respondents and the interviewees on estimable learning assessment practices and learning assessment quality control/enhancement mechanisms used at the universities. Whereas just 173 (46%) respondents reacted properly to the open-ended questions on the issue, eight interviewees (2 from each of the universities) participated. By the open-ended questions, the respondents were directed to list down at least five estimable learning assessment practices or learning assessment quality control/enhancement mechanisms used at their respective universities/departments. Similarly, the interviewees were requested to tell estimable learning assessment practices or learning assessment quality control/enhancement mechanisms used at their respective universities/departments.

In both cases, the qualitative data have brought a number of estimable learning assessment practices. Commonly mentioned both by the respondents and the interviewees as the best/quality learning assessment practices or learning assessment quality control/enhancement mechanisms used at the universities have been listed hereunder.

1. Presence of assessment policy at all the universities;
2. Putting in place different assessment guidelines;
3. Establishing exam committee at all departments;
4. Giving timely quizzes/tests after lessons;
5. Giving mid exams before final exams;
6. Enforcing continuous assessment as motivating and assessment tool;
7. Setting and following up exam standards;
8. Orienting students at early times on the requirements of every courses;
9. Implementing criterion referenced grading system, and orienting students about it before they start courses;
10. Using quality assurance committee to follow up the performances of exam committees, and staff members' feedback giving mechanisms;
11. Preparing learning outcomes and specification of teaching;
12. Giving individual and group assignments/projects;
13. Using peer assessment especially at graduate level;
14. Using a variety of assessment types; and
15. Fixing and timing exam schedules for both mid and final examinations/semesters.

As TR27 indicates, some departments use good practices of assessment including scaffolding by providing learning chances "…for poorly performing students as a tutorial/make up to improve their skills and knowledge; and also supporting needy students economically to let them focus on their academic works". TR14 indicated that "while notifying exam results or showing their exam sheets, there was also a practice of giving chances and opportunities to help students relearn what they had missed before. Furthermore, TR52 indicated that "weighting 20% continuous assessment, 30% mid exam and 50% final" and orienting students before they begin class motivated students to work for achieving the requirements. An interviewee (1IUnv4) indicated that **"[**s]etting and following up exam standards, establishing exam committee, and employing criterion referenced grading system" were the best practices used at his university. For TR13 "[o]bjective orientation, empowering committee members to regulate common course assessments, establishing quality assurance committee, encouraging staff members' feedback, assessment transparency with students; frequent quizzes, assignments, and motivating students to achieve mastery" were among the commendable/estimable learning assessment practices at the university.

TR42 also had the view that "matching learning objectives with assessment items, assigning proper points/marks, and giving [timely] appropriate feedback/correction" were among the good learning assessment practices at his university. In the same vein, ST10 and TR3 had the view that implementing continuous assessment, and providing timely feedback for students, and giving different pedagogical trainings (such as Higher Diploma Program) to newly employed staff members, providing adequate classroom equipment (such as smart classrooms), and provision of digital e-resources library were implemented at their universities resulting in enhancing the quality of learning assessment. TR14 also indicated that his university implemented pre-assessment- by supervising whether classes/chapters/contents were addressed within the assessment items; on assessment- using the course team to contribute questions and/or to test the proposed test/exam itself before administering it, mainly on the common course; and post-assessment-receiving feedback from able students and relevant invigilators (on language, format, and even content) with the purpose to enhance good practice of learning assessment.

Furthermore, a list synthesized from the responses of SR1, SR2, TR43, TR17, TR87, and 2IUnv3, has given the following desirable practices at the universities.

1. Creating a detailed blueprint for the whole assessment tasks,
2. Using explicit/clear assessment tools,
3. Clearly measuring the required learning domains,
4. Embedding assessment tasks in the teaching-learning process,
5. Aligning assessment tasks with learning outcomes,
6. Clearly measuring the required learning domains,
7. Using explicit/clear assessment tools,
8. Employing assessment procedures that generate learners' interest,
9. Designing tasks that assess relevant generic skills along with subject-specific competencies, and
10. Making learning and assessment practices student centered and practical.

For TR3, TR5, TR8, and TR103, there were practices at their respective universities for planning assessment, monitoring content coverage and a variety of assessment tools, maintaining standard mode of delivery of assessment for freshman students, consistently employing fixed grading scale, using feedback for the next assessment tasks to enhance the quality of learning assessment. Some of the subjects further reiterated the importance of implementing desirable learning assessment practices rather than referring to the specific Commendable/estimable learning assessment practices at their universities. For instance, 1IUnv1 had the view that "[i]f we are truly concerned about and strive for education quality, learning assessment should be given due attention". Another respondent indicated that, [a]ssessment is more effective when it reflects an understanding of learning as multidimensional, integrative and revealed in performance over time; works best when the programs it seeks to improve have clear, explicitly stated purposes; and requires attention to outcomes and equally to the experiences that lead to those outcomes (SR2).

