An investigation into the Predictive Validity of English Language Assessment at the Technical College of Civil Aviation & Meteorology (TCCAM) in Libya:

▪ Bashir Ghith Mahfoud*

Abstract:

This correlational study attempted to investigate the predictive validity of English language assessment of the Foundation Program (FP). More specifically, it investigated how well students' scores in the FP assessment could predict their scores in First Year (FY) academic courses in Meteorology, Communication, Air Traffic Control and Air-craft maintenance departments. It also investigated the predictive validity of the assessment of the General English Skills (GES) Academic English Skills (AES). The findings revealed differences in the predictive validity of the FP assessment across the four (specializations) departments by gender, specialization and self-evaluation at the Technical College of Civil Aviation and Meteorology (TCCAM) in Libya. It was carried out over two academic terms, in the first term , the test grades of 174 Foundation Programme students in English language courses were obtained, in the second term, the test grades of 153 First Year students in the four academic courses were obtained. The predictive validity of English language assessment regarding academic achievement was found to be $r=0.3$, $p < 0.01$, and the strength of the predictive validity significantly varied among specialization and self-evaluation groups specifically.

Keywords:, Predictive validity, academic programs, aviation, foundation program.

الملخص:

حاولت هذه الدراسة الارتباطية التحقق من الصلاحية التنبؤية في تقييم اللغة الإنجليزية لدى (برنامج فصل الاساس) (FP). وبشكل أكثر تحديدًا ، فقد بحثت هده الدراسة في مدى ارتباط نجاح درجات الطلاب أثناء تقييم (برنامج فصل الاساس) (FP) في التنبؤ بنتائجهم في الدورات الأكاديمية للسنة الأولى (FY) في الأقسام العلمية و هي: قسم الأرصاد الجوية وقسم الاتصالات وقسم المراقبة الجوية وقسم صيانة الطائرات. كما بحثت هده الدراسة في الصلاحية التنبؤية لتقييم مهارات اللغة الإنجليزية العامة (GES) و مهارات اللغة الإنجليزية الأكاديمية (AES). كشفت نتائج هده الدراسة عن

* lecturer of General English studies, Technical College of Civil Aviation & Meteorology, Tripoli, Libya.

وجود اختلافات في الصلاحية التنبؤية في التقييم لدى (برنامج فصل الاساس) (FP) عبر الأقسام الأربعة العلمية، من ناحية الجنس والتخصص والتقييم الذاتي في كلية تقنية الطيران المدني والأرصاد الجوية (TCCAM) في ليبيا. تم إجراء الاختبار على فصلين أكاديميين ، في الفصل الأول ، تم الحصول على درجات اختبار 174 طالبًا في (برنامج فصل الاساس) في مقررات اللغة الإنجليزية. وفي الفصل الثاني تم الحصول على درجات اختبار 153 طالبًا في السنة الأولى في التخصصات الأكاديمية الأربع. و أثبت الدراسة بأن الصلاحية التنبؤية لتقييم اللغة الإنجليزية فيما يتعلق بالتحصيل الأكاديمي كانت r = 0.3، p <0.01 ، وتفاوتت قوة الصلاحية التنبؤية بشكل كبير بين مجموعات التخصص والتقييم الذاتي على وجه التحديد.

1.Introduction

Language Proficiency in English language assessment and how it is to be assessed has not been given enough attention in higher education research in the EFL Arab context. Despite the fact that English language is widely used as a medium of instruction and is seen as a conditional requirement for entrance to higher education across the world, it is clear that English language assessment in Libya has not been influenced by some modern approaches and techniques in the field of language testing and assessment. Therefore, there has been an urgent need to investigate the predictive validity of English language assessment, more specifically in the foundation program. In the Foundation Program at the technical colleges of sciences and higher technical institutes of science in Libya, all forms of language assessment and tests are used to assess students› language skills. However, although this combination of assessment according to (Hamilton, 2003) increase the assessment validity and results in better academic achievement. Another aspect that brings about different responses is the correlation between language proficiency and academic achievement or what is known as the predictive validity of language assessment. Research studies on this area have conflicting views towards this degree of effect/correlation and that some researchers pointed out that this sort of effect/correlation is not a fruitful line of research. Therefore, these arguments are still debatable and are open to investigation the area of language assessment in higher education and will be the driving force in this study to in investigate the predictive validity of language assessment more specifically in the Foundation Program at the technical college of civil aviation and meteorology, Tripoli, Libya.

