Assessment Practices in the Developing World: Predictors of Assessment Practices in Ugandan Institutions of Higher Learning

Matovu Musa* & Ainol Madziah Zubairi**

Abstract: This study analysed the predictors of assessment practices employed by faculty at selected institutions in Uganda. An Assessment Practices Inventory Modified (APIM) scale was distributed to a sample of 350 academic staff selected from both private and public universities in Uganda. Random sampling was used to select the participants for this study. MANOVA and multiple regression analysis were employed for data analysis. Differences were revealed in faculty assessment practices according to their academic levels and specialisations, and not in the type of universities. Differences in academic levels cut across all the assessment practices sub-scales (design, administration, interpretation, and application) while in specialisations differences were only in assessment interpretation. It was also found out that academic levels and formal assessment course undertaken are the only significant predictors of the academic staff's assessment practices among the many hypothesised predictors (type of universities, specialisations, academic levels, class size, and assessment course). Generally, from the descriptive results of this study it has been noted that academic staff in Ugandan universities lack appropriate assessment skills in assessing their students. This has led to a recommendation that formal assessment training programmes should be made mandatory to all academic staff in universities in Uganda in order to improve their assessment skills to ensure quality in the way they assess students.

Keywords: Assessment practice, higher education, Uganda

_____

* Faculty of Education, Islamic University In Uganda, Mbale, Uganda, Corresponding Email: matovumousa@yahoo.com
** Assoc. Professor. Kulliyyah of Education, International Islamic University Malaysia. Email: ainol@iium.edu.my

Introduction

Assessment is a fundamental institutional structure that offers students academic justice by either locking or unlocking their academic potentials (MOE, 2010; Shohamy, 2004; Arter, 2003; Black & Wiliam, 1998b; Gipps, 1998; Madaus, Raczek, & Clark, 1997). Assessment is a significant component in teaching and learning structures whose proper integration improves the learning process (Alkharusi, 2012; NCCA, 2005; Benson, 2003; Biggs, 2003a; Roeber, 2002), and the effectiveness of teachers' ways of instruction (Ong, 2010; Orzolek, 2006).

Assessment should be non-threatening but affirming, and should have a positive rather than a negative relationship with the teaching and learning outcomes (NCCA, 2005). According to Fraizer (2007), assessment is the measure of instruction normally written by the teacher to evaluate students' achievements before, during, and after the instruction process. The assessment process entails gathering, interpreting, using, and communicating student learning progress and achievements. Assessment demonstrates the quality of teaching and students' learning (**Fatmawati**, 2011; Haken, 2006; Orzolek, 2006; NCCA, 2005; Popper, 2005; Brown, 2003; Alexander et al., 2003), and also used for learning improvement purposes (Martell & Calderon, 2005; Remesal, 2011; Walvoord & Virginia, 1998) among others. In higher education institutions, assessment is commonly through take home assignments, tests, examinations, project work, seminar papers, and tutorial participation (Polytechnic West, 2007).

Good assessment practices among academic staff should focus on unambiguous learning prospects and also ascertain realistic learning that does not propel students into rote learning. Lecturers' assessment practices should provide students with self motivation skills to learning, good study practices, and help them to receive timely feedback of what they have learnt (CCSSO, 2008; Polytechnic, 2007). Assessment must not only be aligned with the assessment design of the curricula and instructional strategies, but also, with assessment practices that measure student learning so as to be related to the course and program outcomes (Haken, 2006). Good student assessment practices among university academic staff should have a mission, reflect the learning system, be faculty driven, stakeholder owned, open and honest, and should be a process not one point event (Chapman University, 2001).

Assessing students in institutions of higher learning has a multifaceted utility that requires the understanding of student learning with familiarity of the principles, purposes, practices, application, and uses of different measures and evaluation (Anderson, 2005). This is because the assessment process is for collecting worth information that grants a foundation to appraise the learning objectives (Buzzetto-More & Alade, 2006). Assessment in higher education institutions does not only mean the traditional end of course marking and grading to measure students' performance, but it also involves consideration to soundness, consistency, stability, and impartiality in measures used to assess learning and the utility of collected information (Marvin et al., 1999; Meyer et al., 2009; Peterson et al., 1999b). It is through the quality of results from assessment that higher education institutions like universities can demonstrate their worth and the value of their graduates (Elander et al., 2006).

In higher education institution different types of assessment are used to assess students by the university academic staff. The most common used types of assessments used in the universities include *assessment for learning* which is also known as formative assessment, *assessment of learning* which is also equated to summative assessment, and *assessment as learning* (McDowell, 2011). In *assessment for learning*, consecutive assessment are undertaken onto the students across the whole study period of a given course. This is usually done to see how students can learn better in relation to the intended course objectives. Assessment for learning or summative assessment gives the students an opportunity to try out exercises several times before the final summative evaluation in order to improve their learning (McDowell, 2011). Also, during formative assessments students tend to scrutinise the details of the subject and understand the criteria for success in a given course. In general assessment for learning helps students to become autonomous in their learning, help them practice the skills they have acquired, gives immediate feedback to students, and learners get immediate guidance from their teaching instructors (Swaffield, 2011; Nicol & MacFarlane-Dick 2006). *Assessment of learning* refers to a single shot evaluation of student learning at the end of a study period to understand whether they have attained the required knowledge during the course (Gipps, 1994). In this case, university assessments can comprise of written examinations or projects that are done to mark the end of a learning session. This

much of a teacher centred approach than a student centred approach in assessing students. On the other hand another type of assessment used to assess students in universities in Uganda is *assessment as learning* which based on students evaluating themselves. Though this is the least used method of assessing students in universities in Uganda, sometimes students are given the opportunity to do self assessment to themselves and peers.

Assessment in universities in Uganda

In higher education institution in Uganda different types of assessment are used to assess students by the university academic staff. The most common used types of assessments used in the universities in Uganda include *assessment for learning* which is also known as formative assessment, *assessment of learning* which is also equated to summative assessment, and *assessment as learning* (McDowell, 2011).

In terms of the enrolment patterns in Ugandan universities, most students admitted are direct entrants from high schools while a few are admitted through mature entry examinations. During enrolment, public universities admit the best students who sit for high school examinations (Uganda advanced certificate examinations) on government sponsorship across various courses. The remaining students who merit university education but are not taken on government sponsorship, they are then shared between the public and private universities on private sponsorship. Both Science and non-Science courses are offered in private and public universities, although, core Science courses (engineering, medicine etc.) are dominant in public universities.

As academic staffs are key players in determining student academic progress through undertaking assessment on their learning, academic staff's assessment practices need to be in a faith of setting clear learning expectations of the student learning. Good assessments should help students attain good learning as well as academic achievements. The university academic staffs need to demonstrate good assessment practices that meet the institutions' mission and vision, reflect the student learning goals, and should be just to the stakeholders.

ASSESSMENT PRACTICES AMONG ACADEMIC STAFF

Assessment practices among academic staff in higher education have attracted the attention of several researchers in past few years and also

take on an essential position in teaching and learning process (Connoley, 2004; James, McInnis & Devlin, 2002). Assessment practices among academic staff in higher education institutions are entrenched in their perceptions and expertise to relate their understanding to the new innovative circumstances in assessing students (Watkins, 1998). Assessment practices vary among academic staff of higher education institutions as well as in their departments, faculties, and universities. Variances in academic staff's assessment practices makes maintenance of assessment standards in higher education institutions not attainable as assessment standards might differ in their own faculties and departments (O'Grady, 2006).

It has been noted that marking and grading practices among the academic staff vary considerably across and within the various universities, departments, and faculties (Warren-piper et al., 1996). If assessment practices can vary in a faculty where almost the content taught to the students is similar, then there should be a considerable variation in assessment practices across different faculties and universities. It should be a must for every university to maintain its assessment standards because they constitute measures to the university standards of teaching (O'Grady, 2006).

