Aga, F. J. (2024). Hurdles to learning assessment quality: Their detrimental effects on student learning. *Intersection: A journal at the intersection of assessment and learning, 5*(2), 48-68.

Hurdles to Learning Assessment Quality: Their Detrimental Effects on Student Learning

Firdissa J. Aga, PhD.

Author Note

Firdissa J. Aga, ORCID: 0000-0001-6602-9541

There are no conflicts of interest to disclose

Intersection: A Journal at the Intersection of Assessment and Learning

Vol 5, Issue 2

Abstract: The study investigated hurdles to the quality of student learning assessment by examining issues related to assessment procedures and practices, learners and learning, learning resources and test constructs, and test admin and feedback. Quantitative and qualitative data were collected from two Ethiopian universities using two types of questionnaires. The results have shown that learning assessment practices and procedures were inundated with a number of hurdles. Among them were academic dishonesty, a dearth of giving timely feedback, assessment procedures' failure to direct learners to higher-order learning outcomes, test construct underrepresentation and irrelevance, and learners' temptation to work for good grades rather than for acquiring the required competencies. These were the major hurdles that had detrimental effects on student learning at the universities. It has, therefore, been recommended that universities and their communities should ensure meaningful and useful assessment of student learning by putting in place effective and quality assessment that constitutes deep learning guided by effective feedback towards the achievement of learning outcomes (LOs).

Keywords: *Assessment Quality, Effects, Ethiopian Universities, Hurdles, Assessment, Learning Outcomes, Learning*

Introduction

This paper describes hurdles to learning assessment quality at two Ethiopian universities. It addresses learning assessment shortfalls related to assessment procedures and practices, learners and learning, learning resources and test constructs, and test admin and feedback. The study is important as it sheds light on the learning assessment quality shortfalls in terms of these constructs. The paper, therefore, outlines conceptual and theoretical background, problem statement, objectives and methodology of the study, results, discussions and conclusions, and recommendations.

Conceptual and Theoretical Background

Assessment has a powerful influence on effective student learning. It is considered effective when it is aligned with Learning Outcomes (LOs), focuses on provision of guidance and feedback towards the achievement of the LOs; results in deep learning and mastery of the required competencies in knowledge, skills, and dispositions; and demonstrates observable levels of cognitive ultimatum put on students (Masuku, Jili, & Sabela, 2021).

Effective assessment is explicit for measuring learning and using assessment data to improve learning as a necessary part of every teacher's work (USDE, 2017). In this sense, assessment is considered as a pedagogical instrument for measuring and promoting deep learning that requires an understanding of the assessment processes, criteria, and standards, and with the involvement of students in the construction of assessment tools (Masuku, Jili & Sabela, 2021). A synthesis from these researchers' claims has given that assessment: 1) should assist students to develop critical thinking skills and enable them to construct knowledge and identify weaknesses in their learning process; 2) has to be structured in such a way that it is possible to observe the levels of cognitive demand placed on students; 3) has to encourage and reinforce a deep approach to learning, increased objectivity in the grading of students, and the provision of timeous feedback; 4) should not only be about grading the performance of students but also should focus more on what students were asked to learn and reflect on the methods that can be used to improve their learning; and 5) has to be flexible, relevant and meaningful to ensure that it nurtures deep learning. Inasmuch as assessment influences deep learning, and at the same time measures and monitors it, its quality has attracted the interests of many individuals, groups, and countries. Consequently, the quality of assessment is considered as an integral part of good teaching practice (Edwards, 2013).

Quality assessment, for Leeuwenkamp, Brinke, and Kester (2017), includes the quality of all aspects of assessment practices, such as test items, tasks, the process of assessing, a program of assessments in a course or a curriculum and the procedures, policies, and administration of the assessment process. For Ainslee (2018, p.1) quality assessment "basically focuses on the targeted areas with complete precision". The author explained that assessment in the education industry should have content validity, reliability, generating interest by the student, and consequential relevance. Reliability with reference to assessment signifies that each and every aspect of the assessment has both a measurable outcome and the quality of being accurately measured without the build-up of any flaw. Ainslee (2018, p.2) further 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 teachers, leave alone [that of] the students. So, assessments should be explicit and creative which does not give a sense of boredom to the candidates".

The need for quality assessment also resonates with the role of higher education "to transform, liberate or free the individuals, giving them a chance to reason and develop reasoning skills and critical thinking skills. Transformation of students is informed by what is taught, how it is taught, how the work is assessed and how students process the information" (Masuku, Jili, & Sabela, 2021, p.277). Inherently, teachers at Higher Education Institutions (HEIs) are expected to know the designed LOs and on what to focus in order to ensure their assessment of student learning is meaningful and useful

for the students' on-going learning and development (Edwards, 2013).

Cognizant of the role of quality assessment for effective learning along with the role of higher education to stand answerable to the taxpayers on what is taught, how it is taught, how students process the information, and how the work is assessed, different countries including Ethiopia, have given due recognition to learning assessment. Particularly, mindful of the fast- changing assessment practices and contexts, Ethiopia has put in place curriculum standards and assessment modalities in which expected LOs are predefined and stated in national and institutional curricula. National and institutional policies and regulations also plead for employing appropriate learning assessment methods to ensure effective implementation of programs and to guarantee effective student learning. HEIs in Ethiopia have, therefore, given serious consideration to competency-based learning and assessment related to LOs (AAU, 2015). Instructors in Ethiopian HEIs are, therefore, acquainted with competency-based learning and assessment, and modularization through Higher Diploma Program (HDP).

Whereas modularization is viewed as a process of designing and implementing a specific and self-contained learning resource, HDP is a practice-based training program for teacher educators at HEIs in Ethiopia. Basically, HDP was designed for teacher educators at the College of Education and Behavioral Studies (CEBS) and had a one-year duration whereby instructors attend two hours of discussion classes for two days per week, supplemented with additional classroom observations and secondary school visits for a week or two. Recently, the Addis Ababa University (AAU) has customized the Program to its context by reducing the duration to a maximum of intensive four months by integrating different competencies, truncating redundant topics, and arranging intensive schedules (Firdissa, 2021).

In both cases (the National framework and that of the AAU), assessment has been considered as a fundamental part- and critical component- of the Ethiopian education system. 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, that of 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, 2015).