•Problem of the study

Some language teachers claim that students' proficiency in English language has an impact on their academic achievement. Therefore, some teachers of academic courses in TCCAM claim that students' underachievement is attributed

to their inadequate English language skills believing a positive correlation between students' proficiency in (English) and their academic achievement. Other language teachers claim that gaining higher scores in English result in a better academic achievement, and that students' failure in FY first year might be attributed to their inadequate English language abilities. As a result, there is a need to investigate the role played by students' language proficiency in academic achievement at the Technical College of Civil Aviation and Meteorology (TCCAM) as this according to the authors constitutes the driving force of this research study.

•Research Questions

This study will, therefore, investigate the predictive validity of FP assessment by correlating students' scores in FP assessment and their scores in FY academic courses.

1. Does students' performance in English language assessment in FP correlate positively with their language performance in FY academic courses assessment?
2. Does the strength of correlation between the language proficiency and academic achievement vary significantly when students' scores in GES assessment or AES assessment are compared?
3. Do the student groups by gender, self-evaluation and specializations show any significant differences in the correlations between language proficiency and academic achievement?

2.Literature Review

With reference to some research articles on English language assessment in EFL higher education, Ross (2008,) states that there is an increasing use of test scores in determining access for admission to higher education, and that proficiency in the English language has also become the key for success in the labour market. Following this phenomenon, proficiency in the English language has been considered a criterion to access most higher education programs in Libya, and the English language assessment plays a critical role in admission to higher education. However, there is always a question about how predictive student scores in English language assessment are of student success in future academic study. This paper investigates the predictive validity of student scores in English language assessment in terms of academic achievement in 4 departments at the technical

College TCCAM in Libya in the following academic courses in Meteorology (M) Communication (C), Air traffic control (ATC) and air craft maintenance (ACM).

•Assessment Validity and Predictive Validity

Test validity involves five separate validities (i.e., face, content, predictive, concurrent, and construct) which constitute the psychometric characteristics of a test. These validities are sometimes viewed as internal, external and construct validities. According to (Martuza, 1977), The internal validity of a test includes face validity and content validity. The external validity of the test reflects its concurrent validity and predictive validity. Hughes (2003) says that face validity of a test signifies its suitability for its purposes, content validity means that an assessment is a reflection of the skills and content that is supposed to test, concurrent validity of a test is established when a test correlates well with another test that similarly measures the same constructs and is taken at the same time. Predictive validity refers to the degree to which a test predicts future performance of test takers, and that construct validity indicates that a test assesses the skills and abilities (i.e., constructs) that it is supposed measure. (Bachman & Palmer, 1996).

The Content validity of a test deals with determining the relationship between test tasks and specific learned content while the construct validity of a test is about identifying the relationship between test tasks and theoretical constructs of language proficiency irrespective of learned materials. (Bachman, L. F. (2004).

The reliability of a test is viewed as a distinct quality from validity but both are necessary for a good test according to Bachman and Palmer (1996). A test's reliability is established if similar scores are obtained when the same test is administered to two groups with the same language abilities or administered to one group at different times (Hughes, 2003).

Also, Harrison (1983) believes that the reliability of a test is its consistency. Messick (1989) redefined validity as a unitary concept that involved multiple facets. He also added that the consequences of a test should be included as an aspect of validity. He stated that the consequences of a test constituted an inherent facet of any evaluative judgment of the adequacy and appropriateness of interpretations and actions based on test scores.