Fornari (2003) highlighted that fair assessment practices among instructors should lead to effective analysis of the institutional programs in achieving their goals and objectives in their disciplines. Also, fair assessment practices should be able to describe the criteria and methods used in an assessment. A formal assessment should be valid for what it is measuring, be reliable, honest and fair, objective, have ambiguity free content, and should be marked using clear and appropriate criteria (Ong, 2010; Silver, Palmer & DiFiore, 2008). For assessments to be valid, they should have considerable utility in terms of timing, managing feedback to students, tracking students' progress, among others. Also, highly achieving assessments should provide precise appraisals of learning outcomes to make it possible for instructors, educational managers, and other education stake holders not simply to inspect the learning process, but to make decisions in order to improve student learning (Fornari, 2003; Dietal, Herman & Knuth, 1991).

Chapman University (2001) reported that good assessment practices should be mission driven, reflect the values of the learning system, and

also, measure outcomes for improving the learning processes. Fairness driven assessment practices should be geared towards faculty goals, be stakeholder owned, open, honest, and should be a process not a project (Chapman University, 2001). O'Grady (2006) mentions that dependable assessment practices are not easily affordable practices and neither can one value them for money. Highly credible assessment practices might require expertise, experience, and availability of resources. Furthermore, time allocated to assessments and institution policies governing the assessment process might influence formal assessment practices among the institutions' academic staff (O'Grady, 2006). Higher education institutions like universities should have training programmes for faculty or departmental academic staff on assessment. These training programmes should suit the nature of the courses assessed in the faculties based on the already approved proficient assessment practices. Training would help the academic staff to acquire expertise and the required experience in assessment (O'Grady, 2006; Warren-Piper et al., 1996).

Training and development courses in assessment are necessary for the academic staff to acquire assessment skills. This is because they hold the sole responsibility of assessing student learning (Ong, 2010). From various studies conducted in assessment it has been highlighted that most academic staff lacked proper assessment knowledge and skills in evaluating student learning (Cizek et al., 1995; McMillan, 2001). With the assessment skills academic staff can make their learning objectives clear to students, involve students in marking, students to set targets learning process, and also students to take part in communicating their own results (Stiggins et al., 2006). Such student learning measures can clearly highlight what students have attained in an assessment, goals of the assessment and course, and the educational needs in relation to the practical situation (QAHE, 2006). In order to achieve the learning goals, assessment practices should be practically fair and should not discriminate students on irrelevant grounds other than study objectives. Academic staff with good assessment practices should provide timely informative and constructive feedback about student learning and also, should use grading processes that are transparent to reflect the actual picture of what students have achieved according to the objectives stated for the course (Polytechnic West, 2007).

Warren-Piper et al. (1996) suggested that in order for universities to maintain good standards of learning they should bench mark this at the quality of knowledge and competences possessed by their lowest passing graduates in the various education programmes. Warren-Piper et al. (1996) mentioned that the quality assurance of assessments is largely controlled by the lecturers' assessment practices. This means that university academic staff is part and partial of quality assurance in student assessment. It is also argued that paying attention to setting of standards for assessment practices among the academic staff is absolutely critical in higher education institutions' assessment (O'Grady, 2006). In another perspective on student assessment practices, favourable assessment practices among academic staff should be based on the Design-Implement-Review-Improve (DIRI) cycle (NIE, 2008). Academic staff with the knowledge of assessment should be able to design proportionate assessments. The academic staff should revisit the assessment process, be objectively driven, should have appropriate criteria in assessing students, and should be queried if the quality of their assessment is not both internally and externally suitable their institutions (NIE, 2008; O'Donovan et al., 2001).

Good policies of student assessment are necessary for good assessment practices among the academic staff (NIE, 2008). In any higher education institution, the absence of assessment policies is an indicator that assessment is not given its due precedence in that particular institution (Peterson et al., 1999). Higher education institutions should ensure that assessment activities are consistent with the learning outcomes, and should clearly be guided by reliable guidelines. Institutions should check on whether their academic staffs are competent in assessing students (Ainsworth & Viegut, 2006). For assessment policies, there should be guidelines for student assessment and standards for marking to which academic staff and students should adhere to. The judgements by the academic staff or examiners taken independently out of the institutional assessment policy should be set out publicly when institutions report on the assessments to the stakeholder (NIE, 2008).

According to Suskie (2002), there are seven steps in assessment practices which make up an excellent assessment. Suskie (2004) mentioned that academic staffs use seven steps to accomplish the assessment process for student learning, to attain success in assessment,

as well as, the learning objectives. Among the seven steps in a proficient assessment, Suskie (2004) highlighted that there is need to have clearly stated learning outcomes before the assessment exercise. The stated learning outcomes should match the learning and teaching of the students. Suskie (2004) also mentioned that academic staff should use as many different measures in assessing student learning. Different assessment methods favour given learning styles over others; such as, rote learning than understanding. Examiners should make sure that they give students different ways of attempting assessments in order to show what they learnt during the assessment process (Suskie, 2004). This can be used to curb assessment inaccuracies and academic staffs to make decisions based on professional judgement as a variety of assessments are done by the students in a particular course (Suskie, 2002). Taking assessment as a mission can make students learn better through preparing them how to do assignments, answer questions, and writing projects (Russell & Haney, 2000; Badger, 1999).

To guarantee that all students can equitably take on the given assessments, assessment regulations and prospects should be made available to students in the form of rubrics. Rubrics are scoring guidelines which explain an instructor's expectations to be used to assess student performance (Walls, 2003; Dodge & Pickette, 2001; Hansen, 1998). This would enhance their academic skills in performing the assessment tasks (Boud & Associates, 2010). Lecturers should provide rubrics which are authentic, put emphasis on application, and use of complex thinking skills (Montgomery, 2002; Andrade, 2000). On top of using genuine assessment practices by the academic staff, academic staff should engage students in order to build their confidence and abilities to undertake assessments (Anderson, 1988). There is need to interpret assessment results appropriately by choosing the most appropriate methods for analysis. For example, it might not be of value comparing students results against their peers' results. This would be most appropriate for out of school or curricular activities, because, it has diminutive justification for disallowing an adequate grade because a good number of his/ her classmates have performed well in the same assessment (Suskie, 2002).

As part of higher education institutions student learning, assessment practices among academic staff should also be based on procedures which give enough evidence, make justifiable inferences,

and summarise students' work accurately in order to attain the maximum grade. This can assist in knowing why students have not performed well, or, to identify where a revision in assessment techniques, pedagogy, or both can be done (Suskie, 2002). Poor results due to poor assessments practices among the academic staff call for the various stakeholder in higher education institutions to change the assessment procedures. This might lead to raising the entry standards for a higher course, or, to halt students who do not fully engage in the study programme (Hersh, 2004). Such measures together with compensation to border line marks indicate that assessment should always bring outstanding results in the learning process (Hersh, 2004).

Instructors should understand that their assessment practices have an effect to whoever encounters the assessment process. It has also been noted that academic staff have many challenges in their assessment practices which need urgent attention to save the students they assess and their institutions. Extant literature has shown that academic staffs in universities have consistently been argued that they need to possess assessment skills, and also, use them appropriately during the assessment process. It has also been highlighted in literature that academic staffs need not only to know how to mark or score tests and examinations, but, they need to understand how to use or apply the assessment results as well. It has again been noted that good assessment practices among academic staff help students to improve their learning by turning students' criticisms into positive suggestions, advising students on their next assignment, encouraging student self reflection, and explaining all comments made to the students' work.

It has been noted in literature that assessment practices among academic staff are manifested in assessment design, interpretation, and use of assessment results. On the other hand extant literature has enlightened that administration of assessment is also crucial in determining academic staff's assessment practices. The existence of literature about assessment administration but not studied as a factor which influences the assessment process and practices of the academic staff has highlighted a gap in the body of knowledge which research has tried to bridge. The finding of this research have bridged the gap by highlighting that now the assessment process comprise of four elements (design, administration, interpretation and application), which also influence the assessment practices among university academic

staff. The lack of assessment skills among academic staff in most of the assessment literature meant that academic staffs had inadequate skills in various assessment practices components. The findings in previous literature which mentioned that academic staffs were inadequate in assessment skills are supported by other literature which highlighted that it is imperative for all academic staff in institutions of higher learning to possess adequate assessment skills, and those who do not have to take on assessment training course.

