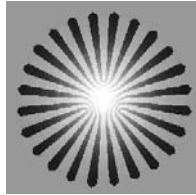


In The Name Of God



**Payame Noor University
Faculty of Humanities**

Department of Linguistics and Foreign Languages

***The Study of Predictive Validity of PNU MA Entrance Exam
vs.***

Sanjesh MA Entrance Exam for PNU TEFL

Submitted in Partial Fulfillment of the Requirements for the Degree of MA in Teaching
English

Supervisor

Dr. F. Vahdany

Advisor

Dr. F. Hemmati

By

Flora Khadivi Seresht

Payame Noor University

September, 2011



شماره
تاریخ
پیوست

تصویب نامه

پایان نامه کارشناسی ارشد رشته آموزش زبان انگلیسی

به شماره پروژه ۷۰۱۸ تحت عنوان:

"مطالعه روایی پیش بینی امتحان فراگیر پیام نور در مقایسه با
روایی پیش بینی امتحان کنکور در مقطع ارشد در رشته
آموزش زبان انگلیسی"

نام و نام خانوادگی: فلورا خدیوی سرشت شماره دانشجویی: ۸۵۷۱۰۳۱۵۰

ساعت: ۱۰-۰۸/۳۰

تاریخ دفاع: ۹۰/۰۶/۲۶

درجه ارزشیابی: عالی

نمره: ۱۹ نوزده تمام

هیات داوران:

امضاء	مرتبه علمی	نام و نام خانوادگی	داوران
	استاد	جناب آقای دکتر فریدون وحدانی	استاد راهنما
	استاد	سرکار خانم دکتر فاطمه همتی	استاد مشاور
	استاد	جناب آقای دکتر متوجه جعفری گهر	استاد داور

Dedicated to my dear parents, brother, and my dear little nephew, who has been learning English since he was four and half, and helped me with the number of lexical word I needed.



Acknowledgement

*I would like to thank **Dr. Freidoun Vahdany**, whose continued assistance and gracious comments was always available throughout this thesis. I would like to expand my sincere thanks to **Dr. Fatemeh Hemmati** whose help and expertise I had at every stage of this dissertation. Her honest help, patience and admirable characteristics, has greatly added to my educational experience. It was their quick brilliant feedback, and encouragement, which shed light to accomplish this work.*

*I also want to express my gratitude to **Tehran PNU TEFL Department**, for having permitted me to conduct this research particularly, **Dr. Iravani** and **Dr. Raee Sharif**.*

*In the process of data collection of the quantitative phase, I am very thankful to the help and support of **PNU Central Organization, Tehran Center of PNU, Faculty of Humanities, and PNU Center of Research and Technology**.*

*Furthermore, this thesis is highly grateful to the honest and sincere corporation of **Sanjesh Organization of Tehran and Karaj** for helping to accomplish the data collection process.*

*Concerning the tremendous part of statistics, I really appreciate Guilan PNU professor, **Mrs. Omolbanin Moosazadeh** for having performed it patiently and carefully.*

*With respect to the item analysis procedure in the qualitative phase, I greatly thank **Dr. Niloofar Mansouri**, Guilan PNU Professor, who earnestly and truthfully accepted to render judgments on the part of linguistic and phonetic items; **Mrs. Moosazadeh** for having kindly evaluated the research items of PNU exam, and **Dr. Abdorreza Tahriri**, Guilan University Assistant Professor, who gave valuable pieces of advice and made an honest final check on all parts of PNU and Sanjesh MA exams items.*

*I would also like to express my gratitude to **the number of professor** whom I got the chance to interview and to furnish the thesis with their appreciating perspectives. It is also the place to thank all good **Faragir students** who eagerly joined the interview and let me have their view points over PNU MA entrance exams as well as academic exams.*

Abstract

This empirical validity study, aimed at examining the predictive validity of PNU Postgraduate Entrance Exam, compared to Sanjesh Postgraduate Entrance Exam on Faragir and Non-Faragir students' academic success in PNU TEFL, comprises two parts: quantitative and qualitative phases.

