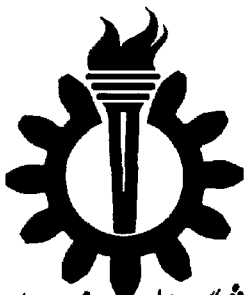


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

۱۳۸۲ / ۵ / ۲۷

مرکز اطلاعات مدرک علمی ایران
تهیه مدارک



دانشگاه علم و صنعت ایران

دانشکده زبانهای خارجی

پیش فرض آزمون دهندگان و برداشت آزمون شوندهگان
از واحدهای آزمون خواندن

فاطمه شاه محمدی

جهت اخذ درجه کارشناسی ارشد آموزش زبان انگلیسی

استاد راهنما: آقای دکتر فرهادی

استاد مشاور: آقای دکتر محمود طباطبائی

بهمن ۱۳۸۱

۴۷۶۳

تقدیم به دخترم،

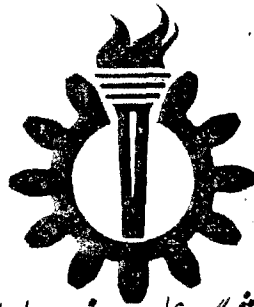
همسرم و

استاد محترم جناب دکتر فرهادی

چکیده

هدف این تحقیق بررسی واحدهای آزمون خواندن ازدو دیدگاه: آزمون دهنده گان و آزمون شوندهگان بوده است؛ در واقع شباهتها و تفاوتهای پیش فرض آزمون دهنده گان و برداشت آزمون شوندهگان از واحدهای آزمون بین المللی خواندن (IELTS) مورد مطالعه قرار گرفته است.

هفده استاد دانشگاه از سراسر جهان و شصت دانشجوی کارشناسی و کارشناسی ارشد رشته زبان انگلیسی از دانشگاههای تهران و مازندران در این پژوهش شرکت داشتند. به منظور انجام تحقیق چک لیستی از مهارتهای خواندن با استفاده از چک لیست های موجود مهارتهای خواندن تهیه و به همراه آزمون خواندن (IELTS) به آزمودنی ها داده شد. چندین آزمون مجذور کای به منظور تحلیل داده ها اعمال شده است. نتایج حاصله نشان می دهد اگر چه میان آزمون شوندهگان و آزمون دهندهگان و همچنین در بین خود آزمون شوندهگان در رابطه با مهارت مورد ارزیابی واحدهای آزمون خواندن (IELTS) ارتباط وجود دارد این ارتباط معنی دار نیست؛ در حالیکه بین آزمون دهنده گان توافق معنی داری وجود دارد. قابل ذکر است که این توافق در رابطه با واحدهایی است که مهارت (SCANNING) را می سنجد. از نتایج حاصله از این تحقیق می توان در تهیه و تدوین مواد آموزشی زبان انگلیسی و در تدریس خواندن و آزمون آن استفاد کرد همچنین این تحقیق توصیه می کند که از آزمودنیها در اعتبار بخشیدن به آزمون ها کمک بگیریم.



دانشگاه علم و صنعت ایران

IRAN UNIVERSITY OF SCIENCE & TECHNOLOGY
DEPARTMENT OF FOREIGN LANGUAGES

Thesis Submitted to the English Department of Iran University
of Science & Technology (IUST) in Partial Fulfillment
of the Requirements for the Degree of Master of Arts in
English Language Teaching

**Testers' Presumption and Test-takers' Perception of
What Reading Test Items Measure**

By:

Fateema Shah Mohammady

Supervisor: *Dr. Hossein Farhady*
Advisor: *Dr. Mahmood Tabatabaee*

February, 2003.

ترتیب اطلاع است و مدرک علمی ایران
تعمیرت مدرک

۴۷۶۰۳

Iran University of Science & Technology
(IUST)

Department of Foreign Languages

WE HEREBY RECOMMEND THAT THE PRESENT
THESIS

BY

Fateema Shah Mohammady

ENTITLED

*Testers' Presumption and Test-takers'
Perception of What Reading Test Items Measure*

BE ACCEPTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS(MA) IN ENGLISH LANGUAGE
TEACHING(ELT)

COMMITTEE OF ORAL PRESENTATION:

.....DR. HOSSEIN FARHADY, PH. D.
(SUPERVISOR)

.....DR. MAHMOOD TABATABAEE,
PH. D. (READER)

.....DR. MAHMOOD TABATABAEE,
PH. D. (HEAD OF DEPARTMENT)

TEHRAN, IRAN, 1381(2003)

۱۳۸۲ / ۱۰ / ۲۷

روزنامه دانش‌آموزان
فصلنامه علمی
موسسه تخصصی

IN THE NAME OF GOD,

THE COMPASSIONATE, THE MERCIFUL

Dedicated to:

My dear daughter, husband and to my great professor Dr. Farhady

Abstract

This study was an attempt to investigate an internationally implemented standard reading test, International English Language Testing System (IELTS) in terms of what it measured from two perspectives: testers' presumption and test-takers' perception of the test items regarding what they measured and that whether there was any agreement within and/or between the two groups. 17 professors specialized in testing from different countries as well as 60 Iranian EFL students at BA and MA levels from Tehran and Mazandaran universities participated in this study.

In order to shed light on the aforementioned issues a reading subskill checklist was prepared based on the existing reading skill checklists and presented to the subjects under study along with the test. Several Chi-square tests were applied. The results indicated that there was not considerable agreement among Iranian EFL students on what the test items measure while a significant agreement could be seen among the testers_ about 2/3 of the testers presumed that IELTS test items measured the test-takers' skill of scanning. Moreover, there was not considerable agreement between the two groups on the skill/s the items measured.