THE PRESENT STUDY

According to various studies done in assessment practices, inconsistencies have been discovered in the findings on the factors influencing university academic staff's assessment practices (Boud & Associates, 2010; Zhang & Burry-Stock, 2003; Eckert et al., 2006; O'Grady, 2006; Bloxham & Boyd, 2007; Fisher et al., 2006; Popham, 2009). Assessment practices inconsistencies among academic staff in universities have been linked to students' increased failure, poor academic accountability, and divergences in grades among students (Bloxham & Boyd, 2007; Gibbs, 2006; Ebersole, 2009; Benjamin & Klein, 2006; Haken, 2006). In the different studies conducted to address the inconsistencies in assessment practices among the academic staff, many have been done on teachers in primary and secondary schools (Zhang & Burry-Stock, 1994, 1999, 2003; Frazer & Burry-Stock, 2008; Eckert et al., 2006), and a few in universities (Postareff et al., 2012; Samuelowicz & Bain, 2002; Xu, & Liu, 2009). In addition, none of the studies done on assessment practices has been directed towards analysing the factors influencing assessment practices among academic staff in Ugandan universities. Conducted studies in assessment practices have been done in areas of teachers' perception of their assessment practices (Plake, Impara, & Fager, 1993; Zhang, 1995; Zhang & Burry-Stock, 1994, 2003), teachers' assessment skills (Stiggin, Griswold, Wikelund, 1989), and their self perceived application of classroom assessment practices (Frazer, 2007). Therefore, there was a need to conduct a study to analyse the factors influencing assessment practices among university academic staff in Ugandan universities.

The study tested for the differences in assessment practices among the academic staff across various academic levels, universities, course specializations, class sizes, and assessment courses. Thus the

research question was: 'Are there differences in assessment practices among academic staff according to their academic levels, universities, and specialisations on the APIM sub-scale scores in Ugandan universities?'

Method

Sampling

The population of the study comprised of all universities in Uganda and their academic staff. Data was collected from both private and public universities in Uganda. There are 32 universities in Uganda; six (6) public and twenty six (26) private universities. Out of the 32 universities, three (3) public universities and three (3) private were selected for this study. Simple random sampling was used to select universities from their clusters on the basis of either being public or privately funded or owned. The population frame for the study comprised of all academic staff in the six randomly selected universities. In the selection of the academic staff from the different categories, the Krejcie and Morgan sample Table was used to select 350 academic staff who participated in the study.

The Assessment Practices Inventory Modified (APIM) was distributed to 350 academic staff of different academic levels and specialisations in the different universities in Uganda. Out of the 350 questionnaires distributed, 329 questionnaires were returned by the participants from whom 321 questionnaires were used for data analysis (see Table 3.1). 321 questionnaires were used for data analysis because some questionnaires (8 cases) which considerable missing information were excluded. The researcher used random sampling method to select participants from the different universities. As the purpose of the study was to analyse the factors influencing assessment practices among university academic staff, the participants were categorised according to their academic levels (teaching assistants, assistant lecturers, lecturers, associate professors, and professors), specialisations (arts, human sciences, science, and education), and also based on the type of the university funding received (Private or Public). This was to understand the differences in assessment practices between the different categories among the university academic staff.

Instrumentation

The Assessment Practices Inventory Modified (APIM) which was used as an instrument of data collection used in this study is based primarily on Assessment Practices Inventory Revised (API$_R$) scale which was developed by Burry-Stock & Frazier (2008). In the APIM scale some of the items were adopted entirely, others were adapted to suit the purpose of the study, while several other items were adopted from the original Assessment Practices Inventory (API) developed by Zhang and Burry-Sock (1994) and from the Institutional Support for Student Assessment Inventory by Peterson et al. (1997). The Burry-Stock & Frazier (2008) API$_R$ scale was adopted because it was the most appropriate and most recently developed assessment practices inventory which was used to collect information on teachers' perceived application of assessment practices.

The APIM which was an instrument used for data collection in this research consisted of 50 statements which described design, administration, interpretation, and application of assessment skills among academic staff in Ugandan universities. The items of the APIM were on a five point likert-scale, rated from not at all skilled to highly skilled (1 = *Not at all skilled,* 2 = A *little skilled, 3 = Some-what skilled, 4 =Skilled,* and 5 = *Highly skilled*) (see appendix 1A for the scale). This is similar to the five point rating likert-scale that was adopted in the API$_R$. The APIM questionnaire comprised of two sections. Section I asked for demographic information such as academic levels, types of universities, specialisations, and assessment courses attended among others. Section II of the APIM questionnaire comprised of 50 items on scale to which participants in the study rated their abilities in designing, administering, interpreting, and applying assessment results in a learning situation.

The APIM scale adopted thirty eight (38) items entirely from the API$_R$ scale without modification. 12 items from the API$_R$ were either modified (re-phrased) or dropped from the APIM scale based on the results of the pilot study which examined their suitability for use in this study. Five (5) of the affected items (Items 3, 6, 10, 17, and 27) were modified by rephrasing them to suit the context of this study, which is the university assessment. The other seven (7) items were dropped from the APIM scale due to similarity to other items, or, not being applicable in assessment at the university level. Out of the seven (7)

items dropped from the APIM scale, three (3) items (Items 5, 25, and 26) were replaced with items from the Zhang and Burry-Stock (1994) original API scale, while the other four (4) items (Items 2, 31, 42, and 44) were replaced with the items from Institutional Support for Student Assessment Inventory (Peterson et al., 1997).

VALIDITY AND RELIABILITY OF THE APIM SCALE

Validity of constructs is referred to as measurement inferences and there consequent usage to the constructs they purportedly measure (McMillan and Schumacher, 2010; Wiggins, 1992). The issue of validation is a must in the development of assessment instruments (Fraizer, 2007). Validation is also used in the revision and improvement of weak instruments, newly developed ones, and those which have been adopted or adapted like the APIM in this study. This study undertook validity and reliability checks on the APIM scale as there are no documented studies on the validity of the API_R scale in measuring assessment practices of academic staff in universities. The API_R was developed for use on secondary school teachers and since the APIM scale was primarily based on the API_R, validity tests were necessary to evaluate the appropriacy of its use in the new context.

The construct validity of the instrument (APIM scale) was tested in the form of convergent validity and discriminant validity. Convergent validity highlights the unidimensionality of the instrument, that is, the extent to which all items in the instrument are measuring the same construct. On the other hand, discriminant validity tests whether there are items which are irrelevant or inappropriate to the measurement of the construct. This helps to identify extraneous variables to the construct in the analysis process (Messick, 1994; AERA, 1999). Construct validity aims at finding out whether the items in the instrument measure or support the theory under investigation. Inadequate explication of the items would suggest that items insufficiently describe the constructs in the scale, and constructs do not properly support theory (McMillan & Schumacher, 2010). Discriminant validity highlights the degree to which items are distinctive from each other on related constructs of their respective variables (McMillan & Schumacher, 2010). Structural equation modelling was employed to validate the measurement model constructs, clarify the meaning of the constructs, and to establish the

extent to which the measurement model reproduces the data (Schmidt et al., 2005).

In terms of reliability, the API and API_R have been used in different studies and their reliabilities have been found to be adequate (Zhang, 1995; Zhang & Burry-Stock, 2003; Burry-Stock & Frazier, 2008). In the studies conducted using the API and API_R, the instruments have been shown to be reliable in measuring the intended construct and population at .940, .976, and .967 according to Zhang (1994), Zhang & Burry-Stock (2003), and Burry-Stock & Frazier (2008) respectively.

Statistical Analysis

Using PASW, Multivariate Analyses of Variance (MANOVA) and multiple regression analysis were conducted to answer research questions one and two respectively. Multivariate Analyses of Variance (MANOVA) was used to determine differences between independent variables on multiple dependent variables (assessment practices sub-dimensions). Testing for the differences, both multivariate and univariate tests were undertaken to determine the differences in assessment practices among the academic staff according to their universities, specialisations, and academic levels based on the APIM sub-scales (design, administration, interpretation, and application). On the other hand, multiple regression analysis was undertaken to test the significance of the regression model, and the effect of the significant independent variables at both the sample and population of the academic staff in universities in Uganda. This also helped to know the statistically significant predictors variables to assessment practices among university academic staff.

Analysis and Results

Are there differences in assessment practices among academic staff according to their academic levels, universities, and specialisations on the APIM sub-scale scores in Ugandan universities?