The subjects, fallen in 10 student-group cohorts of Faragir and Non-Faragir populations, 5 groups each; involved the final number of 141 MA students encompassing 86 Faragir and 55 Non-Faragir students. They were selected in random from among the accessible TEFL students, entered from 1384 to 1388. It required 20 series of data for the 10 groups. MA scores of Faragir, and MA ranks of Non-Faragir students were correlated with their academic grade point average (GPA) of the selected courses of Sem.1 and Sem.2, set as the criterion in each group. Pearson and Spearman's rho correlations were applied to MA scores and MA ranks respectively, having calculated the correlations, first for each academic year in distinct, and then for the entire population in each group. The obtained correlation for the total PNU MA data was significant at the 0.01 level (2-tailed), indicating the predictability of PNU MA data on academic success, in contrast to the inability of rank data due to the poor value of ranks in making academic judgments. The outcome reports, Faragir 85 > Faragir 84 > Faragir 88 > Faragir 87 > Faragir 86, with a significant correlation of 0.01 for Faragir group of 1385. It also involved a T-Test, identifying Faragir students as outperformers with a mean difference of $p. (0.000) < 0.05$.

In the qualitative phase, apart from an interview about PNU MA exams and academic exams; three series of PNU and Sanjesh MA exam items of the randomly selected years of 1384, 1385, and 1387 were examined. The items were detected in terms of content validity regarding their relevance, meaningfulness, and utility; as well as test construction rules, each with respect to its own system. The two systems having exhibited to have taken the same line of work, demonstrated their test items to be content-valid; however, non-standardization was quite observed in both systems with lesser degree towards Sanjesh items. The highest level of difficulty has been identified by PNU MA exam of 1385, in which some items tested two or even three points at a time, stated as a reason for the perceived correlation of Faragir group in this year. This finding can highlight the role of criterion setting on making academic prediction, and validity of the exam.

Key words: Predictive validity, MA Entrance Exam, PNU, Sanjesh, Faragir students, Non-Faragir students

Table of contents

Page

Approval-----	III
Dedication-----	IV
Acknowledgements-----	V
Abstract-----	VI
List of tables-----	XI
Abbreviations-----	XIII

Contents

Chapter one: Introduction-----1

1.1. Overview.....	2
1.2. Statement of the problem.....	4
1.3. Research questions.....	5
1.4. Research hypotheses.....	6
1.5. Significance of the study.....	6
1.6. Purpose of the study.....	7
1.7. Definition of terms.....	8

Chapter two: Review of the Related Literature-----9

2.1. Overview.....	10
2.2. Testing and test evaluation.....	12
2.3. Test validity.....	13

2.3.1. Validity related literature.....	14
2.4. Validity in types.....	16
2.4.1. Content validity.....	17
2.4.2. Face validity.....	18
2.4.3. Construct validity.....	19
2.4.3.1. Convergent/Divergent validation.....	19
2.4.3.2. Factor analysis.....	20
2.4.3.3. Other methods of assessing construct.....	20
2.4.4. Criterion-related validity.....	21
2.4.4.1. Concurrent validity.....	21
2.4.4.2. Predictive validity.....	22
2.4.4.2.1. Predictive validity in Iran.....	23
2.4.4.2.2. Predictive validity in some areas of language tests.....	25
2.4.4.2.3. Predictive validity and IELTS.....	28
2.4.4.2.4. Predictive validity and SAT.....	36
2.4.4.2.5. Predictive validity in other disciplines.....	43
2.5. Item analysis.....	47
2.5.1. Criterion-referenced item analysis.....	48
2.5.2. Norm-referenced item analysis.....	48

Chapter three: Method-----49

3.1. Overview.....	50
3.2. Participants.....	51
3.3. Instrumentation.....	53
3.4. Design.....	54
3.4.1. The quantitative phase.....	54
3.4.2. The qualitative phase.....	55
3.5. Procedure.....	56
3.6. Data analysis.....	57
3.6.1. Quantitative data analysis.....	57