Particularly, Module three on Managing Learning and Assessment, presents the basic elements of assessment with rationales, principles, methods, importance, and grading procedures. AAU (2015) 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 instructors and academic leadership on appropriate parameters of assessment; and to decide which can be left to individual instructors or subject coordinators. In explaining the need for assessing students' learning in a variety of ways, the general provision of the AAU Senate Legislation on examinations indicates that:

> …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. Instructors shall monitor the student's academic performance by keeping track of records (AAU, 2019, p. 78).

In the same vein, AAU (2015) asserts that in the implementation of the modular curriculum, performance of learners should be evaluated in relation to the achievement of the modular-objectives (criterion-referenced) rather than on a competitive basis (norm-referenced) and normal distributions. It further espouses replacing the old system of using the normal curve for determining grades by criterion referenced grading to initially planning the correspondence between number-grades and letter- grades while determining the latter. The same source further asserts that: 1) failing grades for a module can be determined by learner performance below 60 percent of the total; 2) assessment of student work should be continuous, valid, and reliable; and 3) there should be a meaningful and effective system of evaluating, revising, up-grading or phasing out academic programs.

Overall, many of the available national and institutional curricula documents and/or guidelines advocate that students achieve the requirements set out in criterion-referenced assessment modalities and master the LOs which are inherent within the modalities and achieve learning with understanding. These in turn call for aligning competency-based outcomes and learning activities with assessment; intentionally designing curricula around competencies with explicit, measurable, transferable LOs; and integrating with instruction and assessment. With the need for the said alignment, designing and integration have come with due recognition of the role of HEIs to transform their students by transforming what they teach, how they teach, how their students learn, and how they assess the learning (Masuku, Jili, & Sabela, 2021). By implication, HEIs are becoming accountable for the quality of their assessed graduates. Though available policies and guidelines grant institutional autonomy and academic freedom in Ethiopia, inefficiencies including poor assessment practices cannot be tolerated in the current Ethiopian higher education landscape.

Inherent within the heightened interests, concerns, and contexts in assessment matters (globally, nationally, as well as institutionally) is a quest for assuring quality outcomes of HEIs, one of which can be achieved through learning assessment quality. There are, however, a number of obstructions to materialize learning assessment quality. Many of the obstructions are related to assessment procedures and practices, learners and learning, learning resources and test constructs, and test admin and feedback. The study, therefore, sheds light on these issues and ultimately sends messages to instructors and leaders to deliver high-quality assessment practices by planning, delivering, supporting, maintaining standard and consistent procedures, and guaranteeing efficiencies of individual practices.

Problem Statement

Learning assessment is entrusted with influencing effective and deep learning. There are, however, a number of assessment-related hurdles that impact the effectiveness of student learning. Interactive factors such as increasing cohort size/diversity of students joining HEIs, shrinking/declining learning resources, time pressure, unavailability/poorly designed curricula, academic dishonesty, and loose

alignment of competency-based outcomes and learning activities with assessment practices in one way or another affect the effectiveness of student learning.

Issues related to quality learning assessment and the hurdles affecting it 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, p.1). Whereas the need for enhancing the quality of graduates and being sensitive to answerability (accountability) have attracted the interests of many stakeholders, avoiding assessment hurdles and enhancing deep learning seem floundered due to different reasons.

The first is the lack of intentional and regular assessment training to people engaging in assessment work (Horst, & Prendergast, nd, p. 1, citing Hutchings, 2010, & Nicholas & Slotnick, 2018). This in turn has resulted in exerting very little conscious efforts to create proper micro and macro level views and knowledge on learning assessment quality in Ethiopian universities, which leads to abusing assessment practices. James, McInnis, and Devlin (2010) also have the view that misunderstanding and confusion about assessment requirements and the correct attribution of original ideas may result in inadvertent plagiarism.

Second, it seems that there have been mismatches between theories and practices of learning assessment quality in Ethiopian universities. Equally, the diversity of assessment practices across fields of study demand "specifics of change …to be implemented and managed at the level of faculty and department. Getting consensus is not easy and any significant rethinking and change can take a number of years to implement successfully" (James, McInnis, & Devlin, 2010, p. 11).

Third, the fact that there are multiple stakeholders in learning assessment quality has resulted in disagreement on how to establish good/quality learning assessment practices in Ethiopian universities. This in turn has threatened the validity of learning assessment as construct underrepresentation wherein the test is too narrow and the score does not reflect enough aspects of the construct of interest; and construct-irrelevant variance in which case other factors independent of the focal construct consistently affect the scores, while the scores are interpreted only in terms of the focal construct (Firdissa, 2021).

The accumulated effects of assessment shortfalls have remained hurdles to implementing quality assessment and to bring effective learning in Ethiopian universities. Formal and informal observations also have shown that the hurdles to learning assessment quality are related to the level of directing learners to higher-order objectives, balancing assessment practices, inspiring deep learning, learners' efforts to acquire the required competencies, module/learning resource preparation, provision of individual feedback, assessing a diverse mix of learners, avoiding academic dishonesty, orientation of the educational system, connecting testing with learning, test construct representation and relevance, sequencing items, considering affective and psychomotor domains, clarity in test directions and statements, and constructing test items. This study is, therefore, important to shed light on the hurdles related to these constructs so that HEIs' instructors and leaders would take care of these in their planning, implementing and working for continuous improvement of assessment for effective

student learning.

Objective and research question of the study

The study is intended to examine hurdles to the quality of student learning assessment and gauging their detrimental effects and feasible remedies at two Ethiopian universities. To achieve the objective, the study asked: What are the hurdles to the quality of student learning assessment, and to what extent they affect the effectiveness of student learning? In doing so, factors that jeopardize student learning assessment quality related to assessment procedures or practices, learners and learning, learning resources and test constructs, and test administration and feedback have been investigated quantitatively and qualitatively.

Research Methodology

Quantitative and qualitative data were collected using two types of questionnaires from two purposely selected Ethiopian universities. The data for this paper is part of an extended study, which is under full-scale analysis. For the sake of anonymity, the universities have been labeled as U1 and U2. The rationale for taking two different universities and including students and faculty was both for convenience and purpose. Whereas U1 was selected based on its age and productivity in offering postgraduate programs, U2 was selected for convenience purposes. That means, U1 is the oldest and the largest HEI in the country. U2, on the other hand, is a newly emerging and a younger university than U1. The selection of both universities was due to their convenience for data collection. Whereas the researcher was and is working at U1, he was also a guest lecturer at U2 for two weeks. He therefore managed to opportunely collect data both quantitative and qualitative at the two universities supported by two assistant data collectors. Consequently, there was no intent to compare the data of the two universities.