SR4 also had the view that "preparing program goal and specification of learning tasks" enhances the quality of learning assessment. For SR7, giving a variety of assignments, tests, questioning and answering tasks, quizzes, and exams would have values for enhancing the quality of learning assessment. SR8 also indicated that "diagnostic assessment (pre-assessment), formative assessment, summative assessment, and criterion-referenced assessment" enhance the quality of learning assessment.

Five respondents (ST9, TR31, TR65, TR70, & TR73) had the view that using a variety of assessment types, giving timely feedback, attaching assessment with its level of performance, preparing assessment in line with LOs, and initiating discussions on learning assessment quality among students and staff members could serve as catalysts to enhance the good practice of learning assessment. Some of the subjects also indicated their ambivalent attitudes towards learning assessment practices at their universities. For instance, an interviewee (2IUnv4) indicated that his university had put in place different assessment-related guidelines and empowered staff members to own them "as they are the implementers". He, nonetheless, had reservation on the proper implementation of the available policies and guidelines, and he also noted that "there is a huge difference among departments, even I am doubtful about the existence of basic understanding [of the roles of assessment] in some of departments".

Similarly, 2IUnv1 indicated that his college had established an assessment committee. He, nonetheless, lamented that "staff members didn`t have interest, I guess everybody doesn't want to make one`s exams open to others" and consequently "the established committee remained nominal as the staff members refused to send their exams or seek any assistance". "Consequently", he went on mentioning, "…instead of committee we tried to establish a one-to-one peer arrangement in which individuals review one another's assessment, and also "when dealing with common courses, in which a course-team leader coordinates inputs from members and fixes assessment on consensus".

Analyses of both Quantitative and Qualitative data have brought a number of results on the prevalence level of estimable learning assessment practices at the four universities in particular and in Ethiopian universities at large. The quantitative results have shown that the extent of learning assessment quality enhancement mechanisms and/or practices at the universities was rated almost to a medium extent. On average, the extent of using explicit/clear assessment tools, designing tasks that assess relevant generic skills along with subject-specific competencies, employing assessment procedures that generate learners' interest, placing assessment items in increasing order of difficulty, and balancing the ongoing formative and summative- roles of assessment were rated in between medium and greatly, majorly closer to a medium extent. In the same vein, creating a detailed blueprint for the whole assessment tasks, and using assessment quality circles at the universities were rated to a medium extent. Also seen in isolation (other than mean), the extent of employing assessment procedures that generate learners' interest, using assessment quality circle, balancing the ongoing - and summative roles of assessment, designing tasks that assess relevant generic skills along with subject-specific competencies, and placing assessment items in increasing order of difficulty were rated predominantly as medium.

Furthermore, the extent to which the universities had robust internal quality assurance for learning assessment was rated as neither agree nor disagree, and just a few indicated their agreement. On the other hand, the qualitative data have brought a number of estimable learning assessment practices, including, timely and continuous quizzes, assignments/projects, tests, orientations to students; and presence of assessment policy, guidelines, exam committee; quality assurance committee, peer assessment practices, a variety of assessment types, scheduled exam times, scaffolding practices (supporting students), and explicit/clear assessment tools at the universities.

The importance of learning assessment to stand as a catalyst for enhancing education quality was also reiterated by the subjects in a way that assessment is more effective when: 1) it reflects an understanding of learning as multidimensional, integrative and revealed in performance over time; 2) putting in place clear, explicitly stated purposes; 3) attention is given to learning processes and outcomes; 4) preparing program goals and specifications of teaching; 5) using a variety of assessment types; 6) giving timely feedback; 7) attaching assessment with its level of performance; 8) preparing assessment in line with the LOs; 9) preparing task specifications and test blue print, and 10) initiating discussions on learning assessment quality among students and staff members. Different sources also concur with almost all the findings. For instance, Luoma (2001) has the view that writing of a detailed blueprint and defining constructs in it, and operationally defining the tasks to be included in the test enhance estimable practices of learning assessment.

James, Mcinnis, and Devlin (2002:17) have also indicated that "…. it is essential for universities to have robust internal quality assurance for assessment and grading". The same authors (p.9) have the view that weighing assessment tasks by balancing the developmental (formative) and judgmental (summative) roles of assessment, and providing students with feedback enhance the effectiveness of student learning.

