Construct Validity of MSRT Reading Comprehension Module in Iranian Context

Elham Fallahian Sichani[1] & Omid Tabatabaei[1]

[1] English Department: Faculty of Humanities, Najafabad Branch, Islamic Azad University, Najafabad, Iran

Correspondence: Elham Fallahian Sichani, Department of English, Faculty of Humanities, Najafabad Branch, Islamic Azad University, Najafabad, Iran. Tel: 98-913-303-3669. E-mail: elhamfallahian@yahoo.com

Abstract

Many researchers were interested in validity of the language proficiency tests in the previous decades. The present study aims to study the construct validity of the Ministry of Science, Research, and Technology Reading Comprehension module (MSRT) in the Iranian context. After administering a standard language proficiency test (OPT) 65 intermediate EFL learners were selected. The participants passed some 72 hours learning MSRT techniques and the strategies needed to be adopted in the MSRT test sessions. After all these hours, MSRT test was administered to those learners in some institutes. Later EFL questionnaire on reading was run among them and they were asked to check the skills they have applied during MSRT reading comprehension test. Both qualitative and quantitative approaches helped this study to collect and analysis the data. By qualitative approaches, this study collected expert's and test takers' judgments and by quantitative approaches, collected factor analysis. The results revealed that there were a significant consent among the judgments of the experts and an agreement among the test takers on the skills being measured by MSRT reading module. Finally, explanatory factor analysis did not reveal similar findings as those in the judgmental phase of the study. The items in the MSRT reading comprehension tests didn't confirm that MSRT reading parts assess the reading skills in the Iranian context. This study highlighted the importance of designing and using more reliable and valid tests, for researchers and designers, and those who are interested in the use of such tests.

Keywords: MSRT, construct validity, reading comprehension, reading module, factor analysis

1. Introduction

Throughout the history of language testing, measurement has almost always been dealing with three fundamental questions, namely, how to test, what to test, and why to test (Canale, 1988; Bachman, 1990; Douglas & Chapplle, 1993; cited in Tavakoli, 2004). The ultimate goal of testing is to measure learning language ability which is a theoretical construct and cannot be directly assessed. Because of the emergence of communicative competence which emphasized the role of different competencies in using the language, there was a shift of emphasis from linguistic competence to communicative competence. This shift caused the language institutes and universities to provide standardized test such as TOEFL, IELTS, FCE, TOLIMO, MSRT, …. The concept of the test validity has attracted the attention of many language test constructors in the last few decades. The tendency towards validation studies has been more evident in relation to standardized or universal language tests of proficiency. The present investigation explores the construct validity of reading with a specific reference to the reading skills which are claimed to be assessed by the MSRT reading tests. The study considers the use of the MSRT reading tests in the context of Iran about EFL learners.

2. Material Studies

Test constructors have devised different tests based on different purposes; however, they have to make sure about the abilities being measured by individual items in their tests. In the reading comprehension tests, the most important thing is validity. Several researches have posed related issues and problems with reliability when talking about the abilities/skills measured by individual items in reading tests (e.g. Alderson, 1990a, 1990b). Language tests such as MSRT assess the language proficiency of students. It is quiet new to the Iranian context and has been used in assessing the abilities of Ph.D candidates who study different majors other than English at different universities. This present study aims to investigate the way such a test is viewed by Iranian candidates and whether the items in the reading paper of the MSRT measure the skills by the MSRT developing board. The

application of the results of this study might open windows to study the other universal language tests introduced to Iranian educational context. Reading comprehension tests are used in assessing all languages.

2.1 Reliability and Validity

"Reliability refers to the consistency of the students' scores which is received on alternative forms of the same test" (Wells & Wollack , 2003, p. 2). A test is considered reliable if it produces similar results repeatedly. Reliability is an essential characteristic of a good between the scores and the test, because when a test is not reliable, one could not rely on the scores obtained from the test as to be an accurate measure of students' achievement. Based on the results of the research, if there is the subtraction of the standard error of the measurement in the tests, so we will have the addition of reliability. Although reliability is necessary, that's not enough by itself because a test requires to be valid as well in order to be reliable, too. According to Bachman (1995) "reliability is a requirement for validity and the investigation of reliability and validity can be viewed as complementary aspects of identifying, estimating and interpreting different sources of variance in test scores" (p. 239). Traditionally, validity was classified into different types (e.g. content validity, face validity, criterion validity and construct validity), but according to Messick (1980), this view is inadequate (cited in Bachman, 1955 & Wood, 2001). Since this study focuses on construct validity, its nature and background are explained below.