Table 1 MANOVA Descriptive Statistics

Variables			Public			Private		
			Mean	SD	n	Mean	SD	n
Specialisations								
Interpre-tation	Arts		3.37	.667	47	3.38	.555	51

	Human Sciences	3.28	.558	40	3.47	.650	53
	Sciences	3.20	.523	43	3.39	.494	49
	Education	3.75	.723	12	3.94	.658	18
	Total	3.33	.612	142	3.47	.600	171
Academic levels							
Design	T. Assistants	3.17	.410	13	3.32	.689	25
	A. Lecturers	3.01	.545	37	3.31	.631	68
	Lecturers	3.24	.575	85	3.40	.697	66
	Assoc Prof.	3.90	.660	5	3.77	.649	9
	Professors	4.18	.354	2	4.45	.289	3
	Total	3.21	.585	142	3.39	.679	171
Administration	T. Assistants	3.18	.376	13	3.25	.634	25
	A. Lecturers	3.29	.608	37	3.25	.536	68
	Lecturers	3.39	.534	85	3.47	.584	66
	Assoc Prof.	3.86	.509	5	3.74	.525	9
	Professors	3.55	.386	2	2.97	.053	3
	Total	3.36	.547	142	3.36	.579	171
Interpretation	T. Assistants	3.08	.550	13	3.36	.532	25
	A. Lecturers	3.32	.530	37	3.32	.560	68
	Lecturers	3.27	.571	85	3.49	.533	66
	Assoc Prof.	4.64	.490	5	4.52	.504	9

		Mean	SD	n	Mean	SD	n
	Professors	4.00	.386	2	4.09	.158	3
	Total	3.33	.612	142	3.47	.600	171
Application	T. Assistants	2.98	.539	13	3.18	.657	25
	A. Lecturers	3.14	.573	37	3.11	.601	68
	Lecturers	3.27	.619	85	3.26	.513	66
	Assoc Prof.	3.84	.433	5	3.67	.450	9
	Professors	4.05	.064	2	3.91	.158	3
	Total	3.24	.611	142	3.22	.582	171

Note: n = Number, SD = Standard Deviation, T. Assistant = Teaching Assistants, A. Lecturers = Assistant Lecturers, Assoc. Prof = Associate Professor.

MANOVA was used in examining the differences in the university academic staff's academic levels, universities, and specialisations which were compared to the APIM sub-scales. The scores of the APIM scale ranged from 250 which was the highest and 50 which was the lowest for the total score. MANOVA was conducted to establish the multivariate differences among the different academic staff variables on the sub-scales of the APIM. The four factors of student assessment (Design, Administration, Interpretation and Application) were the dependent variables (sub-scale scores) while academic levels, universities, and specialisations were the independent variables. Table 1 shows the descriptive statistics of the independent variables which were significant in the MANOVA analysis.

From the significant MANOVA tests of between-subjects effect results in Table 2, it was revealed that there were statistically significant differences in academic levels according to assessment design (F [4, 313] = 5.400, $p < .05$, $\eta^2 = .072$), administration (F [4, 313] = 3.212, $p < .05$, $\eta^2 = .044$), interpretation (F [4, 313] = 12.162, $p < .05$, $\eta^2 = .149$), and application (F [4, 313] = 3.512, $p < .05$, $\eta^2 = .048$). This indicates that the effect of academic levels on design, administration, interpretation and application of assessment results is different for the teaching assistants, assistant lecturers, lecturers, associate professors,

and professors. From the results in Table 2 it is shown that the difference in the effect of the academic levels on academic staff's assessment practices is different in terms of design, administration, interpretation, and application.

Table 2 MANOVA Tests of Between-Subjects Effect

Variable(s)	DV	F	df	MnSq	Sig	η^2	Power
Academic levels	Design	5.400	4	2.016	.000*	.072	.973
	Administra-tion	3.212	4	.982	.013*	.044	.825
	Interpretation	12.162	4	3.673	.000*	.149	1.000
	Application	3.512	4	1.174	.008*	.048	.861

Note: *$p<.05$, DV = Dependent Variable, MnSq = Mean Square, η2 = Partial Eta Squared, Power = Observed Power

The results of the individual partial eta squared (η^2) reveal a small but significant effect size. The academic levels explained 7.2% in the academic staff's assessment design, and 4.4% in the academic staff's administration of assessments. The partial eta squared also highlighted that the academic levels of the academic staff explained 14.9% of their interpretation of assessment and 4.8% in the way they applied results they got from the assessments.

POST HOC TESTS

The Post Hoc Test results allow an in-depth comparison of the assessment practice variables with the independent variables which are significant. In analysis of the Post Hoc Test results of the academic staff specialisations, the Tukey Post Hoc Test results revealed that statistically significant differences in the interpretation of assessment results existed in the specialisations of Education and Arts ($p = .000$; $p < .05$), Education and Human Sciences ($p = .000$; $p < .05$), and Education and Sciences ($p = .000$; $p < .05$) (Table 4.26). To highlight the differences further, the Bonferroni adjusted critical values to control type I error inflation was used to conduct the follow up. The Bonferroni adjusted critical values also revealed that the differences in the interpretation of assessment

results between Education and Arts (p = .000; p < .013), Education and Human Sciences (p = .000; p < .013), and Education and Sciences (p = .000; p < .013) existed as seen in Table 3. The descriptive results highlighted further that academic staff in the specialisation of Education (M = 3.85, SD = .690) were higher in levels of assessment interpretation than their counterparts in Arts (M = 3.38, SD = .611), Human Sciences (M = 3.38, SD = .604), and Science (M = 3.30, SD = .509) as in Table 1

Table 3 Multiple Comparisons of Interpretation in Specialisations

DV	Equal Variances Assumed	Comparisons		Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Interpretation	Tukey HSD	Education	Arts	.11466	.00*	.1944	.7871
		Education	Human Sciences	.11538	.00*	.1764	.7728
		Education	Science	.11554	.00*	.2666	.8638
	Bonferroni	Education	Arts	.11466	.00**	.1860	.7954
		Education	Human Sciences	.11538	.00**	.1680	.7812
		Education	Science	.11554	.00**	.2582	.8722

*Note: *p< .05, **p< .013, DV = Dependent Variables*

From Post Hoc Test results of academic levels in table 4, the Tukey Post Hoc Test results revealed statistically significant differences in the assessment design between associate professors and teaching assistants ($p = .006$; $p < .05$), associate professors and assistant lecturers ($p = .004$; $p < .05$), and associate professors and lecturers ($p = .026$; $p < .05$). The Tukey Post Hoc Test results confirmed that there were also significant differences in assessment design between professors and teaching assistants ($p = .002$; $p < .05$), professors and assistant lecturers ($p = .001$; $p < .05$), and professors and lectures ($p = .002$; $p < .05$) as in Table 4. According to the Bonferroni adjusted critical value to control type I error inflation in follow up analysis conducted, it was again highlighted that there were significant differences in the assessment design between associate professors and teaching assistants ($p = .007$; $p < .013$) and associate professors and assistant lecturers ($p = .005$; $p < .013$), but there was no difference between associate professors and lecturers ($p = .032$; $p < .013$).

Again, in the Bonferroni adjusted critical values it was observed that significant differences still existed between the assessment design of professors and teaching assistants ($p = .003$; $p < .013$), professors and assistant lecturers ($p = .001$; $p < .013$), and professors and lecturers ($p = .002$; $p < .013$). In an inspection of the mean scores in Table 1 for assessment design in the academic levels, the results indicate that associate professors ($M = 3.83$, $SD = .660$) and professors ($M = 4.32$, $SD = .322$) reported higher levels of assessment design than teaching assistants ($M = 3.25$, $SD = .550$), assistant lectures ($M = 3.16$, $SD = .588$), and lecturers ($M = 3.32$, $SD = .636$). There existed no difference in assessment design among associate professors and professors.