In the same vein, MEE (2015:45) has indicated that "characteristics of assessment should include: a balance between formative and summative assessments within meaningful and authentic contexts". As the same source further indicates, [a] balance of ongoing and periodic assessment opportunities will require learners to demonstrate a body of learning built up over time and to apply their knowledge and skills in different contexts. Mixing a range of learner controlled formative assessment opportunities will allow the learners themselves gauge how they are progressing…. (p.47)

The same source further indicates that "identification of assessment needs before learning experiences are planned allows targeted goals and performance to guide the teacher in their classroom practices. Similarly, implementing an assessment quality circle enhances the quality learning assessment and enhances the effectiveness of student learning (Brown & Knight, 1994). There, however, were reservations to guarantee proper implementation of available policies and guidelines due to different reasons; there were huge differences among the departments even on understanding the role of assessment to enhance effective learning; and staff members' reluctance to be guided by assessment committees, due to fear of exposing their exams to others.

Conclusion

The fact that the majority of the respondents had doctorate degree, the rank of assistant professorship, and above 10 years of teaching and research experiences signifies reliable information for the issues have been secured from the horse's mouth(i.e. from trustworthy sources who had the authority). Equally, the Cronbach's Alpha reliability of the items is high signifies that the items of the questionnaires are correlated and are internally consistent for generating dependable evidence. Moreover, the dominant results on the extent of the prevalence of estimable learning assessment practices and assessment quality control/enhancement mechanisms used at the universities were greatly rated to a medium extent calls for putting in place a robust internal quality assurance for learning assessment by way of writing task specifications and writing blueprint, balancing assessment tasks between the formative and summative roles of assessment within meaningful and authentic contexts, and guaranteeing proper implementation of available assessment policies and guidelines. The findings imply the need to encourage the universities to uphold the prevailed good learning assessment practices as a quest, majorly, nonetheless, the results imply that the universities along with the MOE should make utmost concerted efforts to put in place quality learning assessment practices by way of institutionally putting in place a system of:

1. Robust internal quality assurance for learning assessment;
2. Writing task specifications and writing blueprint;
3. Using assessment quality circle at universities;
4. Balanced tasks between the continuous and final assessment practices; and
5. Proper implementation of available assessment policies and guidelines.

This study has both strengths and limitations. The strengths of the study have been achieved mainly in three ways. First, utmost efforts were made to maintain ethical practices in the journey of the study by: 1) securing consent and cooperation of the data sources by giving clear directions regarding data collection procedures, required time investment to answer the questions, and confidentiality of the responses; and 2) maintaining the anonymity of the research participants so that the information they provided by no means could reveal their identities. This was done by politely warning the respondents (both staff members and students) not to write their names on any page of the questionnaire, by succinctly codding the respondents and the interviewees. Second, efforts have been made to corroborate and triangulate the research results and interpretations by way of mixing the qualitative and quantitative results at the level of discussions and conclusions leading to derive implications. Third, utmost efforts have been made to adhere almost with all the processes of conducting scientific research right from title selection, problematizing it, formulating themes, designing tools, collecting and analyzing data, and deriving conclusions and implications on the issue. This study also has its limitations. In the first place, the study is a bit general in terms of addressing estimable learning assessment practices. Equally, it might be slippery to exclusively treat estimable learning assessment practices without dealing with quality education, which itself is difficult to precisely demarcate. Second, purposely selecting the data sources on the basis of availability of the required cohorts of students in Masters, and PhD programs could have its own limitations. Equally, taking the research participants just from CEBS, and TEFL at the four universities could have effects on the results. Finally, notwithstanding the limitations, utmost efforts were made to use different strategies.

Funding Statements and Declarations

I hereby declare that there are No any financial or non-financial Competing interests that are directly or indirectly related to the work submitted for publication. I also declare that this article is part of a bigger research from which another article has already been published on a different theme. There, however, are inevitable overlaps of citations, and methodology.

REFERENCES

Astin, A.W.; Banta, TW., Cross, K.P; *et al.* (nd). Assessment Toolbox: Nine Principles for Assessing Student Learning, taking from American Association for Higher Education (available at https://teaching.berkeley.edu/sites/default/files/general/9_principles_of_good_practice_for_assessing_student_learning.pdf)

AAU [Addis Ababa University] (2019). *Senate Legislation* (available at file:///C:/Users/user/AppData/Local/Temp/AAU%20Senate%20Legislation%20of%202019.pdf)

AAU (2014 a). Higher Diploma Program: Module two: Modularization and Modular Curriculum (Up)

AAU (2014 b). Higher diploma Program: Module Three: Managing Learning and Assessment (Up)

Ainslee, J. (2018). The Significance of Quality Assessment in the Education System (available at https://elearningindustry.com/quality-assessment-in-the-education-system-significance)

Bloxham, Sue; and Boyd, Pete (2007). Developing Effective Assessment in Higher Education: a practical guide. England: Open University Press.