2.2 Construct Validity

It explains the extent to which performance on tests is consistent with predictions that we make on the basis of a theory of abilities, or constructs. Construct validity in reading is as the ability we wish to test. The present study was concerned with investigating the construct validity of the MSRT test.

2.3 Reading Comprehension

Reading comprehension is a kind of thinking, problem-solving, reasoning and all these mental processes have their roots in past experiences, consequently, success in the process of reading depends on the reader's ability to make a clear relationship between written materials and his/her own past experience. The students can't understand the meaning of the written sentences just in the surface of them.

2.4 Standardized Language Proficiency Testing

A standardized test has standard purposes and criteria that are kept firm the shape of the test to another one. The criteria in large-scale standardized tests are designed to apply to a broad band of competencies that are usually not exclusive to one particular curriculum. The incorrect impression that all standardized tests consists of items that have predetermined responses presented in multiple-choice format is reducing these days. While it is true that most standardized tests have a multiple-choice format, by no means is multiple-choice a prerequisite characteristics. Multiple-choice format provides the test procedure with an "objective" means for determining correct and incorrect responses, and therefore is the preferred mode for large-scale tests.

2.5 MSRT Test

MSRT is an acronym "Ministry of Science, Research and Technology". It is a standardized test of English language proficiency. It is managed by Sanjesh Organization and was established in 2013.The emergence of MSRT goes back to 2002, the first version of an exam called MCHE: Ministry of Culture and Higher Education". It was introduced as a way of testing applicants to graduate from national universities for Ph.D. degrees. It runs every month. The objective of the test is to test overall proficiency. This test has only one type that is paper-based. Its response modes are multiple-choice responses. MSRT reading test forms are similar to other kinds of tests. They consist of three parts: Listening Comprehension, Structure and Written Expression, and Reading Comprehension. This test contains four passages usually followed by 10 questions forming a total of 40 questions, to be administered in 45 minutes. There are a lot of Iranian applicants who should take these tests because of getting their certificates from their universities. Although the candidates spend a lot of time and fortune learning English, they still have some hardship in answering MSRT questions in the exams. It is really useful to know whether or not these tests are suitable for Iranian context. The primary concern for the test reading paper is to focus on various reading passages following instructions, finding main ideas, identifying underlying concept, identifying relationships between the main ideas and drawing logical inferences. These cases are characterized by MSRT (2009) are multiple choice items, identification of writer's views/ attitudes/ claims, finding the main idea of the passage, identifying the true and false sentences, referring to the pronunciations and finding the best synonyms for the special words.

3. Area Description

The objective of this research is to inquiry the construct validity of MSRT reading section interior the situation of Iranian test takers. Due to a number of reasons, only reading skill has been selected. First, reading comprehension is the serious affair methodological cases in the field of Teaching and learning English, second it is the main goal of the learners in the countries like Iran where the people learn English as a foreign language. Third, reading comprehension tests are widely used and more objectively scored than the other skills. This study aims to tackle the problem of construct validity, i.e. the skills measured by the MSRT reading paper, through the process of gathering different types of evidence: (1) EFL expert judgments, (2) the EFL test-takers decisions on the skills assessed by the MSRT reading items and finally (3) factor analysis to analyze items.

The first two are qualitatively oriented and third data collection approach is quantitatively oriented.

3.1 Test Hypotheses

Based on the above problems in the nature of reading construct and reading skills, the present research struggles to answer these three lower questions:

1) Do Iranian EFL testing experts believe that different items on MSRT reading section would actually measure the reading ability of the applicants?

2) Do Iranian EFL learners believe that different items on MSRT reading section would actually measure their reading ability?

3) Does the result of explanatory factor analysis confirm that it would assess the L2 reading skill in the Iranian context?

3.2 Research Question

These three research questions can be formulated in the front of the following null hypotheses.

RH1) Iranian EFL experts do not believe that different items on MSRT reading section would actually measure the reading ability of applicants.

RH2) Iranian EFL students do not believe that different items on MSRT reading section would actually measure their reading ability.

RH3) Result of explanatory factor analysis does not confirm that MSRT reading parts assess the reading skills in the Iranian context.