From both universities, two academic units, namely, College of Education and Behavioral Studies (CEBS) and Teaching of English as a Foreign Language (TEFL), were selected. The selection of CEBS and TEFL and the participants therein were purposeful. This is due to the fact that the researcher himself has a profound background in both areas; both as a researcher, as well as his field of studies, are from these disciplines (two master's in education and a PhD in TEFL). He therefore has a clear idea of the attributes related to learning assessment quality and the likely impediments that would affect the effectiveness of student learning.

Whereas the quantitative data were generated both from instructors and students using close- ended questions, the qualitative questions were generated through open-ended questions on the questionnaires. The data generated both from instructors and students through close-ended questions of the questionnaires were numbered, quantitatively analyzed, and interpreted. They were entered into the Statistical Package for the Social Sciences (SPSS) software version 23 and analyzed. Accordingly, analyses of simple frequency, descriptive statistics, Cronbach's Alpha reliability, comparison of means, and independent sample T-test were made. The data generated through the open-ended questions were qualitatively categorized and analyzed.

Though both the quantitative and qualitative data were collected at the same time, the qualitative data were embedded as secondary (or supportive) evidence within the quantitative ones (Creswell, 2018).

The qualitative data, therefore, served to give meaning to the observed state of the affairs regarding the impediments affecting the quality of student learning assessment at the two universities, and to add depth to the overall results and thereby "putting fleshes on the bones" (Dornyei, 2007, p. 39). From the two universities, 161 respondents (72 instructors and 94 students) properly filled in and returned the questionnaires. All the returned copies of the questionnaires were numbered as: TR1-72, and SR1-94, representing the codes of respondent instructors and students respectively.

Results

Whereas 190 copies of the questionnaires were dispatched at the two universities, just 166 subjects (72 instructors and 94 students) were returned. Forty-nine and twenty-three of the instructors were respectively from U1 and U2. Similarly, fifty-three and forty-one of the student respondents were respectively from U1 and U 2. Of the 53 student respondents from U1, five were PhD students studying at U1 coming from other universities; and two of them indicated that they were not employees of any university.

Biodata of the Respondents. Student-and teacher-respondents were requested respectively to indicate their study programs and qualifications. On this specific question, 85 of the students and 71 of the instructors properly filled and returned the questionnaires as can be seen from Table 1.

Table 1

The students' study programs or levels and teachers' qualifications

		Students' study programs		Teachers' Qualification	
		Frequency	%	Frequency	%
Valid	MSc/MA	52	61	19	27
	PhD	33	39	52	73
	Total	85	100	71	100
Missing	System	9		1	
Total		94	100	72	100

Table 1 shows that 52 (61%) and 33 (39%) of the students were respectively attending MSC/MA and PhD programs. It can further be seen from the Table that the majority (73%) of the teacher respondents had a PhD qualification followed by 27% master's holders. Five copies of the student questionnaires, nevertheless, were jettisoned as they were not properly filled. In analyzing the data, therefore, 161 (72 from instructors and 89 from students) of the properly filled copies of the questionnaires have been used. Requested to indicate their teaching/research experience in years, all the instructors and eighty-five of the students responded properly, as can be seen from Table 2.

Table 2

Respondents' teaching/research experience in years

Experiences		Frequency	Valid Percent
Valid	Under 3	14	9
	3-6 Years	21	13
	7-10 Years	26	17
	Above 10 Years	96	61
	Total	157	100.0
Missing	System	4	
Total		161	

Table 2 shows that the majority of the respondents had teaching/research experience of over 10 years. As can be seen from the Table, 96 (61%) of them had teaching/research experience of above 10 years. A further separate frequency analysis for the same has shown that all those who had under 3 years of teaching/research experience were student respondents. Of those who had over 10 years of teaching/research experience, 85% and 43% respectively were instructors and students.

Teacher respondents were also requested to indicate their respective ranks and the results have been shown in Table 3.

Table 3

Instructor respondents' Ranks

Rank		Frequency	Valid Percent
Valid	Associate Professor	13	18
	Assistant Professor	41	57
	Senior Lecturer	4	6
	Lecturer	13	18
	Assistant Lecturer	1	1

Table 3 shows that the majority (57%) of the respondents had the rank of assistant professorship, followed by 18% associate professorship and equally, lectureship.

Hurdles affecting student learning assessment quality. Seventeen closed questions and one open-ended with three sub-questions, were presented to the respondents to examine the hurdles affecting the quality of student learning assessment at the two universities in particular and at all HEIs in the country in general. The respondents were requested to rate the detrimental effects of some identified hurdles to maintain learning assessment quality on a scale from "1" (no effect) to "5" (major effect), related to assessment procedures/practices, learners and learning, learning resources and test constructs, and admin support and feedback.

Related problems under assessment procedures or practices were failure to direct learners to higher-order objectives, excessiveness of assessment practices, and encouraging shallow learning. The learner and learning related hurdles also include learners' work to earn good grades rather than to acquire the required competencies, deficiency in connecting testing with learning, and ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain). Moreover, included under the learning resources and test constructs related shortfalls, were the absence of well-prepared module/learning resources for the courses taught, test construct underrepresentation and irrelevance, deficiency in sequencing items, lack of clarity in test directions, ambiguous test statements, and poorly constructed test items.

Finally, admin support and feedback related shortfalls include inability to provide high quality individual feedback, failure to to assess large class size, and inability to avoid plagiarism/ academic dishonesty. In one open-ended question, the respondents were requested to list at least three critical hurdles that affected learning assessment quality at their respective universities and/or in Ethiopian universities at large. There was no problem with the participants' understanding of the questions. The questionnaires were prepared anew by the researcher himself and not used anywhere. The questions of the questionnaires were validated earlier as a result of refinements, amendments, additions and omissions of some items.

The Cronbach's Alpha reliability of the closed items is .909. Cronbach's Alpha, if item is deleted (i.e., if the particular item is removed from the scale) has also ranged from .898 to .911 (see Appendix 1).

Whereas the analyses of the frequency, independent sample T-test results for instructors and students, comparison of the means for the academic units (CEBS and TEFL), and independent sample T-test for the Sexes of the quantitative data can be seen respectively from Appendices 2, 3, 4, and 5, a descriptive analysis of the same has been presented in Table 4 differentiated under assessment procedures/practices, learners and learning, learning resources and test constructs, and test admin and feedback.