3.6.2. Qualitative data analysis.....	58
3.7. Abbreviations.....	60
Chapter four: Results and Discussions-----	61
(4.1) Quantitative.....	62
4-1.1. Introduction.....	63
4-1.2. Research questions.....	64
4-1.3. Results.....	64
4-1.3.1. The quantitative data analysis results.....	64
4-1.3.1.1. Investigating H.01.....	65
4-1.3.1.2. Investigating H.02.....	72
4-1.3.1.3. Investigating H.03.....	77
4-1.4. Discussion.....	87
(4.2) Qualitative.....	91
4.2.1. Introduction.....	92
4.2.2. Research question.....	93
4.2.3. The qualitative data analysis results.....	93
4.2.3.1. Investigating H.04.....	93
1. Item analyses of the predictive validity of PMAE and SMAE.....	93
1.1. Item analyses of the predictive validity of PMAE.....	101
1.1.1. Item analysis of PMAE 1384.....	102
1.1.2. Item analysis of PMAE 1385.....	118
1.1.3. Item analysis of PMAE 1387.....	128
1.2. Item analyses of the predictive validity of SMAE.....	140
1.2.1. Item analysis of SMAE 1384.....	142
1.2.2. Item analysis of SMAE 1385.....	153
1.2.3. Item analysis of SMAE 1387.....	161
4.2.3.2. Interview.....	171
4.2.4. Discussion.....	175

4.2.5. Definition of terms.....	179
---------------------------------	-----

Chapter five: Conclusions, Implications, and Suggestions for

Further Study-----180

5.1. Introduction.....	181
5. 2. An overview of the findings.....	181
5.3. Conclusion.....	183
5.4. Implications of the study.....	190
5.5. Suggestions for further researches.....	195
5.6. Limitations of the study.....	197

References-----199

Appendices-----

<i>Appendix A1.</i> (PNU MA Entrance Exam items, 1384)	205
<i>Appendix A2.</i> (PNU MA Entrance Exam items, 1385)	223
<i>Appendix A3.</i> (PNU MA Entrance Exam items, 1387)	236
<i>Appendix A4.</i> (PNU MA Entrance Exam items, 1386)	248
<i>Appendix B1.</i> (Sanjesh MA Entrance Exam items, 1384)	257
<i>Appendix B2.</i> (Sanjesh MA Entrance Exam items, 1385)	265
<i>Appendix B3.</i> (Sanjesh MA Entrance Exam items, 1387)	274

Persian Abstract-----281

List of tables**Page**

Table 2.1. Messick's (1988/ 1989) theory of validity.....	16
Table 4.1.1. Correlations for Faragir 84.....	66
Table 4.1.2. Descriptive Statistics for Faragir 84.....	66
Table 4.1.3. Correlations for Faragir 85.....	67
Table 4.1.4. Descriptive Statistics for Faragir 85.....	67
Table 4.1.5. Correlations for Faragir 86.....	68
Table 4.1.6. Descriptive Statistics for Faragir 86.....	68
Table 4.1.7. Correlations for Faragir 87.....	68
Table 4.1.8. Descriptive Statistics for Faragir 87.....	69
Table 4.1.9. Correlations for Faragir 88.....	69
Table 4.1.10. Descriptive Statistics for Faragir 88.....	69
Table 4.1.11. Descriptive Statistics (for the total Faragir subjects).....	70
Table 4.1.12. Correlations (for the total Faragir subjects).....	71
Table 4.2.1. Correlations for Non Fragir 84.....	73
Table 4.2.2. Correlations for Non Fragir 85.....	73
Table 4.2.3. Correlations for Non Fragir 86.....	74
Table 4.2.4. Correlations for Non Fragir 87.....	74
Table 4.2.5. Correlations for Non Fragir 88.....	75
Table 4.2.6. Correlations (for the total non Faragir subjects).....	76
Table 4.3.1. Group Statistics for Fragir and non Fragir subjects in 1384.....	78
Table 4.3.2. Independent Samples Test for Fragir and non Fragir subjects in 1384.....	79