4. Method

4.1 Participants

To collect the required data, a total of a hundred Iranian EF learners both male and female, age 22 to 40 were randomly selected out of some English institutes in Isfahan. All the participants were getting prepared to take proficiency test like MSRT, to pursue their education.

As regards their educational background, they all had university degree, at least M.A in majors other than English from state or Azad universities all around Iran. Oxford Placement Test (OPT, Allan, 2005) was run since all willing to take the test needed high level of proficiency in English. Based on the scoring system of OPT, Sixty five participants who had intermediate or upper intermediate levels in English were chosen to take the special course aimed at helping them improve their general proficiency. They all passed seventy two hours learning proficiency test techniques similar to those of MSRT and the strategies needed to be adopted in the MSRT test sessions in Jahad-e-Daneshgahi Institute. In general, twenty hours were devoted to improve reading comprehension techniques required for the test.

Twenty five university lecturers were also asked to participate in the study. These highly educated participants, all having at least five years of teaching experience mostly with master and a few with PhD degree, were selected from the state and Azad universities as well as different language institutes.

4.2 Instruments

Instruments of this study consist of three different parts.

4.2.1 Oxford Placement Test (OPT, Allen, 2005)

The OPT (Allen, 2005, See Appendix A) includes 200 items totally.100 items out of 200 items in this test relates to the grammar section. The researchers selected just grammar part for this study. After administering the test, the gathered conclusions were analyzed and the students who received 70 or more in this test were selected as

intermediate and upper- intermediate learners.

4.2.2 MSRT Test

MSRT test was used to elicit the participants' scores. The scoring procedure follows the guidelines provided by the MSRT handbook (1392). Since this study has focused on the MSRT reading module, merely MSRT reading scores were considered. All reading sections had 4 texts usually followed by 10 questions for a total of 40 questions. (Appendix C).

4.2.3 Questionnaire

Measuring the construct validity of the test is the main goal of this study. The students should have some abilities for the reason that they can overcome to all target questions, so they need the construct validity.

To do so, a questionnaire based on Weir's checklist (1997) which was adopted by Barati (2005) was used to elicit both participants' and university lecturers' ideas (Appendix E). The present study used this taxonomy since it is through and has been validated qualitatively.

4.3 Procedures

First of all OPT was administrated to a group of learners (N=100) in order to select 65 intermediate and upper intermediate level learners from State and Azad universities in Isfahan. Their scores were 70 or more.

The selected participants passed seventy two hours learning MSRT techniques and the strategies needed to be adopted in the MSRT test sessions in Jahad-E-Daneshgahi institute. One third of the mentioned hours were completely allocated to improve reading comprehension skills required for the test. After passing all these hours learning MSRT strategies, MSRT test was administrated to the participants in the institutes.

On the other hand, there were twenty five teachers to cooperate this research by taking the MSRT reading section as the expert judges. The learners took this test, too. The experts' judges are so important because they should determine the relationship between the skills in the EFL group of reading skills and individual test items in this test.

5. Results

Conducting a reliable and valid test has always been a goal for test designers, so for this test at first the correspondence according to answer the first and the second research questions, the average of the experts' and learners' judgments over skill/item correspondence were investigated. Exploratory factor analysis was used. At first, descriptive data analysis was conducted to examine the normal distribution of the variables (participants) across the MSRT reading parts. Then the reliability was calculated using Cronbach's Alpha. Finally the reading skills were identified by running explanatory factor analysis.

Table 1. Cronbach's Alpha of the questionnaire

Variable	Number of the questions	Cronbach alpha
Level A	5	0.564
Level B	5	0.544
Level C	11	0.612
Level D	14	0.640
Whole Questionnaire	35	0.799

Note.The whole questionnaire is valid because is slightly near above 0.7.

When Cronbach alpha of the instrument is slightly above the minimum, 0.7 suggested by Nanly (1987), the instrument enjoys an appropriate level of reliability. In this study for estimating the reliability, Cronbach alpha was estimated for each level-of the questionnaire and the whole questionnaire. Cronbach' alpha coefficient of the whole questionnaire showed that it had reasonably high internal consistency. However, the Cronbach's alpha coefficient for the levels A and B was 0.6 and for levels C and D met a moderate alpha Coefficient slightly near 0.7.