Table 4

Descriptive Statistics of the hurdles affecting learning assessment quality

Differentiated List of hurdles	N	Min	Max	$\bar{X}$	SD	Variance
A. Assessment procedures/practices Related hurdles						
1. Assessment procedures fail to direct learners to higher-order objectives	144	1	5	3.4	1.0	1.0
2. Excessive assessment practices	141	1	5	3.2	1.1	1.1
3. Assessment practices encourage shallow learning	161	1	5	3.4	1.2	1.4
Average	149	1	5	3.3	1.1	1.2
B. Learners and Learning Related Hurdles						
4. Learners work to earn good grades, rather than to acquire required knowledge	146	1	5	3.9	1.1	1.2
5. Exam-oriented educational system	146	1	5	3.6	1.2	1.4

6. Deficiency in connecting testing with learning	143	1	5	3.6	1.1	1.2
7. Ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain)	144	1	5	3.6	1.2	1.5
Average	145	1	5	3.7	1.1	1.3
C. Learning resources & Test constructs Related Hurdles						
8. Lack of well-prepared module/learning resources for the courses taught	161	1	5	3.8	1.1	1.2
9. Test construct underrepresentation	144	1	5	3.4	1.1	1.1
10. Test construct irrelevant characteristics	144	1	5	3.3	1.1	1.2
11. Deficiency in sequencing items	142	1	5	3.1	1.1	1.2
12. Lack of clarity in test directions	143	1	5	3.1	1.2	1.4
13. Ambiguous test statements	145	1	5	3.1	1.1	1.3
14. Poorly constructed test items	145	1	5	3.2	1.3	1.6
Average	146	1	5	3.3	1.1	1.3
D. Test administration and Feedback Related Hurdles						
15. Inability to provide high quality individual feedback	161	1	5	3.9	1.0	1.0
16. Challenges to assess a diverse mix of learners (large class)	145	1	5	3.5	1.1	1.1
17. Inability to avoid plagiarism and/or academic dishonesty	161	1	5	4.0	1.1	1.2
Average	**156**	**1**	**5**	**3.8**	**1.1**	**1.1**
Grand average	**149**	**1**	**5**	**3.5**	**1.1**	**1.2**

$\bar{X}$ - Average Mean, SD- Standard Deviation

Table 4 shows that the minimum response was 1 (no effect) and the maximum was 5 (major effect) for all the categories. Whereas the mean of the grand average result is 3.5 with SD of 1.1 and variance of 1.2, the overall results fall almost close to “moderate effect” and by far above “no effect” and “minor effect”. The mean average of the major hurdles are, therefore, closer to “moderate effect” than “no effect” and/or “minor effect”.

Of the categories, test administration and feedback related hurdles have the major effect with the average mean of 3.8, SD 1.1, and variance of 1.1. More specifically, inability to avoid plagiarism and/or academic dishonesty, inability to provide high quality individual feedback, and challenges to assess a diverse mix of learners (large class) respectively have average means of 4.0, 3.9, and 3.5 (see under D in Table 4).

Next, learners and learning related hurdles were the second in affecting learning assessment quality with 3.7, 1.1, and 1.3 average mean, SD, and variance respectively. A closer look at the Table shows that learners who work to earn good grades rather than to acquire the required knowledge have a mean average of 3.9 with 1.1, and 1.2 SD and variance respectively. On the other hand, the categories of exam-oriented educational systems, deficiency in connecting testing with learning and ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain) all have the average mean of 3.6 with varying SD and variance.

Table 4 further shows that the average means of assessment procedures/practices and learning resources and test constructs related hurdles are equally 3.3, with SD of 1.1(see under A & C in Table 4). When seen separately, assessment procedures/practices related hurdles, assessment procedures fail to direct learners to higher-order objectives, and assessment practices encourage shallow learning have equal average means of 3.4 with 1.0 and 1.2 SD respectively.

Regarding learning resources and test constructs related hurdles, lack of well-prepared module/learning resources for the courses taught, test construct underrepresentation, test construct irrelevant characteristics, and poorly constructed test items have average means of 3.8, 3.4, 3.3, and 3.2 respectively with similar SD except the last which has 1.3 SD. Equally, deficiencies in sequencing items, lack of clarity in test directions, and ambiguous test statements all have 3.1 average means.

Analysis of the comparison of means of the academic units (CEBS and TEFL) has shown that the ratings by CEBs has greater average mean than that of TEFL with 3.5 and 3.4 respectively (see Appendix 4).

Furthermore, a comparison of the means of the instructors and students, using analysis of independent sample T-test, has given close average means, SDs, and Std. Error Means (see Appendix 3). The average means and SDs for assessment procedures/practices related hurdles are 3.1 for the instructors and 3.4 for the students with 1.1 and 1.0 SD respectively (see under A in Appendix 3).

The Table further shows that the average means for learners and learning related hurdles are 3.8 for the instructors and 3.6 for the student respondents with a similar SD of 1.1 (see under B in Appendix 3). For learning resources and test constructs related hurdles, the average means are 3.2 and 3.4 with SDs of 1.1 and 1.2 respectively for the instructors and student respondents (see under C in Appendix 3). When it comes to test administration and feedback related hurdles, the average means are 3.9 and 3.7 respectively for the instructors and student respondents with equally 1.1 SD (see under D in Appendix 3).

A comparison of means of the sexes using analysis of independent sample T-test has shown that males highly rated the effects of the listed hurdles to maintain learning assessment quality. As can be seen from Appendix 5, grand average means for males and females are 3.6 and 3.5 respectively. When seen separately too, the average means of males are higher than that of the females except in one are, i.e. for assessment procedures/practices related hurdles. The average means for males and females are 3.3 and 3.6 respectively.

Furthermore, requested to list hurdles to maintain the quality of learning assessment at their respective universities in particular and/or in Ethiopian universities in general, 97 (60%) of the respondents listed salient hurdles that affected the quality of learning assessment. A synthesis from 26 (16%) respondents indicates that the major hurdles could be brought under: 1) content- related, 2) test administration–related, 3) student-related, 4) teacher-related, and 5) learning resources- and policy–related issues and/or factors (SR7, SR15, SR22, SR26, SR45, SR41, SR51, SR57, SR59, SR68, SR75, SR81, SR90, SR92; TR4, TR6, TR10, TR17, TR33, TR51, TR58, TR63, TR66, TR68, TR70, TR72).