Table 4 Multiple Comparisons of Design in Academic Levels

DV	Equal Variances Assumed	Comparisons	Std Error	Sig	95% Confidence Interval	
					Lower Bound	Upper Bound
Design	Tukey HSD	Assoc Prof T. Assistants	.19103	.006*	.0230	1.0721

	Group 1	Group 2				
	Assoc Prof	A. Lecturer	.17385	.004*	.1362	1.0910
	Assoc Prof	Lecturer	.17070	.026*	.0387	.9762
	Professors	T. Assistants	.29068	.002*	.2759	1.8722
	Professors	A. Lecturer	.27969	.001*	.3722	1.9081
	Professors	Lecturer	.27774	.002*	.2713	1.7966
Bonferroni	Assoc Prof	T. Assistants	.19103	.007**	.0070	1.0881
	Assoc Prof	A. Lecturer	.17385	.005**	.1217	1.1055
	Assoc Prof	Lecturer	.17070	.032	.0244	.9905
	Professors	T. Assistants	.29068	.003**	.2515	1.8966
	Professors	A. Lecturer	.27969	.001**	.3487	1.9316
	Professors	Lecturer	.27774	.002**	.2480	1.8199

Note: *$p < .05$, **$p < .013$, DV = Dependent Variable, T. Assistant = Teaching Assistants, A. Lecturers = Assistant Lecturers, Assoc. Prof = Associate Professor

In other Post Hoc Test results as seen in Table 5, the associate professors differed in the administration of assessment with teaching assistants and assistant lecturers. The Tukey Post Hoc Test results revealed statistically significant differences in the assessment administration between associate professors and teaching assistants ($p = .013$; $p < .05$), and associate professors and assistant lecturers ($p = .010$; $p < .05$). The follow up results of the Bonferroni adjusted critical values which were to control type I error inflation showed that there were differences in the administration of assessment between associate professors and assistant lecturers ($p = .012$; $p < .013$), but no differences existed between associate professors and teaching assistants ($p = .015$; $p > .013$) as in Table 5. Assessment administration mean scores in the

academic levels showed that associate professors ($M = 3.80, SD = .517$) had higher skills of assessment administration than teaching assistants ($M = 3.22, SD = .505$), and assistant lectures ($M = 3.27, SD = .572$) as seen in Table 1.

Table 5 Multiple Comparisons of Administration in Academic Levels

DV	Equal Variances Assumed	Comparisons		Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Administration							
	Tukey HSD	Assoc Prof	T. Assistants	.17286	.013*	.0797	1.0290
		Assoc Prof	A. Lecturer	.15732	.010*	.0832	.9471
	Bonferroni	Assoc Prof	T. Assistants	.17286	.015	.0652	1.0435
		Assoc Prof	A. Lecturer	.15732	.012**	.0700	.9603

Note: *p< .05, **p< .013, DV = Dependent Variable, T. Assistant = Teaching Assistants, A. Lecturers = Assistant Lecturers, Assoc. Prof = Associate Professor

Looking at academic levels, the Tukey Post Hoc Test results highlighted that there were differences in the interpretation of assessments between the associate professors and professor with teaching assistants, assistant lecturers, and lecturers (Table 6). The Tukey Post Hoc Test results revealed statistically significant differences in the interpretation of assessments between associate professors and teaching assistants ($p = .000; p < .05$), associate professors and assistant lecturers ($p = .000; p < .05$), and associate professors and lecturers ($p = .000; p < .05$). Again, the Tukey Post Hoc Test results of interpretation in the academic level were statistically significant in the interpretations of assessments between professors and teaching assistants ($p = .022; p < .05$), professors and assistant lecturers ($p = .031; p < .05$), and professors and lecturers ($p = .049; p < .05$).

The Bonferroni adjusted critical values in Table 4.29 to follow up the Tukey Post Hoc Test results also indicated that there were differences between associate professors and teaching assistants ($p = .000$; $p < .013$), associate professors and assistant lecturers ($p = .000$; $p < .013$), and associate professors and lecturers ($p = .000$; $p < .013$). On the contrary, the Bonferroni adjusted critical values for interpretation in academic levels results indicated that they were insignificant in the interpretation of assessments between professors and teaching assistants ($p = .027$; $p > .013$), Professors and Assistant Lecturers ($p = .038$; $p > .013$), and professors and lectures ($p = .062$; $p > .013$) as seen in Table 4.29. For the descriptive statistics of interpretation of assessment results in academic levels it was noted that associate professors ($M = 4.58$, $SD = .497$) and professors ($M = 4.05$, $SD = .272$) were more skilled in assessment interpretation than teaching assistants ($M = 3.22$, $SD = .541$), assistant lecturers ($M = 3.32$, $SD = .542$), and lecturers ($M = 3.38$, $SD = .552$) as in Table 4.23. Associate professors and professors did not have much difference in interpreting assessment results.

Table 6 Multiple Comparisons of Interpretation in Academic Levels

DV	Equal Variances Assumed	Comparisons		Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Interpretation							
	Tukey HSD	Assoc Prof	T. Assistants	.17181	.000*	.8235	1.7670
		Assoc Prof	A. Lecturers	.15635	.000*	.8079	1.6666
		Assoc Prof	Lecturers	.15353	.000*	.7708	1.6140
		Professors	T. Assistants	.26143	.022*	.0736	1.5092
		Professors	A. Lecturers	.25154	.031*	.0426	1.4240
		Professors	Lecturers	.24979	.049*	.0026	1.3744
	Bonferroni	Assoc Prof	T. Assistants	.17181	.000**	.8091	1.7814

Assoc Prof	A. Lecturers	.15635	.000**	.7948	1.6797
Assoc Prof	Lecturers	.15353	.000**	.7580	1.6268
Professors	T. Assistants	.26143	.027	.0516	1.5311
Professors	A. Lecturers	.25154	.038	.0215	1.4451
Professors	Lecturers	.24979	.062	-.0183	1.3953

Note: *p < .05, **p < .013, DV = Dependent Variable, T. Assistant = Teaching Assistants, A. Lecturers = Assistant Lecturers, Assoc. Prof = Associate Professor

Finally, in the analysis of the Post Hoc Test results, the Tukey Post Hoc Test results again showed statistically significant differences in the application of assessment results by the different academic levels among the academic staff in universities. It is observed in Table 4.30 that the differences in academic levels in the application of assessment results existed between associate professors and teaching assistant (p = .007; p < .05), associate professors and assistant lecturers (p = .002; p < .05), and associate professors and lecturers (p = .035; p < .05). Differences also existed between professors and teaching assistants (p = .018; p < .05), professors and assistant lecturers (p = .014; p < .05), and professors and lectures (p = .048; p < .05) as shown in Table 7.

The Bonferroni adjusted critical value follow up results also explained that there were differences in the application of assessment results between associate professors and teaching assistant (p = .007; p < .013) and associate professors and assistant lecturers (p = .003; p < .013). There were no differences between associate professors and lecturers (p = .024; p > .013) in application of assessment results. Additionally, there were no statistically significant differences that existed between professors and teaching assistants (p = .021; p > .013), professors and assistant lecturers (p = .016; p > .013), and professors and lectures (p = .061; p > .013) as it is highlighted in the results of the Bonferroni adjusted critical values in Table 7. According to the descriptive mean scores in table 1, the associate professors (M = 3.76, SD = .442) differed in application of assessment results with teaching assistants (M = 3.08, SD = .598), assistant lecturers (M = 3.13, SD = .587), and lectures (M = 3.27, SD = .566). Again, the mean scores highlighted that the professors

(M = 3.98, SD = .111) were highly skilled in application of assessment results than teaching assistants (M = 3.08, SD = .598), assistant lecturers (M = 3.13, SD = .587), and lectures (M = 3.27, SD = .566) as in Table 1. Associate professors and professors did not greatly differ in application of the assessment results according to the revealed mean scores.