Table 4.3.3. Group Statistics for Faragir and non Faragir subjects in 1385.....	80
Table 4.3.4. Independent Samples Test for Faragir and non Faragir subjects in 1385.....	80
Table 4.3.5. Group Statistics for Fragir and non Fragir subjects in 1386.....	81
Table 4.3.6. Independent Samples Test for Fragir and non Fragir subjects in 1386.....	81
Table 4.3.7. Group Statistics for Faragir and non Faragir subjects in 1387.....	82
Table 4.3.8. Independent Samples Test for Faragir and non Faragir subjects in 1387.....	82
Table 4.3.9. Group Statistics for Fragir and non Fragir subjects in 1388.....	83
Table 4.3.10. Independent Samples Test for Fragir and non Fragir subjects in 1388.....	83
Table 4.3.11. Group Statistics for the two groups.....	84
Table 4.3.12. Independent Samples Test for the two groups.....	85

Abbreviations

PNU: Payame Noor University

PMAE: PNU MA Entrance Exam

SMAE: Sanjesh MA Entrance Exam

GPA: Grade Point Average

Chapter One

Introduction

Chapter One

Introduction

1.1. Overview

Higher education, which meets the needs of millions of educators over years, is experiencing a growing demand by passage of time. Due to the expected surge and demand, university education either on-site or remote is likely to face overcrowding. Hence, the composition of our universities calls for attention in order to firmly decide what kind of students they want, to test their potential success properly, and to provide them with exactly the entrance exam resources which lead them to succeed.

The practice of language testing drawing upon and contributing to all disciplines within applied linguistics, has shown to be different in that, it is all about researching and creating better tests. According to Fulcher and Davidson (2007), sociolinguists do not create "sociolinguistic things". Discourse analysts do not create discourses. Phonologists do not create spoken utterances. Language testing, in contrast, is about *doing*. It is about *creating tests* in which art is reflected. According to Fulcher and Davidson (2007), the movement set within a new approach to testing brings testing practice, theory, ethics, and philosophy together. Focal to this new approach, is the concept of effect-driven testing which is a highly

pragmatic view of test validity with emphasis on the outcome of testing activities. This type of validity concerns the impact of tests on individuals and societies exerted by test results, being pleasant/unpleasant, acceptable/unacceptable.

To guarantee the students' success in the accepted field, universities have to decide what criterion in each field of study is needed. The issue of criterion-setting is a worthy and interesting debate in the area of testing and to walk away from this challenge means to spoil the students' future. It is regretful to say; sometimes the overriding population in a field of university education turns out to be highly different from what is supposed to be. Specialized education plays a vital role throughout lives, and the individual candidate finds achieving this target the sole guarantee of his future life. To pursue education, he devotes himself to gain access to any form of it even the online education, regardless of age, color, ethnicity, etc.

Amongst the modes of university education there exist mega universities spread all over the world, offering long-distance education. One such university is Payame Noor University (PNU) which its predictive validity of MA Entrance Exam in the field of English Language Teaching is of the main concern in this study. Students in this mega university, fall within the two categories of Fragir and Non-Fragir cohorts.

The mode of education offered in PNU, encountered praise as well as criticism. It is praised for raising independent and self-contained students who feel responsible toward their own learning. Students in PNU context, though faced with restricted period of instruction, have proven to be more self-reliant compared to other university students. The opposing view clarifies increasing dissatisfaction resulted by the mode and the length of

education, when the materials appear to be so heavy or a two/four year degree takes too long that the educator seems no longer interested in it. Based on the research, the point is mostly observed by the incorrectly accepted number of applicants considered as false-positives by the present study, who underperform in the university. The study considers, the point can signal presence of deficiencies in PNU Postgraduate Entrance Exam and the shortcomings of a quantitative approach of the admission system.

As with Fragir students, the case is also observed across Non-Fragir students who are admitted via the system of Sanjesh MA Exam. Hence; this study pays a look at the entrance exams of the two systems to determine the extent to which they can predict the future performance of the incoming populations.

1.2. Statement of the problem

Tests and testing, as well as the quality and appropriateness of examinations remain an issue of intense debate. Testing becomes more challenging when it comes to be presented to the university education particularly, in case of specialized education. According to Eric Gustafsson (n.d.), selection for higher education is necessary when the number of applicants is high. He says that the selection instruments must be objective, reliable, valid and fair enough to avoid negative consequences for certain groups of applicants. Selection should offer objective and precise tests which can enable to assign reliable and valid grade scores. His paper mentions that grade scores tend to have better predictive validity than tests. In some situations grades change lives and a student should live with a grade throughout his/her life, while a test may be repeated overtime.