Test Validity according to Messick, (1989, p.6) is known as "a unified though faceted concept", and validation is seen as a "scientific enquiry into score meaning". Also Bachman (2004) supports the premises of validity in Messick's

view stating that test validity reflects the quality of the interpretation not scores, and that validity is a question of a degree and is not always static. He also adds that test validity is specific to a particular use, and that validity of a test consists of a comprehensive evaluative judgment. In this view, test validation is viewed as the process of collecting information that reflects the appropriateness and correctness of the interpretations of the test scores (Messick, 1989; Bachman & Palmer, 1996).

•Predictive Validity of IELTS & TOEFL tests.

Although there is a widespread theoretical consensus of the unitary view of validity that consists of several 'aspects', research studies on the predictive validity of language assessment and testing are still conducted for their own purposes. In other words, estimating students' performance by correlating results on two different assessment instruments separated by a specific time difference.

Over the past two decades, Graham (1987) stated that the results obtained from predictive validity studies on language tests as inconsistent, and today the same conclusion can be made based on the following summary which consists of some research studies on the predictive validity of internationally standardized language tests as gatekeepers to higher education universities and colleges, namely (i.e., IELTS, TOEFL) language tests.

Table 1. Studies on Predictive Validity of IELTS

Study	Country	Number of participants	Type of correlation	Correlation Strength
Elder (1993)	Australia	32 International Students	IELTS & Administrator Ratings	0.5*
Cotton & Conrow (1998)	Australia	33Undergraduate & Postgraduate Students	IELTS & GPA -	-0.24*
			IELTS & Staff Ratings	0. 15*
			IELTS & Student Self-assessment	-0.28*
Huong (2001)	Australia	320 Vietnamese Post- & Undergraduate Students	IELTS & GPA	0.30*
Kerstjen & Nery (2000)	Australia	113 International Students	IELTS & GPA	Non-Significant

Feast (2002)	Australia	101 International Students	IELTS & GPA	0.39*
Woodrow (2006)	Australia	62 Students 15Teachers in Faculty of Education	IELTS & Teacher Evaluations	0.40*
Breeze & Miller (2008)	Spain	289 Undergraduate Spanish Students	IELTS & GPA (Humanities)	0.34*
			(Law)	0.28 **
			(Medicine)	0.25*
Yen & Kuzma (2009)	Britain	61Chinese Students (Bussiness)	IELTS & GPA	0.46**

* p<0.05, ** p<0.01

Table 2. Studies on Predictive Validity of TOEFL

Study	Country	Number of Participants	Type of Correlation	Correlation
Vinke & Jochems (1993)	Netherlands	90 Indonesian Students (Engineering)	TOEFL & GPA	TOEFL< 450 =0.09**
				TOEFL > 450 = 0.5**
Cho & Bridgeman (2012)	USA	2594 Graduate & Undergraduate Students	TOEFL & GPA	Graduate Students =0.16*
				Undergraduates = 0.18*
Al-Musawi & Al-Ansari (1999)	86 Undergraduate Students (English Language Studies)	TOFEL & GPA/ENGPA***		GPA = 0.50**
				ENGPA =0.70**
Maleki & Zangani (2007)	Iran	50 Undergraduate Students (English language studies)	TOFEL & GPA	0.48*

*p < 0.05, **p < 0.01, *** Students' GPA in English Language Major

• Effects of the Predictive Validity of Language Assessment Specializations

Several studies have reported a variance in the strength of the predictive validity values of language assessment across different specializations. In a study by Jochems et al. (1996) showed that the value of the predictive validity varied considerably from r = 0.32 to r = 0.46 in Computer Sciences and Engineering majors. Their study investigated the correlations between Dutch language proficiency as a second language (Dutch was the medium of study) and academic achievement. Another study by Lynch (2000) revealed that there was some variance in the correlation coefficient between the English language test used at the University of Edinburgh and students' average scores in the academic courses across the students' different areas of study. For example, the correlation coefficients in the Arts and Veterinary Medicine were non-significant, whereas, the coefficients in Social Sciences, Law, Science and Engineering were r = 0.23, r = 0.32 and r = 0.24 respectively. Likewise, a correlational study by Huong (2001) who claimed that the correlation between language proficiency and academic achievement in the linguistically demanding disciplines (e.g., TESOL) was stronger than it was in the less linguistically demanding disciplines (e.g., Engineering). Also, Woodrow (2006) found in his study that the correlation coefficient between the students' bands in IELTS and their GPA in TESOL courses to be $r = 0.4, p < 0.01, n = 62$. In the English language domain. Similarly, Cope (2011) reported that the value of the correlation varied between different specializations when he investigated the predictive validity of three types of English language entry programs.

