۱۳۸۱/۴/۱۰

مرکز اطلاعات و مدارک علمی ایران
تهیه کننده آرک

University for Teacher Education

DEPARTMENT OF FOREIGN LANGUAGES

Construct Validation of Reading Comprehension Tests

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Teaching English as a Foreign Language (TEFL).

By:
Hamid Reza Moeini Asl

Advisor: Dr. Hossein Farhady
Reader: Dr. M.R. Ataei

TEHRAN, IRAN
February 2002

۳۲۶۸۶

UNIVERSITY FOR TEACHER EDUCATION
DEPARTMENT OF FOREIGN LANGUAGES

WE HEREBY RECOMMEND THAT THIS THESIS

BY
HAMID REZA MOEINI ASL

ENTTTLED BY THE CONSTRUCT VALIDATION OF READING COMPREHENSION TESTS BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS IN TEEL

Committee on Final Examination:

..... Dr. Hossein Farhady,
Advisor

..... Dr. Ataei,
Reader

..... Dr. Mohammad Hossein Keshavarz
Chairman of the English Department

۳۲۶۵۶

Dedicated to who :

Smiles whenever I feel embarrassed.

Whishes me the best whenever I feel frustrated.

Crosses her fingers whenever I ask for prayers.

Pushes me forward whenever I would rather move backward.

Keeps loving me whenever I cease to love.

Acknowledgements

I would like to express my deepest gratitude to Dr. Farhady for his thought- provoking ideas, initial encouragement, helpful insights, valuable assistance, useful comments as well as his meticulous reading of the final draft of this thesis.

I am also grateful to Dr. Ataei for his helpful suggestions during the project and his careful reading of the final draft.

Further, I would like to express my appreciations to all those who have stimulated and guided my thinking over the years by pertinent observation and critical comments. Of these may I single out a few, and apologize for omitting the others: Dr. Farhady, Dr. Keshavarz, Dr. Vossoughi, Dr. Ataei, and Miss Saetti.

I am indebted to Mr. Ahmady and Mr. Hesami for invaluable assistance in developing the reading comprehension test.

I am also indebted to the following people for their assistance in the data collection process: Dr. Maarefat, Dr. Asadi, Dr. Borzabady, Mis Saetti, Mr Aghvami, Mr Kayvani, to name just a few.

Last but not least, I wish to express my gratitude to my wife who has not only been too much patient during the years but also has encouraged me to move forward whenever I felt frustrated.

Abstract

The purpose of the present study was to investigate the latent pattern underlying reading ability. Some 272 male and female participants, English majors at the BA level, participated in the study.

Two valid tests were used in this study, the reading test developed exclusively for the purpose of this project and academic reading section of IELTS. To investigate the possible latent underlying traits, factor analysis was conducted on the data. The final analysis showed that a nine-factor solution and eight-factor solution were the best explanations of the factor patterns. It is concluded that the reading ability consists of eight or nine major skills which are distinguishable and testable.

Table of Contents

Title I
Dedication III
Acknowledgement IV
Abstract V
Table of Contents VI-VII

Chapter One: Introduction

1.1. Overview 1-2
1.2. Statement of the Problem and the Purpose of the Study 2-6
1.3. Significance and Justification of the Study 7-9
1.4. Research Questions and Hypotheses 10
1.5. Definition of Key Terms 10

Chapter Two: Review of Literature

2.1. Introduction 11-13
2.2. Definitions of Reading 13-16
2.3. Purposes of Reading Comprehension 16-17
2.4. Levels of Reading Comprehension 17-18
2.5. Models of Reading 18
2.5.1. Process Models 19-29
2.5.1.1. Bottom-up Approaches 19-21
2.5.1.2. Top-down Approaches 21-24
2.5.1.3. Critique of the Bottom-up models 24-26
2.5.1.4. Critique of Top-down Models 26-27
2.5.1.5. Interactive Approaches 27-29
2.5.1.6. Interactive-Compensatory Approaches 29
2.5.2. Componential Models 29-35
2.5.2.1. Two-Component Models 30-32
2.5.2.2. Three-Component Models 32-35
2.6. Different Kinds of Reading 35-38
2.7. Reading Comprehension Skills 38-40
2.8. Taxonomies of Reading Comprehension Skills 40-51
2.9. Divisibility of Reading 52-59