Content-related issues that affected the quality of learning assessment include shortfalls in: properly identifying key competencies of the courses, employing assessment for learning and learning-oriented assessment, aligning course content with assessment tasks, covering lesson contents (not planning lesson time), test item construction, maintaining uniformity of different tests on the same course by different instructors, preparing items from contents covered, avoiding students' test wise-ness, and

asking high level cognitive questions beyond facts and figures.

In the same vein, test administration-related issues/factors that affected the quality of learning assessment include inability to: avoid plagiarism and/or academic dishonesty, properly manage large number of students in a class, balance high teacher workload, give due attention to continuous assessment, put in place proper record keeping system of test related information, halt mobile uses during examination, and give adequate support to instructors and to students as well.

Furthermore, student-related factors that affected the quality of learning assessment include students' poor background, hopelessness, lack of interest to learn and acquire knowledge, poor language in the medium of instruction, carelessness, and lack of preparedness for deep learning.

Instructor-related factors that affected the quality of learning assessment include instructors' lack of: basic knowledge and skills on learning assessment, basic experience and skills on quality test construction, interest to teach and do research on assessment matters, strictness in controlling exam undertakings, sufficient awareness on learning assessment tasks, fortitudes towards assessment requirements, capability and commitment to provide high quality individual feedback, and decisiveness on grading.

Finally, teaching learning resources and policy-related issues that affected the quality of learning assessment were attributable to unavailability and/or shortage of: modules for students of undergraduate programs, learning materials in [English] lab class, adequate infrastructure and/or learning facilities, and a well communicated assessment policy.

The majority of the respondents had the view that academic dishonesty and/or plagiarism was the most common phenomenon at universities in Ethiopia. A further scrutiny of the results has shown that academic dishonesty takes different forms, including: 1) cheating in exams, which has "now become unmanageable" (TR50), 2) copying from a friend or from notes (TR56), 3) plagiarism during assignments (TR56), and 4) lack of clear accountability and poorly managing group assignments.

In relation to the last issue, TR15 had the view that there "is a challenge to ensure individual accountability in group assignments/group work". In the same vein, TR50 indicated that "working group projects by individuals greatly impeded team learning". TR15 further indicated that "low staff knowledge and skills in assessment of learning" highly affected the quality of student learning assessment.

For TR59, "assessment items lack uniformity in terms of difficulty level". The same respondent further indicated that "[t]hough teachers assess real performances of students in tests, they hardly assign 'bad' grades".

For TR57 "change of policy from normative to criterion without developing rubrics, and criterion approach of testing" have influenced the quality of student learning assessment. TR65 also had the view that "absence of good educational background" on the part of students contributed to poor

quality of their test taking.

Discussion and Conclusions

Discussion. The majority (61%) of the instructors and students who properly completed and returned the questionnaires had teaching/research experience of above 10 years. No one from the teacher respondents had under 3 years teaching/research experience. Also, of those who had above 10 years of teaching/research experience, 85% and 43% were instructors and students respectively. By all measures, the respondents had reasonable teaching/research experiences that would enable them to judge hurdles that affected the quality of learning assessment.

Similarly, the majority of the teacher respondents (73%) had a PhD qualification, followed by 27% master's holders. When it comes to their rank, the majority (57%) of the teacher respondents were assistant professors, followed by 18% associate professorship and equally a lectureship. This, therefore, shows that the instructor respondents had reasonable qualifications and rank that would enable them to evaluate the level of student learning assessment vis-à-vis available theories and criteria.

The Cronbach's Alpha reliability analysis of the data on the hurdles to quality learning assessment has given a high reliability result (.909) signifying that the items in the questionnaire are correlated and are internally consistent for generating dependable evidence. Equally, analysis of the comparison of means of the academic units (CEBS and TEFL) has shown that the ratings of the CEBS has greater average mean than that of the TEFL with 3.5 and 3.4 respectively see Appendix 4).

It has also been found from the descriptive analysis that test administration and feedback related hurdles had the major effect, specifically failure to avoid plagiarism and/or academic dishonesty, inability to provide high quality individual feedback, and challenges to assess a diverse mix of learners (large class) respectively having average means of 4.0, 3.9, and 3.5. This is followed by learners and learning related hurdles in affecting learning assessment quality with average mean of 3.7, wherein learners work to earn good grades rather than to acquire required knowledge, has the highest average mean of 3.9 respectively with 1.1, and 1.2 SD and variance, and exam- oriented educational system, deficiency in connecting testing with learning, and ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain) each with average mean of 3.6 with varying SD and variance.

Furthermore, assessment procedures/practices, and learning resources and test constructs related hurdles have equal average mean and SD of 3.3 and 1.1 respectively. Regarding learning resources and test constructs related hurdles, lack of well-prepared module/learning resources for the courses taught, test construct underrepresentation, test construct irrelevant characteristics, and poorly constructed test items have average means of 3.8, 3.4, 3.3, and 3.2 respectively with similar SD except the last which has 1.3 SD. Equally, deficiency in sequencing items, lack of clarity in test directions, and ambiguous test statements have equally 3.1 average means.

Finally, the qualitative results have also provided a number of hurdles to maintain the quality of learning assessment at the universities in particular and/or in Ethiopian universities in general. A synthesis from 26 (16%) respondents indicates that the major hurdles could be brought under: 1) content-related, 2) test administration–related, 3) student-related, 4) instructor-related, and 5) learning resources- and policy–related issues and/or factors.

Overall, therefore, the quantitative and the qualitative results have given that: 1) failure to avoid plagiarism and/or academic dishonesty, 2) inability to provide high quality individual feedback, 3) lack of well-prepared module/learning resources for the courses taught, 4) learners' temptations to work for good grades, 5) failure of assessment procedures to direct learners to higher-order learning outcomes, 6) encouraging shallow learning, 7) learners' negligence to work for acquiring the required competencies, 8) challenges to assess a diverse mix of learners (large class), 9) deficiency in connecting testing with learning, 10) exam-oriented educational system, 11) test construct irrelevant characteristics, and 12) deficiency in sequencing items were the major hurdles to the quality of student learning assessment at the two universities in particular and at all the HEIs in the country in general.