• Self-Evaluation of Language Skills

Few studies on predictive validity have attempted to investigate the potential effect of the students' self-evaluations to the strength of the predictive validity of language assessment (Powers, Kim, & Weng, 2008). In another study by Xu (1991) who investigated the correlation between students' self-evaluations of their language proficiency and self-reported academic difficulties found some correlation between TOEFL scores and self-reported academic difficulties. His finding revealed that the students' self-evaluation was a better predictor of the perceived academic difficulties than were their TOEFL scores. although Xu's main purpose of his study was on perceived academic difficulties, his findings shed some light on the role of self-evaluation in understanding possible future academic difficulties.

3.Methodology

According to Messick (1989) validation is best described as a scientific inquiry into score meaning. This study will, therefore, investigate the score interpretation that assumes a positive correlation between student scores in English language assessment and their scores in academic courses taught in English.

In this study, students' grades in the Foundation Program (FP) assessment at the technical College of civil aviation and meteorology (TCCAM) are correlated with their grades in the academic courses of their First Year (FY). Students started the FP in February 2017 and then began their academic FY in September 2017. The predictive validity of FP assessment will be investigated with reference to gender, self-evaluation, and specialization.

• Data collection

This correlational study is quantitative in nature. In other words, only tests were used to collect the data about the students' grades over two terms.. The sample started out with 174 students on the FP, and then it decreased to 153 students in the FY due to the students' inability to pass the foundation program. The size of the sample included in the statistical tests to investigate the correlations was about *N*=153. Therefore, the sampling technique included only 153 students in four specializations (departments) in the technical college of civil aviation & meteorology (TCCAM).

Table 3. Assessment Instruments in the Foundation Program Courses

Course	Assessment instruments	% of course total	% of foundation program total
General English	Midterm test	40%	50%
	Final test	60%	
Academic English	Presentation	50%	50%
	Report writing	50%	

The FP is a pre-sessional program that consists of two hours of mathematics and/or computer skills courses in each semester. The English language program is divided into two major courses, the General English skills (GES) and Academic English Skills (AES)

AES assessment includes continuous assessment (i.e., a report writing and presentation) as shown in the Table 3. GES assessment includes tests which were centrally developed, although individual teachers at the college (TCCAM) participated in the process of writing, reviewing and rewriting these tests. The teachers participated in standardization and moderation training sessions before marking the writing component of GES tests.

Though these tests were constructed and reviewed following rigid procedures. They were not trailed before use and their reliability is uncertain. Similarly, AES assessment used rating scales to evaluate student performances in report writing and presentation. However, no sessions in standardizing the implementation of the rating scales in AES were given to teachers. For the purposes of this study, the term (FP) refers to the English component only (i.e., GES and AES). Student scores in Physics, Mathematics or Computer Skills were not included in this predictive validity study of FP assessment.

'Proficiency' and 'Achievement'

Before investigating the relationship between the students' language proficiency and their academic achievement, it is crucial to explain how the concepts 'language proficiency' and 'academic achievement' were operationalized. Students' English language proficiency was represented by their average grades on the two FP English language courses (i.e., AES and GES). Likewise, the students' achievement in academic courses was represented by their average grades on the FY academic courses in the first semester.