Chapter Three: Method

3.1. Subjects 60
3.2. Instrumentation 60-63
3.3. Design 64
3.4. Procedures 64-65
3.5. Data Analysis 65

Chapter Four: Results and Discussion

4.1. Introduction 66-67
4.2. Analysis No.1 67-70
4.3. Analysis No.2 71-76
4.4. Analysis No.3 77-81

Chapter Five: Conclusion

5.1. Restatement of the Problem 82-87
5.2. Pedagogical Implications 88-89
5.2.1. Implication for the EFL/ESL Testing 88
5.2.2. Implications for Teaching 88-89
5.2.3. Implications for Materials Development 89
5.3. Suggestions for Further Research 89-90

References 91-101

Appendix

1. IELTS TEST 102-113
2. Reading Comprehension Test developed for this study: The pilot study version 114-127
3. Reading Comprehension Test developed for this study: The final version 128-136
4. Table of specifications for the test 137

CHAPTER

ONE

Introduction

1.1. Overview

Reading is clearly one of the most important skills; in fact, in many instances around the world it is argued that reading is the most important foreign language skill (Eskey,1988). This is particularly the case where students have to read English materials for their own specialist subject but may never actually have to speak the language itself (ibid). Furthermore, "it not only serves as a source of information and pleasurable activity, but also as a means of consolidating and extending one's knowledge of the language" (Rivers, 1987).

Despite the "critical role of effective reading in academic progress and achievements" (Carrel,1988,p.1) reading research is just a little more than a hundred years old. Hence, not too much progress has been made in terms of determining the nature of reading process. Although there has been a lot of research during this short period, as a result of the double difficulty describing and understanding reading in second language, compared with reading in one's native language, the problem still remains (Chun & Plass, 1997). The problem rises as a consequence of the large number of factors that interact in a non-linear and non-sequential manner in

reading process. On the other hand, employers, admission officers, teachers, and other end-users of test data often want information especially about a candidate's reading ability (Weir, 1994). To supply such data, attempts should be made to reduce the possibility of muddied measurement and focus on testing the construct of reading and control the contamination of the measurement of this ability (ibid). As a result of this desire to measure reading ability and nothing else language testers are obliged to be as explicit as possible concerning the nature of the ability they want to test.

1.2. Statement of the Problem and the Purpose of the Study

Every test is intended to measure one or more constructs. A construct is a psychological concept, which derives from a theory of the ability to be tested (Alderson, 2000). Considering construct validity as the core of validity, Bachman and Palmer (1996) state: " a construct is the specific definition of an ability that provides the basis for a given test task and for interpreting scores derived from this task"(p. 64). Whatever the purpose of a test is, there's no way to avoid interpreting a language test score (Bachman, 1990).

But it should be kept in mind that there are some limitations which restrict the ability to make such inferences. A major concern

of language test development, therefore, is to minimize the effects of these limitations. To accomplish this, Thorndike and Hagen, (1977), cited in Bachman (1990), suggest three steps: (1) identifying and defining the construct theoretically; (2) defining the construct operationally, and (3) establishing procedures for quantifying observations. The first step in measurement of a given language ability, therefore, is to distinguish the constructs by defining it clearly, precisely, and unambiguously.

It can be concluded here that construct validity_ whether the test actually measures aspects of the behavior under consideration _is of particular importance if one is to rely on test scores (Powell, 1989). As tests are, in manner of speaking, operational definition of constructs, Brown (1994) insists that a teacher needs to be satisfied that a particular test is an adequate definition of a construct.