Conclusions. With the purpose of investigating the hurdles to the quality of student learning assessment and their detrimental effects on student learning at two Ethiopian universities, issues related to assessment procedures or practices, learners and learning, learning resources and test constructs, and test admin and feedback were examined. The fact that the majority of the respondents had above 10 years of teaching/research experience, reasonable qualifications, and rank shows that they are appropriate to reasonably judge hurdles that could affect the quality of student learning assessment. The fact that the questions of the questionnaires had very high reliability (.909) also led to identifying a number of hurdles to quality learning assessment.

The fact that the grand average means the ratings of the instructors and students are equal (i.e., 3.5) with similar SD and variance of 1.1 and 0.1 (see Appendix 3) shows that the hurdles to the quality of learning assessment and their effects were widespread and commonly understood by pertinent parties at the two universities. Equally, the average means of CEBS is greater than that of TEFL signifying that the participants at the CEBS were aware of assessment practices and tenets due to their exposure to HDP, which is a practice-based training program for teacher educators at HEIs in Ethiopia (Firdissa, 2021; AAU, 2015).

Based on the quantitative and qualitative results and the discussions made so far, it can be concluded that the impediments related to assessment procedures/practices, learners and learning, learning resources and test constructs, and test admin and feedback affected the quality of student learning assessment at the two universities, in particular, and at all the HEIs in the country in general.

The case portends that very few conscious efforts were made by pertinent stakeholders to ascertain the quality of learning assessment at the universities. The assessed tasks were not aligned to clearly articulated LOs for the levels of achievement. As a result, assessment practices still focus on testing knowledge and comprehension and ignore the practices of developing and accessing judgments; and also fail to consider assessment peculiarities. There is also a failure to avoid plagiarism

and/or academic dishonesty, which is pilfering someone else's work without proper attribution, and it has real legal implications (Kirkpatrick, & Zang, 2011). The results can also be interpreted as a failure to fit the purposes of the different stakeholders including students, instructors, government, the society, employers, and other pertinent parties who have interest in education (Gerritsen-van Leeuwenkamp, Brinke, & Kester, 2017).

Failure to fit the purposes of the different stakeholders may be termed as inferior assessment quality. Inferior assessment quality for Gerritsen-van Leeuwenkamp, Brinke, and Kester (2017, p.1) "is a problem that has serious consequences for students, instructors, government, and society" at all levels of education.

Specifically, the consequences of inferior assessment quality at HEIs have significant detrimental effects on student learning, and their functioning in society. This could eventually lead to dichotomizing assessment, teaching, learning processes, and perceiving assessment as a one- shot activity instead of viewing it as an integral part of the overall teaching- learning process (Sanga, 2016). This in turn could promote inferior assessment quality.

Overall, the findings can be construed as a failure to put in place an effective assessment system, which constitutes deep learning with the focus being on providing guidance and feedback towards the achievement of learning outcomes. The case calls for reiterating a recommendation that Masuku, Jili and Sabela (2021) forwarded regarding the need for instructors to develop and explain to students a grading tool to ensure the quality of the assessment process wherein the assessment tasks align with the stated LOs and marking criteria are made clear to the students before the actual assessment.

Recommendations

The results have shown that the universities were inundated with a number of hurdles to learning assessment quality, which had detrimental effects on student learning and their effectiveness in life. Our era demands ensuring meaningful and useful assessment of student learning. This in turn calls for putting in place effective and quality assessment that constitutes deep learning guided by effective feedback towards the achievement of explicitly stated LOs. It has, therefore, been recommended that:

1. universities should make the utmost efforts to avoid plagiarism and/or academic dishonesty, devise mechanisms to make class size manageable, and ensure proper execution of test administration;
2. instructors should take the drivers' seat to direct learners to higher-order objectives (beyond shallow learning) by way of connecting testing with learning, and providing meaningful feedbacks that guide effective learning; and
3. the Ministry of Education along with the universities should put in place: a) strategies to avoid plagiarism and/or academic dishonesty, and b) well-prepared module/learning resources for the courses taught.

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.

References

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

AAU (2015). *Handbook for higher diploma program (HDP): Booklet-One, Module 1-3*. Addis Ababa: AAU Printing Press.

Ainslee, J. (2018). *The significance of quality assessment in the education system*. https://elearningindustry.com/quality-assessment-in-the-education-system-significance

Creswell, J.W. and Creswell, J.D (2018). *Research design: Qualitative, quantitative, and mixed methods approaches (Fifth Edition)*. SAGE Publications, Inc.: California 91320.

Dornyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies.* New York: Oxford University Press.

Edwards, F. (2013). Quality assessment by science teachers: Five focus areas. In *Science Education International, 24(2),* 212-226.

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.

Gerritsen-van Leeuwenkamp, K.J.; Brinke, D.J.; and Kester, L. (2017). Assessment quality in tertiary education: An integrative literature review. *Studies in Educational Evaluation, Vol. 55,* 94-116. https://dx.doi.org/10.1016/j.stueduc.2017.08.001

Horst, S. J. and Prendergast, C.O. (nd). The Assessment skills framework: A taxonomy of assessment knowledge, skills and attitudes. *In Research and Practice in Assessment, 15(1),* 1-25.

James, R.; McInnis, C.; and Devlin, M. (2010). *Assessing learning in Australian universities*. Australia: Centre for the Study of Higher Education.

Kirkpatrick, R. & Zang, Y. (2011). The negative backwash of exam-oriented education on Chinese high school students (53-70). In Kirkpatrick, R., Hlaing, H.L (eds.). *Language Testing in Asia* (LTA), 1(3), ISSN 2229-0443

Leeuwenkamp, G.; Brinke, D.; Kester, L. (2017). Assessment quality in tertiary education: An integrative literature review. *Studies in Educational Evaluation, Vol. 55*, 94-116. http://dx.doi.org/10.1016/j.stueduc.2017.08.001

Masuku, M.M.; Jili, N.N.; & Sabula,Th. (2021). Assessment as a pedagogy and measuring tool in promoting deep learning in institutions of higher learning. *International Journal of Higher Education, 10(2)*, 174-283. http://ijhe.sciedupress.com

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

Sanga, Ph. L. (2016). Implications of teacher educators' practices in assessment for student learning in Tanzania. *Makerere Journal of Higher Education, 8 (1),* 1–24. DOI: https://dx.doi.org/10.4314/majohe.v8i1.1

USDE [U.S. Department of Education]. (2017). *Reimagining the Role of Technology in Education: National Education Technology Plan Update.* https://tech.ed.gov

About the Author

Firdissa Jebessa Aga, Associate Professor, Addis Ababa University, Email: Firdessa.jebessa.aau.edu.et, Mobile Phone: (+251) 91152847, P.O. Box 1156, Ethiopia.