Regarding the distributed questionnaire for each question, 3 choices have been determined. For analyzing the data, a triple index has been used to have valid statistics based on Linkert-scale (Appendix E) Yes with the score

of 3, No Idea with the score of 2, and No with the score of l. In order to determine that the population has agreed on employing any factors, the averages were compared by using one-tailed t-test. Therefore when HO is confirmed, it means that most of the population did not use that technique or at least had no idea about using it. On the other hand, when HO is rejected, it means that most of the students or experts have applied that strategy. Here the experts' and learners' agreement on using the factors have been investigated respectively. In order to answer the first null hypothesis, 25 EFL Ph.D. or M.A. holders were asked to check the beneficial factors in answering the MSRT reading comprehension tests. It can be said that the first null hypothesis was rejected.

Table 2. Investigating experts' agreement on using total factors

Variable	Mean	Sig. (2tailed)	Mean Difference	Lower Bound	Upper Bound		
Level_A	2.2070	.001	.39200	.1863	.5977	Reject H0	Agreement
Level_B	2.0105	.000	.68800	.5509	.8251	Reject H0	Agreement
Level_C	2.0877	.000	.52000	.3962	.6438	Reject H0	Agreement
Level_D	2.1817	.000	.59714	.4889	.7054	Reject H0	Agreement

To address the second null hypothesis, the present study asked 65 students to participate in MSRT tests and after answering the reading comprehension questions, they were asked to check the factors which were useful for answering them. Based on the findings, the second null hypothesis of this study was also rejected.

Table 3. Investigating students' agreement on using total factors

Variable	Mean	Sig. (2-tailed)	Mean Difference	Lower Bound	Upper Bound		
Level_A	2.3920	.008	.20702	.0569	.3571	Reject H0	Agreement
Level_B	2.6880	.864	.01053	.1117	.1327	Confirm H0	Disagreement
Level_C	2.5200	.074	.08772	.0088	.1842	Confirm H0	Disagreement
Level_D	2.5971	.000	.18170	.0954	.2680	Reject	Agreement

Based on the result of this study, there is an agreement between the students on the skills to be assessed by the MSRT reading comprehension tests. The qualitative judgments of this study were against with Tavakoli & Barati (see 2.7.5).

Table 4. Rotated component matrix[a] for masters

	Component												
1	2	3	4	5	6	7	8	9	10	11	12	13
								0.487				
				0.631								
0.456												
0.688												
		0.507										
		0.678										
		0.615										
		0.544										
			0.46									
									0.342			
									0.482			

		0.496			
		0.283			
			0.576		
				0.524	
			0.578		
				0.371	
	0.517				
					0.498
			0.666		
0.648					
0.568					
0.742					
		0.636			
		0.541			
		0.407			
0.665					
			0.296		
					0.329
0.622					
0.555					
0.493					
0.547					
0.503					
0.455					

Note. Extraction Method: Principal Component Analysis.

a. 13 components extracted.

Table 5. Comparing the student' and expert's agreement on using total factors

Factors	Experts	Students
Level A	agreement	agreement
Level B	agreement	disagreement
Level C	agreement	disagreement
Level D	agreement	agreement
Total	agreement	agreement

In order to address the third null hypothesis, qualitative approach was used. To this end, explanatory factor analysis was run on both the experts' and students' answers in the questionnaire. Based on the total variance explained for the experts the most used 13 components were extracted.

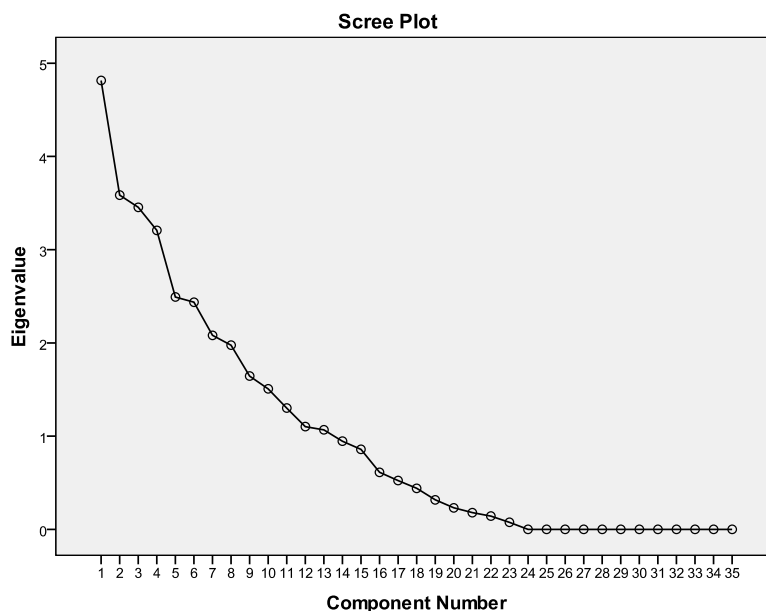Based on this plan, we had a pause after the thirteen parts.