Another point to clarify is how the Grade Point Average (GPA), used in TCCAM to report students' achievement was employed in this study. GPA stands for "the Grade Point Average of the numeric value of the entire results that the student has passed or failed in that semester" (TCCAM, 2017). To calculate the GPA, student scores were transformed from numeric grades to grade points ranging from 0 to 4 using the scale in Table 4, which was also the standard scale for calculating GPA in TCCAM. The crude GPA form of the FY was deemed to be unsuitable for this study as it included the average results of all of the courses taken in a specific semester. This study investigated only the English language medium courses that were related in content to the students' academic specializations. Therefore, only the grade points of the academic courses that were taught in English and related in

content to the students' academic study were included in the GPA used to represent academic achievement.

Table 4. Conversion Table for Scores Used in TCCAM*

Numeric grade	<50	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-100
Grade point	0	1.0	1.3	1.7	2.0	2.3	2.7	3.0	3.3	3.7	4.0
Letter grade	F	D	D+	C-	C	C+	B-	B	B+	A-	A

*. from the Registration Office at TCCAM, , 2017.

One problem was faced here was that the students' scores in the academic courses were only available in a grade point system, while their scores in the FP assessment were available in a numeric system. To solve it having the grades in two different forms, scores in the FP were converted to grade points using the scale used in TCCAM as shown in Table 4. For example, if a student's score in FP is ranging between 80 and 84, then this score will be converted to a grade point of 3.0.

•Data Analysis

Statistical Analyses Used with the Student Scores

This research is a correlational study of the predictive validity of English language assessment in the Foundation Program (FP). It investigates the correlation between students' English language proficiency in the FP and their academic achievement in the First Year (FY). It also attempts to investigate whether the strength of the correlation was affected by the different groups of students. Two types of statistical analyses were used: the correlational analysis using Spearman's rho and the difference in means analysis using Mann-Whitney U test and the Kruskal Wallis Test. These tests were used to identify significant differences between student scores in different groups when the predictive validity varied amongst the groups. The distribution of the scores was negatively skewed and the sizes of the group samples were not equal.

4.Results

FP Assessment Predictive Validity

Students' grades in the Foundation Program (FP) English language courses

and their average grades in the First Year (FY) academic courses were tested for normality of distribution using Kolmogorov-Smirnov, Shapiro-Wilk tests. These results revealed that the students' scores were all negatively skewed as shown in Table 5 below.

Table 5. Descriptive Statistics and Skewness of Student Scores in FP and FY Academic Courses assessment

Courses		N	Minimum	Maximum	Mean	Std. Deviation	Skewness
FP Assessment	**AES** assessment	153	1.70	4.00	3.22	.49	-.36
	GES Assessment	153	.00	3.70	2.23	.61	-.72
	(**AES** + **GES**)	153	1.0	4.0	2.77	.47	-.62
FY Academic Courses Assessment		154	.50	3.90	2.71	.66	-1.08

First Year (FY) academic courses were all taught in English language and were taken as core courses in the Meteorology, communication , air traffic control and Air-craft maintenance academic programs.

As table 6 shows above, the results showed a high significance, but weak a correlation between the two variables, rho=0.31, $p < 0.01$ Also, the difference in the predictive validity of each of the FP courses (i.e., GES and AES) was investigated. The students' grades in the GES assessment had a weak correlation with their average grades in the academic courses, rho =0.37, $p < 0.01$. Also, the correlation between the students' grades in the AES assessment and in the academic courses assessment was weaker, rho=0.27, $p < 0.01$. In other words, the students' grades in the FP assessment were clearly a weak predictor of their grades in the academic courses.

Table 6. Correlations between Students Grades in Academic Courses, Foundation Program assessment, General English Skills Test and Academic English Skills Assessment

Courses	Academic Courses (N =153)	FP (GES +AES) (N =153)	GES (N =153)	AES (N =153)
Academic courses	1.000	.312**	.369**	.271**
FP	.312**	1.000	.806**	.826**
GES	.365**	.807**	1.000	.476**
AES	.271**	.826**	.476**	1.000

Significant correlation at the 0.01 level (2-tailed).