Appendices

Appendix 1

Cronbach's Alpha Reliability Statistics and Item-Total Statistics

Cronbach's Alpha	N of Items
.909	17

Hurdles to maintain learning assessment quality	Cronbach's Alpha if Item Deleted
1. Assessment procedures fail to direct learners to higher-order objectives	.906
2. Excessive assessment practices	.907
3. Assessment practices encourage shallow learning	.909
4. Learners work to earn good grades, rather than to acquire required knowledge	.905
5. Lack of well-prepared module/learning resources for the courses taught	.909
6. Inability to provide high quality individual feedback	.907
7. Challenges to assess a diverse mix of learners (large class)	.901
8. Inability to avoid plagiarism and/or academic dishonesty	.911
9. Exam-oriented educational system	.901
10. Deficiency in connecting testing with learning	.898
11. Test construct underrepresentation	.902
12. Test construct irrelevant characteristics	.899
13. Deficiency in sequencing items	.901
14. Ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain)	.900
15. Lack of clarity in test directions	.902
16. Ambiguous test statements	.901
17. Poorly constructed test items	.900

Appendix 2

Frequency analysis results

Impediments listed	No effect		Minor effect		Neutral		Moderate effect		Major effect	
	cnt	%	cnt	%	cnt	%	cnt	%	cnt	%
1. Assessment procedures fail to direct learners to higher-order objectives	4	2.8%	27	18.8%	39	27.1%	59	41.0%	15	10.4%
2. Excessive assessment practices	7	5.0%	33	23.4%	44	31.2%	43	30.5%	14	9.9%
3. Assessment practices encourage shallow learning	11	6.8%	26	16.1%	39	24.2%	54	33.5%	31	19.3%
4. Learners work to earn good grades, rather than to acquire required knowledge	5	3.4%	12	8.2%	26	17.8%	51	34.9%	51	34.9%
5. Lack of well-prepared module/learning resources for the courses taught	3	1.9%	17	10.6%	39	24.2%	47	29.2%	55	34.2%
6. Inability to provide high quality individual feedback	3	1.9%	11	6.8%	38	23.6%	51	31.7%	58	36.0%
7. Challenges to assess a diverse mix of learners (large class)	8	5.5%	16	11.0%	46	31.7%	51	35.2%	24	16.6%
8. Inability to avoid plagiarism and/or academic dishonesty	3	1.9%	16	9.9%	30	18.6%	45	28.0%	67	41.6%
9. Exam-oriented educational system	10	6.8%	17	11.6%	34	23.3%	49	33.6%	36	24.7%
10. Deficiency in connecting testing with learning	6	4.2%	17	11.9%	32	22.4%	58	40.6%	30	21.0%
11. Test construct underrepresentation	5	3.5%	23	16.0%	45	31.3%	47	32.6%	24	16.7%
12. Test construct irrelevant characteristics	8	5.6%	30	20.8%	38	26.4%	49	34.0%	19	13.2%
13. Deficiency in sequencing items	13	9.2%	25	17.6%	45	31.7%	49	34.5%	10	7.0%
14. Ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain)	10	6.9%	20	13.9%	31	21.5%	47	32.6%	36	25.0%
15. Lack of clarity in test directions	15	10.5%	33	23.1%	35	24.5%	42	29.4%	18	12.6%
16. Ambiguous test statements	12	8.3%	30	20.7%	48	33.1%	37	25.5%	18	12.4%
17. Poorly constructed test items	16	11.0%	31	21.4%	32	22.1%	40	27.6%	26	17.9%

Table Note: Cnt= Count

Appendix 3

Group Statistics of independent sample T-test for instructors and students

Differentiated List of Impediments	Respondent Group	N	$\overline{X}$	SD	Std. Error Mean
A. Assessment procedures/practices Related Impediments					
1. Assessment procedures fail to direct learners to higher-order objectives	Instructors	68	3.46	.999	.121
	Students	76	3.30	.994	.114
2. Excessive assessment practices	Instructors	68	2.88	1.100	.133
	Students	73	3.44	.943	.110
3. Assessment practices encourage shallow learning	Instructors	72	3.32	1.243	.146
	Instructors	68	3.46	.999	.121
Average Means	**Instructors**	**69**	**3.2**	**1.1**	**0.1**
	Students	**79**	**3.4**	**1.0**	**0.1**
B. Learners and Learning Related Impediments					
4. Learners work to earn good grades, rather than to acquire required knowledge	Instructors	69.0	4.3	1.0	0.1
	Students	77	3.6	1.1	0.1
5. Exam-oriented educational system	Instructors	69.0	3.7	1.2	0.1
	Students	77	3.4	1.1	0.1
6. Deficiency in connecting testing with learning	Instructors	69.0	3.6	1.1	0.1
	Students	74	3.6	1.0	0.1
7. Ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain)	Instructors	67.0	3.5	1.3	0.2
	Students	77	3.6	1.2	0.1
Average Means	**Instructors**	**69**	**3.8**	**1.1**	**0.1**
	Students	**76**	**3.6**	**1.1**	**0.1**
C. Learning resources & Test constructs Related Impediments					
8. Lack of well-prepared module/learning resources for the courses taught	Instructors	72	3.8	1.1	0.1
	Students	89	3.9	1.0	0.1
9. Test construct underrepresentation	Instructors	68	3.3	1.1	0.1
	Students	76	3.6	1.0	0.1
10. Test construct irrelevant characteristics	Instructors	68	3.2	1.1	0.1
	Students	76	3.4	1.1	0.1
11. Deficiency in sequencing items	Instructors	67	2.9	1.1	0.1
	Students	75	3.3	1.0	0.1
12. Lack of clarity in test directions	Instructors	69	2.9	1.3	0.2
	Students	74	3.3	1.1	0.1
13. Ambiguous test statements	Instructors	69	3.0	1.2	0.1
	Students	76	3.2	1.1	0.1
14. Poorly constructed test items	Instructors	69	3.1	1.2	0.1
	Students	76	3.3	1.3	0.2
Average Means	**Instructors**	**69**	**3.2**	**1.2**	**0.1**
	Students	**77**	**3.4**	**1.1**	**0.1**
D. Admin and Feedback Related Impediments					
15. Inability to provide high quality individual feedback	Instructors	72.0	4.0	1.0	0.1
	Students	89	3.9	1.0	0.1
16. Challenges to assess a diverse mix of learners (large class)	Instructors	68.0	3.5	1.2	0.1
	Students	77	3.4	1.0	0.1
17. Inability to avoid plagiarism and/or academic dishonesty	Instructors	72.0	4.2	1.0	0.1
	Students	89	3.8	1.1	0.1
Average Means	**Instructors**	**71**	**3.9**	**1.1**	**0.1**
	Students	**85**	**3.7**	**1.0**	**0.1**
Grand average Means	**Instructors**	**69**	**3.5**	**1.1**	**0.1**
	Students	**79**	**3.5**	**1.1**	**0.1**