Despite the importance of defining the construct clearly and the significant increase in test usage across the world, numerous issues surrounding the testing of reading have remained to be settled (Johnson, 1986). One of the issues and probably a major one is lack of clarity of the reading process. As Grabe (1991) states: "Among language skills reading has been considered as the most studied yet the least understood process in language teaching and testing" (p.376). In reality, even if they know the reading process, it does not

mean at all that test developers are able to examine its occurrence in somebody through his performance.

Meanwhile, it should be taken into account that the act of reading consists of the deployment of a range of separate skills, abilities, or strategies. Therefore, researchers have long tried to establish what skills are essential to text comprehension. This has involved identifying a priori what skills are theoretically needed, and then devising tests aimed at measuring such skills (Alderson,2000). Typically the relationship between the items testing the different skills is then analyzed, to see to what extent they can be empirically isolated (and therefore presumably tested and taught)(ibid).

Much controversy surrounds such research. There is contradictory evidence as to whether these skills are separately identifiable (Rost , 1993). Different analyses of the same databases of skills have resulted in more or fewer factors that appear to underlie skilled understanding. Davies (1968) claimed that the skills tested in his list were empirically distinguishable, whereas Thorndike (1974), who reanalyzed the data claimed that only one skill (world knowledge) could be distinguished from the others. Spearitt (1972) also did a re-analysis of skills in Davies's list, and

claimed there were four separate factors: recalling word meanings, drawing inferences from the context; recognizing a writer's purpose, tone and mood; following the structure of the text.

As it will be seen in great detail in Chapter Two, there are statistical and judgmental reasons for doubting whether skills can be measured separately or whether subskills of reading can be shown to exist and be related to the ability to answer particular sorts of test questions.

Although there is no empirically established theory illustrating the separability or inseparability of reading skills, the issue is of crucial importance to the assessment of reading. And the issue is: "if we are not able to define what we mean by the 'ability to read', it will be difficult to devise means of assessing such abilities" (Alderson,2000).

It is noteworthy to mention here that a number of issues related to the testing of skills thus far have been investigated: how many underlying factors, or empirically separable skills, are there? Is reading simply one unitary skill? Can judges distinguish which skills the items are testing? Which skills contribute most to performance on reading tests? (Weir and Urquhart,1998).

Regarding the factors underlying reading, some believe in a general factor (Thrndike, 1917a;1917b;1917c; Thurstone, 1946; Thorndike, 1973-74;Goodman, 1976, Rost,1985,among others). However, there are some scholars who support the multiple factor view of reading (Gray, 1919; Davis, 1944; Greoben, 1982). Moreover, some others take a third position in that they believe in two factors underlying the reading ability, one accounting for vocabulary and literal reading and another accounting for general comprehension or inferential reading (Johnson & Reynolds, 1941; Stocker & Kropp, 1960; Vernon, 1962; Grabe,1997;to mention a few).

Considering the component skills of reading, some scholars have proposed at least six general component skills and knowledge areas (Grabe; 1991). These components and skills will be discussed in great detail in the review of literature section.

1.3. Significance and Justification of the Study

It is commonplace in theories of reading to seek to identify skills which underlie or contribute to the reading process. Sometimes, the skills identified relate to linguistic features of the text in general (Munby, 1978), sometimes these skills relate to different sorts of meaning in the text, and sometimes they relate to supposedly different levels of the understanding that readers can derive from the the text (see, Gray, 1960). The theoretical nature and status of these skills and their interrelationships are far from clear, and there are many issues to be resolved. However, "it is a common practice among teachers, testers and researchers of reading to assume that reading skills can be identified, taught, tested and researched" (Alderson,1991). Using different approaches to explore the nature of these skills have produced different and sometimes contradictory results. For example, Carr and Levy (1990) argue that reading can be broken down into underlying skills and strategies for the purpose of teaching and testing. This view held by many language testers and teachers is labelled as the divisibility view. Carrol (1993) re–analyzes over 30 factor–analytic studies and identifies four common factors in reading: (general) reading comprehension, special reading