Student scores in AES assessment were much higher than their scores in GES assessment as Table 5 shows; Also, the predictive strength of GES assessment is higher than AES assessment. The only explanation for the higher correlation between student scores in GES assessment and Academic Courses assessment is the kind of assessment instrument used (i.e., tests). Therefore, it is worth saying that at this point the GES assessment like FY assessment contained standardized tests while the AES assessment included performance assessment tasks.

•Differences between Gender Groups

The correlations between the students' scores in the FP assessment and their grades in the FY academic courses assessment did not show a significant difference between the gender groups. So the Spearman coefficient for the male group was rho = 0.31 and for the female group was rho = 0.33.

Table 7. Correlation between Scores in FP and FY assessment by Gender

Gender	Correlation	Sig.	N=153
Male students	.31*	.07	51
Female students	.33**	.000	102

*. significant correlation at the 0.05 level (2-tailed).

**. significant correlation at the 0.01 level (2-tailed).

Foundation Program (FP), First Year (FY)

Differences among Self-evaluation Groups

The students were required to self-evaluate their language proficiency using the descriptors: weak, average, good, very good and excellent. The Spearman correlation between students grades in the FP assessment and their grades in FY academic courses assessment ranged from rho = 0.17 for the average group to rho = 0.88 for the excellent Group as shown in Table 8.

This indicates that the higher the students self-evaluated their language proficiency, the stronger the predictive validity coefficient of FP assessment was, and that the more their performance in the academic courses assessment was predictable by their performance in the FP assessment.

Table 8. Correlations between Scores in the FP and FY Assessment by to self-evaluation groups

Self-Evaluation	Correlation	Sig.	N = 163
Average	.17	.59	15
Good	.25*	.02	85
V. Good	.39**	.005	51
Excellent	.88**	.009	12

*. significant correlation at the 0.01 level (2-tailed).

*. significant correlation at the 0.05 level (2-tailed).

Foundation Program (FP), First Year (FY)

Differences among Specialization Groups

The strength of the predictive validity of the FP assessment varied depending on the students' specializations. Table 9 shows that the students' grades in meteorology and air traffic control courses were less well predicted by their grades in the FP assessment than were their grades in Communication and Air-craft maintenance courses.

The predictive validity of FP assessment in the specialization groups varied considerably from rho = 0.18, p = 0.12 for the air traffic control group to rho = 0.64, p = 0.002 for the Communication group.

Table 9. Correlations between Scores in the FP and FY Assessment by Specializations

Specialization	Correlation	Sig.	N = 153
Meteorology	.41*	.008	31
Communication	.64**	.002	21
Air Traffic Control	.18	.12	78
Air-craft maintenance	.57**	.005	23

*. significant correlation at the 0.05 level (2-tailed).

**. significant correlation at the 0.01 level (2-tailed).

Foundation Program (FP), First Year (FY)

The difference in the predictive validity for the four different groups of students at (TCCAM) could be explained by the type of specializations taught in each of the Departments and the size of student samples represented by each specialization in this study as shown in Table 9. The participants from the College were specialized in meteorology, communication and air traffic control and air craft maintenance.

It is true that most of the College participants were communication students (66.93% of the sample), and that the predictive validity of FP assessment for the communication group was non-significant, could very well explain the non-significant result obtained for the predictive validity of the FP assessment in the college.

Table 10. The FP assessment Predictive Validity by Specialization

	Specialization	Correlation	Sig.	n
Technical college of Aviation & Meteorology (TCCAM)	Meteorology	27	.27	18
	Communication	.73**	.000	21
	Air-traffic control	.11	.31	78
	Air-craft maintenance	.66**	.001	23

The findings of this study suggested that the predictive validity of FP assessment is weak. Also, the strength of the predictive validity varied according to the student specializations and self-evaluations; the predictive validity of FP assessment was found to be stronger for Communication students and the students who evaluated their language skills as higher.