Brown, Sally; and Knight, Peter (1994). *Assessing Learners in Higher Education.* London & Newyork: RoutledgeFalmer.

Bryan, K. & Clegg, K. (eds.) (2006).*Assessment in Higher education*. New York: Rutledge

Firdissa J. (2023). Learning Outcomes, Instruction, and Assessment Alignment: The case at two Ethiopian Universities. *In International Journal of Research Publications*, 35(1), 206-217; doi:.10.47119/IJRP1001351 1020235573

Firdissa J. (2022). Assessment Quality at two Ethiopian Universities: Requirements versus Practices. *Research in Education Assessment and Learning*, 7(1), 1-14, DOI: https://doi.org/10.37906/real.2022.2

Firdissa J. (2021). Learning Assessment Practices in Ethiopian Universities: Quests versus Upheavals. *Research in Education Assessment and Learning*, 6(2), 1-14, ISSN: 1947-5497, doi.org/10.37906/real.2021.6 1

Firdissa J. (2009).Teachers' roles in quality management systems at universities. Dialogue, The Addis Ababa University Teachers' Association Journal. Addis Ababa: Alpha Printing PLC,1(1), 15-35.

Gerritsen-van Leeuwenkamp, K.J.; Brinke, D.J.; andKester, L. (2017).Assessmentqualityintertiaryeducation:Anintegrative literaturereview. *Studiesin Educational Evaluation*, Vol. 55, 94-116, http://dx.doi.org/10.1016/j.stueduc.2017.08.001

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

Luoma, S. (2001). *What does your test measure? Construct definition in language test development and validation*. Jyvaskyla, Finland: Centrefor Applied Language, ISBN 951-39-0897-6

MEE [Ministry for Education and Employment] (2015).*Educators' guide for pedagogy and assessment. Using a learning outcomes approach*. Floriana: Directorate for quality and standards in education

MoE [Ministry of Education] (2006).*Higher Diploma for Teacher Educators. A Training Portfolio* (unpublished).

Muñoz, A.P.and Álvarez, M.E. (2009). Wash back of an oral assessment system in the EFL classroom. *Language Testing*, 27(1)1–17, DOI:10.1177/0265532209347148

OET [Office of Educational technology] (2017). Reimagining the Role of Technology in Education: National Education Technology Plan Update (available at http://tech.ed.gov)

Özturgut, O. (2011). Standardized Testing In The Case Of China and the Lessons to Be Learned For the U.S. *Journal of International Education Research*, 7(2), 1-6

Redecker, Ch. & Johannessen, Ø. (2013).Changing Assessment-Towards a New Assessment Paradigm Using ICT. *European Journal of Education*, 48(1), 77-96

Sanga, Ph. L. (2016). Implications of Teacher Educators' Practices in Assessment for Student Learning in Tanzania. In *Makerere Journal of Higher Education*, 8 (1), 1–24; DOI: http://dx.doi.org/10.4314/majohe.v8i1.1

Scott, S. Scott, D., Webber, Ch. (eds., 2016).*Assessment in Education: Implications for Leadership*. New York: Springer ISSN2198-2651, DOI 10.1007/978-3-319-23398-7

Stiggins, R. (2007). Assessments through the student's eyes. *Educational Leadership, 64*(8), 22-26.

Appendix: Cronbach's Alpha overall Reliability, and Item-Total Statistics

Cronbach's Alpha		Cronbach's Alpha Based on Standardized Items		N of Items	
.90		.90		8	
Item-Total Statistics Issues	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
1. Employ assessment procedures that generate learners' interest	23.46	27.049	.718	.553	.884
2. Use explicit/clear assessment tools	23.38	26.745	.671	.521	.888
3. Place assessment items in increasing order of difficulty	23.48	27.298	.644	.455	.890
4. Design tasks that assess relevant generic skills along with subject-specific competencies	23.47	26.555	.734	.564	.882
5. Use assessment quality circle	23.66	25.731	.784	.646	.877
6. Create a detailed blueprint for the whole assessment tasks	23.72	25.669	.717	.569	.884
7. Balance the developmental (ongoing formative)- and judgmental (periodic summative)- roles of assessment	23.50	25.356	.784	.628	.877
8. Your university (where you study) has robust internal quality assurance for learning assessment. To what extent do you agree?	23.50	28.847	.444	.221	.908

****