Therefore, tests need reflection and revision to yield in reliable and fruitful results. Tests should allow the testers to render meaningful judgments from test scores about the testees' future performance and the likelihood of his/her success. The impact a test exerts upon individuals can give us an image about the effect-driven testing issue stated in Fultcher (2007), that how a good test taker may have to come along with a poor grade score an entire life and conversely, how a chance score is granted to a poorly-prepared examinee after a life-altering exam due to the lack of comparability of the test. The study thinks this pragmatic framework gives way to ponder on the influential phenomenon of testing and based on the encapsulation model of achievement grade scores in Eric's paper; make the best predictions obtained by grade scores in combination with measures of domain-specific knowledge. Reliable predictions made from test scores support the validity of tests. Back to Eric Gustafsson (n.d.), some scores are corrupted to reflect something else than knowledge and skills. They may exhibit to be suffering from lack of comparability, reliability, and validity.

This research is going to unravel the degree of measurement on the basis of which PNU MA Entrance Exam can predict Fragir students' possibility of arriving at the academic success criterion in TEFL, with a look at that of Sanjesh MA Entrance Exam power on the academic success of Non-Fragir students, studying in PNU TEFL.

1.3. Research questions

This study, launched to estimate the predictive ability of PNU Postgraduate Entrance Exam and its Sanjesh counterpart on Fragir and Non-Fragir students' academic success, raises the questions below:

- 1) Is there any significant relationship between Fragir students' PNU MA exam results and their academic average in TEFL?
- 2) Is there any significant relationship between Non-fragir students' Sanjesh MA exam results and their academic average in PNU TEFL?
- 3) Compared, is there any significant difference found between the two groups, with respect to their academic performance in PNU TEFL?
- 4) Is there any difference between the test items in the two systems in terms of their quality?

1.4. Research hypotheses

- H.01: There is no significant relationship between Fragir students' PNU MA exam results and their academic average in TEFL.
- H.02: There is no significant relationship between Non-fragir students' Sanjesh MA exam results and their academic average in PNU TEFL.
- H.03: Compared, no significant difference is found between the two groups with respect to their academic performance in PNU TEFL.
- H.04: There is no difference between the test items in the two systems in terms of their quality.

1.5. Significance of the study

This empirical validity study signifies the part played by the criterion-referenced testing in demonstrating the possible relationship between the predictor, in this case PNU MA exam results and the predicted component i.e. the extent to which the academic goal is achieved.

This achievement signifies the predictive power of PNU MA Entrance Exam which is being

investigated for the first time by this study on PNU postgraduate students in TEFL. It intends to probe the roots of the deficiencies observed in PNU MA admission system in TEFL, with a relative emphasis on Sanjesh MA system. The research wants to propose that in the applied and linguistic demanding discipline of Teaching English as a Foreign Language (TEFL), the entrance exam which by itself needs revision and standardization is only one of the factors, which contribute to the selection of applicants particularly, when it comes to a high number of candidates in specialized education.

The study wants to explore the drawbacks of the admission scores and to convey; mere reliance on such a quantitative criterion should not satisfy PNU TEFL admission system in picking up the most qualified and competent candidates. There can be involved a process, in which value is given to the role of some intervening factors to elicit rich knowledge structures within particular domains which may be also selective and predictive of the students' academic success. In this respect, we can refer to what Dooley (1999), in her paper says that much of the literature in the area of the predictive validity of language tests suggests "English proficiency is one among many other factors that affect academic success" (Graham 1987).

1.6. Purpose of the study

The rationale behind the study is to **(a)** evaluate the predictive validity of PNU MA Entrance Exam in TEFL, since it has been offered i.e. 1384 onward, **(b)** assess the predictive ability of Sanjesh MA Entrance Exam administered in the same years, **(c)** make a comparison between the predictive validity of PNU and Sanjesh systems, by looking at the high achievers in each group, **(d)** find the qualitative differences between the test items in the two systems, **(e)** offer