5.Discussion

• Predictive Validity of FP

The findings of the predictive validity of the FP English language assessment demonstrated a significant but weak correlation between the students' grades in the FP English language assessment and their FY grade in academic courses. Also, Students' grades in GES assessment showed slightly a stronger correlation coefficient with their grades in the academic courses assessment than did their grades in the AES assessment. This finding suggests that language proficiency is not predictor of students' academic achievement.

This finding is in the line with the similar research conducted on the predictive validity of various English language tests that are used as gatekeepers to higher education institutions such as IELTS, TEAM, and various local tests (Davies, 1990; Elder, 1993; Cope, 2011; Lynch, 2000).

This finding indicates that the predictive validity of FP assessment accounts only for about 16% of the variance of students' performance in academic courses assessment. Also, this finding raises some questions about the policies on accepting students with different language proficiency levels in higher education institutions not only in Libya but also in other international institutions. The difference in the strength of the predictive validity of GES and AES also raises some questions about the reliability of performance assessment and consistency in using marking scales.

• Predictive Validity of FP across Specializations

The finding of this study showed that the strength of the correlation between the students' language proficiency and academic achievement varied considerably depending on the students' specializations. These different predictive validity values for the specializations could be partly explained by the Communication assessment instruments and test tasks seemed to focus on students' language skills more than did those of the Meteorology or Air traffic control assessment instruments. In Communication, students are required to write a 1000 word reports, write essays in the final exam and conduct presentations, all of which require a certain level of English language mastery that is less required by the assessment tasks in other specialization.

This finding is similar to the findings of other research studies (e.g.,

Lynch, 2000; Huong, 2001) which constitutes a model suggesting that there is a variation in language skills requirements of academic disciplines not only in the Libyan higher education, but also in many other international higher education institutions. Despite the fact that many international higher education institutions require different levels of language proficiency for different academic disciplines, these requirements are usually not based on predictive validity studies. Therefore, the TCCAM should increase the entry level of English language requirements for the students who are willing to study, specifically, Communication Studies.

- **Predictive Validity of FP across Self-Evaluation Groups**

The correlations between language proficiency and academic achievement showed a variance according to the students' self-evaluations of their language proficiency levels. The higher the students evaluated themselves, the stronger the correlation between their grades in FP assessment and academic courses assessment was. However, few research studies investigated the impact of self-evaluation on academic achievement/ difficulties, but in a study by (Xu, 1991) reported self-evaluation as a good predictor of academic difficulties. This study suggests that more emphasis should be devoted to study the role of self-evaluation in predictive validity in future research. Also, self-evaluations can be used in higher education institutions as an investigating tool to probe more into students' academic achievement and/or difficulties.

6.Conclusion

This correlational study explored the predictive validity of the Foundation Program assessment by correlating students' scores in its assessment with their scores in the First Year academic courses. The findings showed that language proficiency in English is a moderate predictor of academic achievement in general. However the power of the predictive validity was found to differ with regards to students' self-evaluations and specializations, but not according to their gender. The higher the students evaluated their language proficiency, the higher the FP assessment predictive validity was. The predictive validity of FP assessment was strong for the Communication and Air craft maintenance groups, moderate for the meteorology group and non-significant for the air traffic control group. The findings of this study revealed moderate to low predictive validity of English language assessment according to academic

achievement, but students' proficiency in the English language plays a major role in accessing Libyan higher education.

Therefore, this study recommended that in admission to a higher education institution, English language proficiency should be viewed as a criterion along with students' academic achievement, but used fairly variously. however, higher education institutions that employ English as a medium of instruction request a specific level of language proficiency in high school English language courses that is equivalent to that language level required in academic courses.

Moreover, the AES assessment revealed a lower value of its predictive validity than did the GES tests. This finding should be applied in cases where students' scores are very close to the cut-off point (50 out of 100). However, the present procedure taken currently is that if a students' score is between 47 and 49, it is then added up to 50, which is the passing score. English language assessment plays an important role in higher education and its impact is evident in higher education admission policies. Therefore, this study suggests that these policies should be re-evaluated and interpretations made of student scores in English language assessment should be examined carefully with reference to the findings of the predictive validity of FP assessment.