Table 7 Multiple comparisons of Application in Academic Levels

DV	Equal Variances Assumed	Comparisons		Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Application							
	Tukey HSD	Assoc Prof	T. Assistants	.18077	.007*	.1209	1.1136
		Assoc Prof	A. Lecturer	.16451	.002*	.1569	1.0604
		Assoc Prof	Lecturer	.16154	.035*	.0206	.9077
		Professors	T. Assistants	.27507	.018*	.0983	1.6089
		Professors	A. Lecturer	.26467	.014*	.1183	1.5718
		Professors	Lecturer	.26283	.048*	.1293	1.4222
	Bonferroni	Assoc Prof	T. Assistants	.18077	.007**	.1057	1.1288
		Assoc Prof	A. Lecturer	.16451	.003**	.1431	1.0742
		Assoc Prof	Lecturer	.16154	.044	.0071	.9213
		Professors	T. Assistants	.27507	.021	.0752	1.6320
		Professors	A. Lecturer	.26467	.016	.0961	1.5940
		Professors	Lecturer	.26283	.061	.0832	1.4443

Note: *p< .05, **p< .013, DV = Dependent Variable, T. Assistant = Teaching Assistants, A. Lecturers = Assistant Lecturers, Assoc. Prof = Associate Professor

In a univariate Analysis of Variance (ANOVA) done to follow-up the revealed differences in each of the independent variable on the dependent variables, statistically significant differences were noted as in Table 8. Assessment practices variables which were found to be statistically significant under the MANOVA were significant when tested independently using ANOVA. The design and academic levels (F [4, 277] = 5.400$, p < .001, \eta^2 =. 072$), and administration and academic levels (F [4, 277] = 3.212$, p < .05, \eta^2 = .044$) were statistically significant. Significant differences were also found using ANOVA among the academic staff assessment practices between interpretation and academic levels (F [4, 277] = 12.162$, p < .001, \eta^2 = .149$) and in application and academic levels (F [4, 277] = 3.512$, p < .01, \eta^2 = .048$). In analysing the results of the partial eta squared participants rated the variable of interpretation of assessments higher in importance ($\eta^2 = .149$) than any other variable in the model of the university academic staff assessment practices (Table 8).

Table 8 Univariate Tests of Between-Subjects Effects

Dependent variable	Independent variable	Sum of Squares	df	MnSq	F	Sig	η^2	Power
Design	Academic level	8.064	4	2.016	5.400	.000***	.072	.973
	Error	103.416	277	.373				
Administration	Academic level	3.927	4	.982	3.212	.013*	.044	.825
	Error	84.683	277	.306				
Interpretation	Academic level	14.691	4	3.673	12.162	.000***	.149	1.000
		83.651	277	.302				
Application	Academic level	4.497	4	1.174	3.512	.008**	.048	.861
		92.608	277	.334				

*Note: *p< .05, **p<.01, ***p<.001, df = Degrees of Freedom, MnSq = Mean Square, η^2 = Partial Eta Squared, Power = Observed Power*

The observed power shows the dangers which would be incurred when a null hypothesis is accepted (type II error). Also, the observed power reflects that the probability results would be significant in a sample drawn from a population within an effect size equal to the effect size of the APIM sample (Fraizer, 2007; George & Mallery, 2006). The results of the observed power in table 4.31 indicate that design had a 97.3% chance, administration had an 82.5% chance, interpretation had a 100 % chance, and application had an 86.1 % chance to find the same partial eta squared significant with a similar university academic staff sample. The observed power also highlighted that if a type II error was made in interpreting the results (accepting a null hypothesis when it is false), the probability of the type II error would be 97.3% for design, 82.5% for administration, 100% for interpretation, and 86.1% for application.

Discussion

The second research question inquired about the differences between assessment practices among the academic staff in Ugandan universities according to their academic levels, universities, and specialisations on the APIM sub-scales. The existence of four dependent variables (APIM sub-scales) and three independent variables (universities, specialisations, and academic levels) made it necessary to conduct a MANOVA in order to find out the differences (see Weinfurt, 1995). To understand the differences in assessment practices among the university academic staff, MANOVA was conducted to establish if there was any significant difference in assessment practices among the academic staff on the APIM sub-scales. The differences were determined according to the type of universities, specialisations, and academic levels of the academic staff in Ugandan universities.

The results of the difference between the assessment practices among the academic staff on the APIM sub-scales revealed that differences occurred in assessment design, administration, interpretation, and application with the academic levels of the university academic staff. Also, there were differences in interpretation of assessment results among the academic staff in terms of their specialisations. The differences that occurred in academic levels might have been due to the experience and training the different academic staff at different academic levels had acquired. Academic staffs at higher academic level are usually more experienced than those at lower academic levels. For

the differences that were observed in specialisations might have been due to the different trainings exposed to the academic staff during their university studies, or, as part of their staff development programmes. Some academic staff might have been exposed to assessment skills as a course while others might have not.

The MANOVA results revealed that there were significant differences between academic levels in assessment practices among the academic staff in universities. Also, the interaction between specialisations, universities, and academic levels yielded a significant difference in assessment practices among the university academic staff. This meant that academic staffs of different specialisations and with different academic levels were different in assessing students. In the ANOVA analysis, the difference in the interaction between specialisations, universities, and academic levels was realised to have occurred in specialisation where academic staff from education were different in assessing students from other academic staff in other specialisations. On the contrary, there were no statistically significant differences in assessment practices among academic staff in the different universities. According to the realised differences in specialisations, the academic staff from the education specialisation interpreted assessment results differently from those in other specialisations. This was because the academic staffs in the education specialisation undertake assessment courses during their university education. This was reflected in the MANOVA descriptive results as their assessment practices in interpretation of assessment results were different from other specialisations. The education specialisation assessment practices descriptive results were higher than those of the academic staff in the specialisations of arts, human sciences, and sciences.

According to the ANOVA results for the dependent variables as a follow up on the multivariate analysis, statistically significant variables in MANOVA were also noted to be significant in contrasts obtained using Post Hoc Test of both Tukey HSD and Bonferroni adjustment. From the Tukey HSD Test which was used for all the dependent variables and the Bonferroni adjustment critical values which were used as a follow up to compare the dependent variables, they were all found to be significant. This means that the differences existed in the independent variables on the dependent variables even at a more critical level. The Bonferroni adjustment was used to look at the results

of the significant dependent variables at a more critical level. Once the significant differences were established using the general model, further analysis was done to highlight where the significant difference was among the dependent variables with respect to the academic staff's characteristics (universities, specialisations, and academic levels). From the Post Hoc Test results of this analysis, it is highlighted that there are significant differences among the academic staff variables with respect to their assessment practices.

The results of Tests of Between-Subjects Effects revealed that significant differences existed between academic levels in all the sub-scales of the APIM (Design, administration, interpretation, and application). In the designing of assessments professors and associate professors differed with teaching assistants, assistant lecturers, and lecturers in their assessment practices. In the interpretation and application of assessment results, again, professors and associate professors differed in assessment practices with the teaching assistants, assistant lecturers, and lecturers. This reflected that professors and associate professors designed and interpreted assessments differently than the teaching assistants, assistant lecturers, and lecturers. This might have been due to their experience in undertaking assessment on students or, having taken a formal assessment course.

In the MANOVA Post Hoc Test results of administration of assessment it was revealed that only associate professors differed with teaching assistants and assistant lecturers in the way they administered assessments to students. But associate professors did not differ with professors and lecturers. In terms of specialisations a significant difference was discovered in the interpretation of assessment results between academic staff in the specialisation of education with academic staff from arts, human sciences, and science specialisations. From the Post Hoc Test results it can generally be concluded that academic staff of higher academic levels (professors and associate professors) differed in assessment practices with those having lower academic levels (teaching assistants, assistant lecturers, and lecturers). According to the academic levels, it was noted that the professors who had the highest academic levels differed with teaching assistants who had the lowest academic levels. From the research results it is revealed that the higher the academic level the better in assessing students. So, universities should help their academic staff to progress in their academic careers to

attain higher academic levels because it might improve their assessment practices. Also, taking a formal assessment course like the academic staff in the specialisation of education would improve the assessment practices of the academic staff in other specialisations (Arts, Human Sciences, and Sciences).

Conclusion

This study has set a milestone in the literature of university academic staff assessment practices as there are no studies that have been done before in Uganda, and using the APIM scale. Previous studies on assessment practices have been done in secondary schools (Herman & Choi, 2008; Fraizer, 2007; Mertler, 2004; Volante & Fazio, 2007; Wiliam et al., 2004; Stiggins, Griswold & Wikelund, 1989), and in other parts of the world. But this study has provided a national profile of the current assessment practices among the academic staff in Ugandan universities. So, this research has become a source of information for university administrators, faculties, departments and planning units for improving assessment practices among their academic staff. Again from the results of this research, a model has been postulated to explain the factors influencing assessment practices among academic staff in Ugandan universities.

Finally, this research has added knowledge to the existing information about the factors influencing assessment practices among university academic staff in Ugandan universities. At the same time, based on the recommendations of this research future researchers will be able to carry out further research on the factors influencing assessment practices among academic staff in Ugandan universities or elsewhere.