Figure 1. Expert's eigenvalue

Based on the component matrix for the students, the rotated component matrix for students was distinguished. Those factors which had the coefficient less than 0.4 were eliminated because they did not have internal consistency.

The analysis of explanatory factor analysis illustrated that the students' ideas did not match with the category of MSRT reading comprehension tests. Based on the expert's ideas, there were 14 factors out of 35 factors with variance above 1. Those 14 factors were found useful while they were answering the MSRT reading comprehension tests.

Table 6. Component matrix[a] for students

	Component													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Q1	.521													
Q2		.509												
Q3												.360		
Q4		.585												
Q5	.475													
Q6		.514												
Q7														.345
Q8			.553											
Q9			.499											
Q10			.428											
Q11			.442											
Q12				.552										
Q13													.364	
Q14	.387													
Q15	.319													
Q16			.303											
Q17							.409							

Q18	.198				
Q19			.297		
Q20	.517				
Q21			.484		
Q22	.	0.414			
Q23	.428				
Q24	.613				
Q25				.460	
Q26					.348
Q27			.478		
Q28			.314		
Q29				.356	
Q30					.340
Q31					.369
Q32	.363				
Q33	.504				
Q34				.480	
Q35	.332				

Note. Extraction Method: Principal Component Analysis.

a. 14 components extracted.

All these factors are displayed in Screeplot of the students' ideas in the Figure 2. These 14 factors had Eigenvalue above 1.
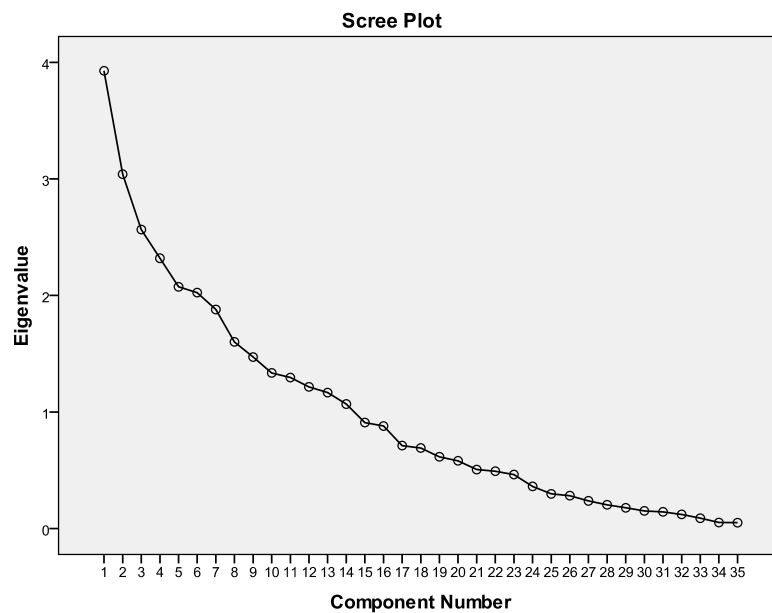

Figure 2. Student's eigenvalue

There were 14 factors out of 35 factors with variance above 1. Based on the conclusions of the Factor Analysis the items in MSRT comprehension tests were not conform on different components.

Bachman (1995, p. 237) "validity can be viewed as complementary aspects of identifying, estimating, and interpreting different sources of variance in test scores". The results of factor analysis indicated that the MSRT