• References

Al-Musawi, N. M., & Al-Ansari, S. H. (1999). Test of English as a Foreign Language and First Certificate of English Tests as Predictors of academic success for undergraduate students at the University of Bahrain. System, 27, 389-399.

Bachman, L. F. (2004). Statistical analyses for language assessment. Cambridge: Cambridge University Press.

Bachman, L., & Palmer, A. (1996). Language testing in practice: Designing and developing useful language tests. Oxford: Oxford University Press.

Breeze , R., & Miller, P. (2008). Predictive validity of the IELTS listening test as an indicator of student coping ability in Spain. IELTS Research Reports. 12, 1-34.

Cho, Y., & Bridgeman, B. (2012). Relationship of TOEFL iBT scores to academic performance: Some evidence from American Universities. Language Testing, 29(3), 421-442.

Cope, N. (2011). Evaluating locally-developed language testing. Australian Review of Applied

Linguistics, 34(1), 40-58.

Cotton, F., & Cornrow, F. (1998). An investigation of the predictive validity of IELTS amongst a group of international students studying at the University of Tasmania. IELTS Research Reports, 1, 72-115.

Davies, A. (1990). Principles of Language Testing. Oxford: Blackwell.

Elder, C. (1993). Language proficiency as a predictor of performance in teacher education. Melbourne Papers in Language Testing, 2(1), 68-89.

Feast, V. (2002). The impact of IELTS scores on performance at University. International Education Journal, 3(4), 70-85.

Harrison, A. (1983). A language Testing Handbook. London: McMillan.

Hughes, A. (2003).Testing English for Language Teachers. Cambridge. UK: Cambridge University Press.

Huong, T. (2001). The predictive validity of the international English language test system (IELTS). The Post Graduate Journal of Education Research, 2(1), 66-96.

Jochems, W., Sinppe, J., Jan Smid, H., & Verweij, A. (1996). The academic progress of foreign students: Study achievement and study behaviour. Higher Education, 31(3), 325-340.

Kerstjen, M., & Nery, C. (2000). Predictive validity in the IELTS test: A study of the relationship between IELTS scores and students' subsequent academic performance. IELTS Research Reports, 86-108.

Lynch, T. (2000). An evaluation of the revised test of English at matriculation at the University of Edinburgh. Edinburgh Working Papers in Applied Linguistics, 10, 61-71.

Maleki, A., & Zanjani, E. (2007). A survey on the relationship between English Language proficiency and the academic achievement of Iranian EFL students. Asian EFL Journal, 9(1), 86-96.

Martuza, V. R. (1977). Applying Norm-Referenced and Criterion-Referenced Measurement in Education. Boston: Allyan and Bacon.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. Educational Researcher, 18(5), 5-11.

Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Assessment. Educational Measurement. 14 (4). 5-8.

Powers , D. E., Kim, H. J., & Weng, V. Z. (2008). The redesigned TOEIC (listening and

reading) test: Relations to test-taker perceptions of proficiency in English. ETS Research Report.

Ross, S. J. (2008). Language testing in Asia: Evolution, innovation, and policy challenges. Language Testing, 25(1), 5-13.

Vinke, A., & Jochems, w. (1993). English proficiency and academic success in international postgraduate education. Higher Education, 26, 275-285.

Woodrow, L. (2006). Academic success of international postgraduate education students and the role of English proficiency. University of Sydney Papers in TESOL, 1, 51-70.

Xu, M. (1991). The impact of English language proficiency on international graduate students' perceived academic difficulty. Research in Higher Education, 32(5), 557-569.

Yen, D. A., & Kuzma, J. (2009). Higher IELTS score, higher academic performance? The validity of IELTS in predicting the academic performance of Chinese students. Worcester Journal of Learning and Teaching, 1-7.

Zabihi, R. (2011). Personality in English proficiency and achievement. Continental Journal of Education Research, 4(1), pp.1-6