FUTURE RESEARCH

Although the results of this research have been found to be robust, they would require being followed-up with more studies. Having highlighted that the APIM scale could be used in determining assessment practices among university academic staff, it would be worthwhile to determine whether the interpretations from the APIM scale contribute to changing of the academic staff's assessment practices. More importantly, this would be useful to attempt to determine the best ways in which we can improve the assessment practices among the academic staff. This study permits further research into the other causal factors to assessment

practices among the university academic staff. It would also been useful to determine if academic staff assessment practices could be unearthed in their way, that is, academic staff revealing their actual assessment practices. This calls for more studies on factors influencing assessment practices among university academic staff using qualitative methods.

Considering methodology, most of the studies conducted on assessment practices have been done using quantitative research methods. There are a wide variety of qualitative approaches that could be used to determine the factors influencing assessment practices among academic staff in universities. The qualitative research methods use in-depth inquiry approaches which can go deeper into understanding why differences in assessment practices occur among the academic staff of the various academic levels and specialisations. More methods could also be used to find conclusive evidence about assessment practices among university academic staff according to their universities, academic levels, and specialisations. This research is a portion into the analysis of the factors influencing assessment practices among academic staff in Ugandan universities. It is hoped that at the completion of this research and other more researches to come the gap of missing information about the assessment practices among the academic staff in Ugandan universities would be bridged.

This research on the factors influencing assessment practices among university academic staff collected information from academic staffs who were presently teaching in Ugandan universities. It was important to examine the APIM scale and the hypothesized model across different categories of assessment stakeholders in higher education institutions other than the academic staff. In having participants from a wider scope it would give an appropriate model based on a heterogeneous population. This would also raise the number of participants as the population of the participants will be bigger than the population of only the academic staffs who teach in classrooms. So, more studies on the factors influencing assessment practices among university academic staff would be conducted on other university stake holders such as administrators, students among others. It can be noted that this research is not conclusive but an eye opener into more researches in assessment practices among the academic staff in universities in Uganda. There is a need for more studies because there are no major studies that have

been conducted on the factors influencing assessment practices among university academic staff out there, more so in the Ugandan context.

REFERENCES

Ainsworth, L., & Viegut, D. (2006). *Common formative assessments: How to connect standards based instruction and assessment*. California: Corwin.

Alexander, W. A., Trudy, W. B., Patricia, K, C., El-Khawas, E., Ewell, P. T., Hutchings, P., Theodore, J. M., Kay, M. M., Mentkowski, M., Miller, M. A., Thomas, M. E., & Wright, B. D. (2003). *Best Practices in Academic Assessment*. AAHE Assessment Forum, Exxon Education Foundation.

Alkharusi, H. (2012). Educational assessment attitudes, competences, knowledge, and practices: An exploratory study of Muscat teachers in the Sultanate of Oman. *Journal of Education and Learning, 1*(2), 217-232.

Anderson, H. M., Anaya, G., Bird, E., & Moore, D. L. (2005). A review of educational assessment. *American Journal of Pharmaceutical Education, 69*(1), 12.

Anderson, J. A. (1988). Cognitive styles and multicultural populations. *Journal Teacher Education, 24*(1), 2-9.

Arter, J. A. (2003). Assessment for learning: Classroom assessment to improve student achievement and well-being. *RIEPR2004, 24*.

Badger, E. (1999). Finding one's voice: A model for more equitable assessment. In A. L. Bailey & K.V. Drummond (Eds.). Who is at risk and why? Teachers' reasons for concern and their understanding and assessment of early literacy. *Educational Assessment, 11*(4), 149-178.

Benjamin, R., & Klein, S. (2006). *Assessment versus accountability in higher education: Notes for reconciliation*. Retrieved January 10

Benson, A. D. (2003). Assessing participant learning in online environments. *New Directions for Adult and Continuing Education, 100*, 69-78.

Biggs, J. (2003a). Aligning teaching and assessment to course objectives. Paper presented at ICHEd Conference: Teaching and Learning in Higher Education: New Trends and Innovations, April 13-17, in University of Aveiro, Portugal. Retrieved January 10, 2012.

Biggs, J. (2003b). *Teaching for quality learning at university: What the student does. Society for research into higher education (2nd)*. Buckingham: Open University Press.

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher education*: *A practical guide*. New York: McGraw Hill.

Boud, D., & Associates. (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council. Retrieved February 13, 2012.

Braney, B. T. (2010). An Examination of Fourth Grade Teachers' Assessment Literacy and its Relationship to Students' Reading Achievement. Unpublished PhD Dissertation. University of Kansas.

Brown, G. T. L. (2003). *Teachers' instructional conceptions: Assessments relationship to learning, teaching, curriculum, and teacher efficacy*. Paper presented at the Joint Conference of the Australian and New Zealand Associations for Research in Education (AARE/NZARE). Auckland: University of Auckland. Retrieved March 18, 2012.

Burry-Stock, J. A., & Frazier, C. H. (2008). *Revision of the Assessment Practice Inventory Revised (API$_R$): A combined exploratory factor analysis and polytomous IRT approach*. Paper Presented at the American Educational Research Association. New York: NY.

Burry-Stock. J. A., & Frazier, C. H. (2005). *Assessment practice inventory revised (API$_R$)*. Tuscaloosa, AL: University of Alabama.

Buzzetto-More, N. A., & Alade, A. J. (2006). Best practices in e-assessment. *Journal of Information Technology Education, 5,* 1-19.

Cizek, G. J., Fitzgerald, S. M., & Racher, R. E. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment, 3,* 159-179.

Connoley, R. (November, 2004). *Criterion-referenced assessment*. Faculty of Business & Law, Deakin University. Retrieved August 21, 2011.

Dodge, B., & Pickette, N. (2001). *Rubrics for web lessons*. Retrieved January 13, 2012.

Dunca, C. R., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *The Alberta Journal of Educational Research, 53*(1), 1-21.

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning.* Thousand Oaks, CA: Corwin Press.

Ebersole, T. E. (2009). Postsecondary assessment: Faculty attitudes and levels of engagement. *Assessment Update*, *1*(2), 13-14.

Eckert, T. L., Codding, R. S., Begeny, J. C., & Kleinmann, A. E. (2006). Assessment of mathematics and reading performance: An examination of the correspondence between direct assessment of student performance and teacher report. *Psychology in the Schools, 43*(3), 247 - 265.

Elander, J., Harrington, K., Norton, L., Robinson, H., & Reddy, P. (2006). Complex skills academic writing: A review of evidence about the types of learning required to meet core assessment criteria. *Assessment & Evaluation in Higher Education*, *31*(1), 71-90.

Fatmawati, A. (2011). *Using Portfolio to Assess Student Learning of Problem Solving.* Geelong-Victoria: Deakin University.

Fisher, D., Lapp, D., Flood J., & Moore, K. (2006). Linking literacy teaching with assessment: A continuing professional development initiative for secondary schools. *Literacy, 40*(2), 115-122.

Fraizer, C. H. (2007). Investigating Teachers' Self-Perceptions Applications of Classroom Assessment Practices using Assessment Practice Inventory Revised (API$_R$). Doctoral Dissertation, University of Alabama.

Gibbs, G. (2006). Why assessment is changing. In C. Bryan & K. Clegg, (Eds.), *Innovative Assessment in Higher Education* (pp. 11-22). London: Routledge.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment.* London: Falmer Press.

Gipps, C. (1998). Socio-cultural aspects of assessment. *Review of Research in Education, 24,* 355-392.

Haken, M. (2006). Closing the loop-learning from assessment. Presentation made at the University of Maryland Eastern Shore Assessment Workshop. Princess Anne: MD.

Hansen, E. (1998). Creating teachable moments and making them last. *Innovative Higher Education, 23*, 7-26.

Herman, J. L., & Choi, K. (2008). *Formative assessment and the improvement of middle school Science learning: The role of teacher accuracy.* Los Angeles, CA: Centre for Assessment and Evaluation of Student Learning. Retrieved January 10, 2012.

Hersh, R. H. (2004). Assessment and accountability: Unveiling value added assessment in higher education. AAHE National Assessment Conference, Denver, CO. Retrieved January 16, 2012.