$\overline{X}$ - Average Mean, SD(s) - Standard Deviation(s)

Appendix 4

Report of the Analysis of the Means for the CEBS and TEFL

Hurdles to learning assessment	**CEBS**	**Mean**	**variance**	**SD**
1. Assessment procedures fail to direct learners to higher-order objectives	CEBS	3.5	1.0	1.0
	TEFL	3.1	0.9	0.9
2. Excessive assessment practices	CEBS	3.2	1.2	1.1
	TEFL	3.1	1.0	1.0
3. Assessment practices encourage shallow learning	CEBS	3.5	1.3	1.2
	TEFL	3.3	1.4	1.2
4. Learners work to earn good grades, rather than to acquire required knowledge	CEBS	3.8	1.3	1.1
	TEFL	4.0	0.9	0.9
5. Lack of well-prepared module/learning resources for the courses taught	CEBS	3.9	1.0	1.0
	TEFL	3.7	1.5	1.2
6. Inability to provide high quality individual feedback	CEBS	4.0	0.9	2.0
	TEFL	3.7	1.3	1.1
7. Challenges to assess a diverse mix of learners (large class)	CEBS	3.5	1.1	1.1
	TEFL	3.3	1.2	1.1
8. Inability to avoid plagiarism and/or academic dishonesty	CEBS	3.9	1.2	1.1
	TEFL	4.1	1.2	1.1
9. Exam-oriented educational system	CEBS	3.6	1.4	1.2
	TEFL	3.4	1.5	1.2
10. Deficiency in connecting testing with learning	CEBS	3.6	1.2	1.1
	TEFL	3.7	1.2	1.1
11. Test construct underrepresentation	CEBS	3.4	1.1	1.1
	TEFL	3.4	1.0	1.0
12. Test construct irrelevant characteristics	CEBS	3.3	1.2	1.1
	TEFL	3.3	1.2	1.1
13. Deficiency in sequencing items	CEBS	3.2	1.2	1.1
	TEFL	3.0	1.1	1.1
14. Ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain)	CEBS	3.7	1.5	1.2
	TEFL	3.3	1.4	1.2
15. Lack of clarity in test directions	CEBS	3.1	1.4	1.2
	TEFL	3.3	1.4	1.2
16. Ambiguous test statements	CEBS	3.2	1.3	1.1
	TEFL	3.0	1.2	1.1
17. Poorly constructed test items	CEBS	3.3	1.7	1.3
	TEFL	3.0	1.3	1.1
Average means	**CEBS**	**3.5**	**1.2**	**1.2**
	TEFL	**3.4**	**1.2**	**1.1**

Appendix 5

Group Statistics of independent sample T-test for the Sexes

Differentiated List of Impediments	Respondent	N	$\overline{X}$	SD	SDEM
A. Assessment procedures/practices Related Impediments					
1. Assessment procedures fail to direct learners to higher-order objectives	M	133	3.4	1.0	0.1
	F	10	3.8	0.6	0.2
2. Excessive assessment practices	M	130	3.1	1.0	0.1
	F	10	3.6	1.2	0.4
3. Assessment practices encourage shallow learning	M	143	3.4	1.2	0.1
	F	14	3.5	1.0	0.3
Average Mean	M	135	3.3	1.1	0.1
	F	11	3.6	0.9	0.3
B. Learners and Learning Related Impediments					
4. Learners work to earn good grades, rather than to acquire required knowledge	M	135	3.9	1.1	0.1
	F	10	3.5	1.2	0.4
5. Exam-oriented educational system	M	134	3.6	1.2	0.1
	F	11	3.4	1.0	0.3
6. Deficiency in connecting testing with learning	M	132	3.6	1.1	0.1
	F	10	3.7	0.7	0.2
7. Ignoring the affective and psychomotor domains in assessment practices (failing to go beyond cognitive domain)	M	133	3.6	1.2	0.1
	F	10	3.3	0.8	0.3
Average	M	134	3.7	1.2	0.1
	F	10	3.5	0.9	0.3
C. Learning resources & Test constructs Related Impediments					
8. Lack of well-prepared module/learning resources for the courses taught	M	143	3.8	1.1	0.1
	F	14	3.6	1.1	0.3
9. Test construct underrepresentation	M	133	3.4	1.1	0.1
	F	10	3.9	0.6	0.2
10. Test construct irrelevant characteristics	M	133	3.3	1.1	0.1
	F	10	3.2	1.0	0.3
11. Deficiency in sequencing items	M	131	3.1	1.1	0.1
	F	10	3.1	0.9	0.3
12. Lack of clarity in test directions	M	133	3.6	1.2	0.1
	F	10	3.3	0.8	0.3
13. Ambiguous test statements	M	132	3.1	1.2	0.1
	F	10	2.7	1.3	0.4
14. Poorly constructed test items	M	134	3.2	1.1	0.1
	F	10	2.9	1.2	0.4
Average	M	134	3.4	1.1	0.1
	F	11	3.2	1.0	0.3
D. Admin and Feedback Related Impediments					
15. Inability to provide high quality individual feedback	M	143	3.9	1.0	0.1
	F	14	4.0	1.0	0.3
16. Challenges to assess a diverse mix of learners (large class)	M	133	3.5	1.1	0.1
	F	11	3.6	0.9	0.3
17. Inability to avoid plagiarism and/or academic dishonesty	M	143	4.0	1.1	0.1
	F	14	3.8	1.2	0.3
Average	M	140	3.8	1.1	0.1
	F	13	3.8	1.0	0.3
Grand average Mean	**M**	**136**	**3.6**	**1.1**	**0.1**
	F	**11**	**3.5**	**1.0**	**0.3**

$\overline{X}$ - Average Mean, SD(s) - Standard Deviation(s), SDEM-Std. Error Mean