The findings of this study can be applied in material preparations, teaching reading courses, and test development.

Acknowledgements

Although not possible at all, I'd like to express my gratitudes to all those who sincerely assisted and supported me throughout my MA studies.

First of all, my special thanks go to my supervisor, Dr. Farhady, without whose support and insightful comments I could never ever complete this task.

I'd also like to thank Dr. Tabatabaee, my advisor and the head of the English Department, who sincerely guided me during my research study.

Furthermore, I'd like to appreciate Dr. Birjandi and Dr. Vaezee for agreeing to present as the external and internal examiners of my presentation.

I am indebted to all my professors at IUST and at other universities, Dr. Maftoon, Dr. Fallahi, Dr. Bidahri, Dr. Eslami, Dr. Vaezee, Dr. Birjandi, Dr. Miremadi, Dr. Mirhasani.

Last but not the least, I would like to express my wholehearted appreciation to my family, my husband and my daughter who tolerated my lengthy absence from home.

TABLE OF CONTENTS

CHAPTER I

Title	Page
1.1 Introduction.....	1
1.2 Statement of the problem and the purpose of the study.....	3
1.3 Significance and justification of the study.....	4
1.4 Research questions.....	5
1.5 Null hypohese.....	6
1.6 Delimitations of the study.....	7
1.7 Definition of the related term.....	7

CHAPTER II

2.1 Introduction.....	8
2.2 Models of reading comprehension.....	10
2.3 Testing reading skill or language.....	17
2.4 Reading skill or skills.....	18
2.5 The state of the Art in Language Testing.....	30
2.5.1 Historical overview.....	31
2.5.2 Recent development in language testing.....	35
2.5.3 Communicative language testing.....	41
2.5.4 The IELTS test.....	43

CHAPTER III

METHOD

3.1 Participants.....	49
3.2 Instrumentation.....	50
3.3 Procedures.....	50
3.4 The checklist.....	51
3.5 Data analysis.....	53

CHAPTER IV

RESULTS AND DISCUSSION

4.1 Analysis I.....	54
4.2 Analysis II.....	60
4.3 Summary of the results.....	64
4.4 Discussion.....	66

CHAPTER V

CONCLUSION AND IMPLICATION

5.1 Pedagogical implications and applications.....	69
5.2 Suggestions for further research.....	71

REFERENCES

APPENDIXES

CHAPTER I

INTRODUCTION

Tests should be labeled just like dangerous drugs, use with care! (Spolsky, 1981: 20)

Although the above statement might seem to be an extreme position, it indicates that teachers, testers and all those who are concerned with test preparation, test taking processes, and score interpretations are to be cautious and regard the point that what they are dealing with may have important effects on the test takers' lives. Farhady(1999) also asserts "...tests in general, language tests no exception, have turned into extremely powerful weapons in the hands of not only teachers, but also administrators, economists, politicians, and some other authorities in the society"(Farhady, 1999:22).

The significance of testing and tests has resulted in a bulk of research studies in different areas of language testing whose intention was to explore the testing processes and find better ways to prepare valid and reliable tests. One of the areas of research which is getting special attention by researchers is *Judgement* in language testing. Judgement can be traced in different points of testing including content specifications, item writing, scoring, and score interpretation and this is not new. A new area to which judgement seems to have come nowadays is post-judgement. That is, judging a test after the routine objective,

statistical validation procedures which can be done using either expert or nonexpert informants. Several studies including Alderson and Lukmani(1989), Alderson(1990a and 1990b), Lumley(1993), and Hajipournejad(2002) have employed expert or nonexpert *judgement* so as to better understand the reading comprehension processes.

“Alderson(1986, 1990), Alderson and Lukmani(1986), Alderson and Henning, and Lukmani(1987) have been investigating a) the extent to which “experts” agree in their judgement about what specific skills EFL reading test items measure, and at what levels, and b) whether these expert judgements about ability levels are related to the difficulty of items”(Alderson,1990b:674). The studies show contradictory results and employment of post subjective judgement in language testing is an open question which merits further studies.

The present study concerns with both expert and non-expert judgement in testing reading. In fact, the point the study probes is how the consumers of a test perceive or judge a given test in terms of what it tests. “The perception of teachers and students play a large role in the success of a learning event”(Brady, 1998: 102) from which testing cannot be isolated. Johnson(1995) suggests perception “acts as a filter for what the student actually learns[or do]”(mentioned in Brady, 1998: 102). In this research study it is assumed that reading skill, as some reading theorists(Nuttall 1996, Grabe 1997&1991, deLopez, et al. 1997, Lumley 1993, Alderson 1990) suggest, consists of some components/skills which can be developed in people through instruction.

Based on this assumption a reading test should test the *reading ability* of the test-takers. In fact, the present study is an endeavor to investigate the potential correlation between the students' perception and the teachers' presumption of what a reading test measures_ a comparative study between the judgement of two groups of people, testers and test-takers. To probe what the testers might require a test to do with the test-takers in a reading test and that whether this corresponds with the students' own perception of what that test does require them to do.

1.2 Statement of the problem and purpose of the study

The complexity of reading could not prevent researchers from venturing into its territory and its testing. Indeed, testing can be considered a way into reading construct/s. The point we should bear in mind as Alderson(2000) states is we must start from somewhere, if we wait until we have a perfect understanding of the constructs before we begin to devise an assessment instrument, then we will never begin to test reading which is necessary for the real world purposes. "Testing builds upon the assumption that it is possible to discover what ability a test or an item measures"(Alderson 1990, p. 437). Tests items are constructed based on detailed guidelines and specifications and after going through extensive validation processes, they are put into action. But "what matters in reading tests is what an item is testing, rather than what its content specifications claim"(Alderson 1990, p. 430).