James, R., McInnis, C., & Devlin, M. (2002). Assessing learning in Australian universities: Ideas, strategies and resources for quality in student assessment. Centre for the Study of Higher Education.

Jeffrey, B., & Woods, P. (1996). Feeling de-professionalised. *Cambridge Journal of Education, 26*(3), 325-343.

Madaus, G., Raczek, A., & Clark, M. (1997). The historical and policy foundations of the assessment movement. In A. Goodwin (Edn.), *Assessment for equity and inclusion: Embracing all our children,* (pp. 1-34)*.* New York: Routledge.

Martell, K., & Calderon, T. (2005). Assessment of student learning in business schools: What it is, where we are, and where we need to go next. In K. Martell & T. Calderon (Eds.), *Assessment of student learning in business schools: Best practices each step of the way* (pp. 1-22). Tallahassee, Florida: Association for Institutional Research.

Marvin, W. P., Marne, K. E., Catherine, H. A., & Derek, S. V. (1999). *Designing student assessment to strengthen institutional performance in doctoral and research institutions.* Centre for the Study of Higher and Postsecondary Education University of Michigan. Retrieved March 9, 2012.

McDowell, L., Wakelin, D., Montgomery, C., & King, S. (2011). Does assessment for learning make a difference? The development of a questionnaire to explore the student response. *Assessment & Evaluation in Higher Education, 36*(7), 749-765.

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research, 95*(4), 203-213.

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*(1), 20-32.

McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (7th edn.). Boston: Pearson.

Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education, 33*(1), 49-64.

Meyer, H. A., Rees, M., Fletcher, R. B., Johnston, P. M., & Anderson, H. (2009). *Tertiary assessment and higher education student outcomes:*

Policy, practice and research: An agenda for action. Retrieved March 7, 2012.

Montgomery, K. (2002). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching, 50*(1), 34-39.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

O'Donovan, B., Price, M., & Rust, C. (2001). The student experience of criterion referenced assessment through the use of a common criteria assessment grid. *Innovations in Learning and Teaching International, 38*(1), 74-85.

O'Grady, G. (2006). *Guide for good practice in assessment.* Ngee Ann Polytechnic, Singapore. Retrieved February 12, 2012.

Orzolek, D. C. (2006). The paradox of assessment: Assessment as paradox. *Research and Issues in Music Education, 4*(1). Retrieved September, 2011.

Peterson, M. W. (1988). The organizational environment for student learning. In J. S. Stark & L. A. Mets (Eds.), *Improving teaching and learning through research.* San Francisco: Jossey-Bass.

Peterson, M. W. (1997). *Inventory of institutional support for student assessment.* National Centre for Research to Improve Postsecondary Teaching and Learning. Michigan, Ann Arbor: University of Michigan, NCPI.

Peterson, M. W., & Spencer, M. G. (1990). Understanding academic culture and climate. In W. G. Tierney (Eds.), *Assessing academic climates and culture*. San Francisco: Jossey-Bass.

Peterson, M. W., Cameron, K. S., Knapp, A., Spencer, M. G., & White, T. H. (1991). *Assessing the organizational and administrative context for teaching and learning: An institutional self-study manual.* Michigan, Ann Arbor: University of Michigan, NCPI.

Peterson, M. W., Cameron, K. S., Mets, L. A., Jones, P., & Ettington, D. (1986). *The organizational context for teaching and learning: A review of the research literature.* Michigan, Ann Arbor: University of Michigan, NCPI.

Peterson, M. W., Einarson, M. K., Augustine, C. H., & Vaughan, D. V. (1999a*). Institutional support for student assessment: Methodology and results of a national survey.* National Centre for Postsecondary Improvement. Stanford, CA: Stanford University, NCPI.

Peterson, M. W., Einarson, M. K., Augustine, C. H., & Vaughan, D. S. (1999b). *Designing Student Assessment to Strengthen Institutional Performance in Comprehensive Institutions.* National Canter for Postsecondary Improvement. Stanford, CA: Stanford University, NCPI.

Peterson, M. W., Einarson, M. K., Trice, A. G., & Nichols, A. R. (1997). *Improving Organizational and Administrative Support for Student Assessment: A Review of the Research Literature.* National Centre for Postsecondary Improvement. Stanford, CA: Stanford University, NCPI.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.

Pilcher, J. K. (1994). The value-driven meaning of grades. *Educational Assessment, 2*(1), 69-88.

Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment Competencies of Teachers: A National Survey. *Educational Measurement: Issues and Practice, 12*(4), 10-12.

Polytechnic West (2007). *Higher Education Assessment.* Higher Education/International Business. Retrieved March 3, 2012.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4-11. DOI: 10.1080/00405840802577536.

Popper, E. (2005). Learning goals: The foundation of curriculum development and assessment. In K. Martell & T. Calderon (Eds.), *Assessment of student learning in business schools: Best practices each step of the way, 1*(2), 1-23. Tallahassee, Florida: Association for Institutional Research.

Postareff, L., Virtanen, V., Katajavuori, N., & Lindblom-Ylanne, S. (2012). Academics' conceptions of assessment and their assessment practices. *Studies in Educational Evaluation. 38*(3), 84-92.

Russell, M., & Haney, T. (2000). Bridging the gap between testing and technology in schools. *Education Policy Analysis Archives.* Retrieved March 3, 2012.

Samuelowicz, K., & Bain, J. D. (2002). Identifying academics' orientations to assessment practice. *Higher Education, 43*(2), 173-201.

Schmidt, G. A., Hoffmann, G., Shindell, D. T., & Hu, Y. (2005). Modelling atmospheric stable water isotopes and the potential for constraining cloud processes and stratosphere-troposphere water exchange. *Journal of Geophysical Research, 110*, D21314, DOI:10.1029/2005JD005790.

Silver, B., Palmer, C., & DiFiore, F. (2008). Using common assessment data to predict high stakes performance: An efficient teacher-referent process. Paper presented at the annual meeting of the American Educational Research Association New York, New York.

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, *86*(1), 22-27.

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22*(4), 271-286.

Stiggins, R. J., & Chappuis, J. (2005). Using student-involved classroom assessment to close the achievement gaps. *Theory into Practice, 44*(1), 11-18.

Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. J. (2006). *Classroom assessment for student learning: Doing it right, using it well.* Oregon: Educational Testing Service.

Stiggins, R. J., Griswold, M. M., & Wikelund, K. R. (1989). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement, 26*(3), 233-246.

Suskie, L. (2002). Fair assessment practices: Giving students equitable opportunities to demonstrate learning. *Adventures in Assessment, 14*.

Suskie, L. (2004) *Assessing student learning: A common sense guide*. Bolton, MA: Anker.

Swaffield, S. (2011). Getting to the heart of authentic Assessment for Learning. *Assessment in Education: Principles, Policy & Practice, 18*(4), 433-449.

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education, 30*(3), 750-770.

Walls, K. (2003). *Developing rubrics. The assessment quarterly: A learning, assessment, and evaluation tool. Heartland community college assessment committee.*

Walvoord, B. E., & Virginia, J. A. (1998). *Effective grading: A tool for learning and assessment.* San Francisco: Jossey-Bass

Warren-Piper, D., Nulty, D. D., & O'Grady, G. (1996). Examination practices and procedures in Australian universities. Higher Education Division, Department of Employment Education and Training. Evaluations and Investigations Program. AGPS, Canberra.

Watkins, D. (1998). Assessing university students in Hong Kong: How and why? In D. Watkins, C. Tang, J. Biggs & R. Kuisma (Eds.), *Assessment of university students in Hong Kong: How and why, assessment portfolio, students' grading* (pp. 5-27). Hong Kong, ESEP: City University of Hong Kong.

Weinfurt, K. P. (1995). Multivariate analysis of variance. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate analysis*. (pp. 245-276). Washington, DC: American Psychological Association.

Wiggins, G. (1998). *Educative assessment*. San Francisco: Jossey-Bass.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education, 11*(1), 49-65.

Zhang, Z. (1995). *Investigating teachers' perceived assessment practices and assessment competencies on the API.* Tuscaloosa, AL: The University of Alabama.

Zhang, Z., & Burry-Stock, J. A. (1994). *Assessment practices inventory.* Tuscaloosa, AL: The University of Alabama.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342.