reading comprehension module used in this study was not valid enough to be used with confidence in order to measure EFL learners' reading ability. The moderate level of Cronbach' alpha coefficient of questionnaire 0.799 >0.7 showed that the questionnaire was an appropriate instrument to be used. The findings also showed that the 2.60% of the Ph.D and M.A. holders agreed on applying the fourth skill on answering MSRT test in Reading Expeditiously for Global Comprehension which was getting the main idea of the paragraph. On the other hand, the students agreed on employing the first skills in this level which, showed that 2.33% of the learners first had looked at the topic. In level B the 2.84% of experts came to agreement that the tenth skills was identified by MSRT reading comprehension tests in Reading Expeditiously for Local Comprehension which was looking for examples. 2.07% of learners came to agreement that the seventh skills identified by MSRT reading comprehension test in level B which looking for specific information. Investigating the level Reading Carefully for Global Comprehension 2.26% of the students found seventeenth skill was really important, being able to answer the questions which related to key information of the text, without which the comprehension of the text would be difficult. However 2.84% of the experts agreed on the importance of the twentieth skill being able to make on outline of the main points. Finally 2.38% of the students believed that 30th skill was really important in level D Reading carefully for Local Comprehension. On the other hand, 2.88% of the experts had applied the 30th skill more than the others in level D, too.

6. Discussion

The aim of this study was assessing the construct validity of MSRT test so three hypotheses and research questions were formulated. Regarding the first research question the normality of distribution and the Cronbach' alpha coefficient of questionnaire were estimated. The present study used judgmental approach to examine the identifiability of MSRT reading module. MSRT reading questions were handed to 25 EFL experts. They were asked to decide about the correspondence between MSRT reading comprehension tests and the reading skills in the EFL questionnaire. The experts agreed on using all levels while answering MSRT reading comprehension. The average for their agreement for level A was 2.568 for level B, level C and for level D was out of 3.

In summary, there is an agreement among the greatest of experts' judgment on the reading part being assessed by the items in MSRT Reading module. T0 answer the first question which stated that: do Iranian EFL testing experts believe that different items on MSRT reading section would actually measure the reading ability of the applicants, was yes. Based on their agreement on applying those variables in MSRT reading comprehension tests, the researcher came to the conclusion that different items on MSRT reading section would measure the reading ability of the applicants.

To address the second research question and examine the EFL experts' idea on the relation between the items in MSRT reading module and the skills from the EFL questionnaire, a group of 65 EFL Iranian students were asked to identify which skill from the questionnaire they thought that they had applied while answering the MSRT reading comprehension tests. The second question of this study stated that: do Iranian EFL learners believe that different items on MSRT reading section would actually measure their reading ability? The answer is yes. According to these factors Iranian learners believed that MSRT reading comprehension test would actually measure their reading skills. The idea is that what EFL learners cannot come to agreement on is the skills being assessed at the times of the test. There may be a problem in determining what the items are really testing. Their agreement here could be assumed as a piece of evidence for the validity of MSRT. They Agreed on using the 4 variable when answering the tests, level A with the mean of 2.2070, level B with the mean of 2.0105, level C with the mean of 2.0877, level D with the mean of 2.1817 and the total mean of 2.1217 which illustrated that majority of learners agreed on the skills assessed by MSRT reading comprehension test. Comparing with experts' idea with the total mean of 2.5493 and level A with the mean of 2.3920, level B with the mean of 2.6880, level C with the mean of 2.5200, and level D with the mean of 2.5971, it seemed that there was a considerable agreement. The key findings of the present study indicated that the MSRT reading skills assess Iranian EFL learners' ability qualitatively by EFL experts' and learners' ideas but not via FA. The Iranian context and rejected the first and second null research hypotheses.

To address the third question which stated that, does exploratory factor analysis confirm that MSRT reading module would assess the L2 reading skill in the Iranian context? Factor analysis was run. The experts shared in thirteen factors from the questionnaire being assessed by the MSRT reading module. That is nearly all judges agreed on the skills to be assessed by these 13 factors in the questionnaire.

In summary, the answer of the third question which states that does exploratory factor analysis confirm that MSRT reading module would assess the L2 reading skill in the Iranian context was no. Because as it was mentioned above, the experts have merely found that 13 factors had factor load and the students have agreed that

14 factors had factor loads. The items in reading comprehension tests don't test 13 reading ability of Iranian MSRT candidates.

7. Conclusion

For examining the validity of MSRT reading test, qualitative and quantitative approaches were used. Based on the quantitative phase of the study, 65 EFL students' answers and 25 expert's answers on the MSRT reading comprehension test and Factor Analysis were examined. Based on the qualitative phase of the study, EFL expert judges' and the EFL test takers' decisions on the relation between the reading skills in the EFL questionnaire and the items in the MSRT reading comprehension test.

Based on the results of this study, MSRT reading skills assess Iranian EFL learners' ability qualitatively by EFL experts and learners' ideas but not via FA. Hence the quantitative findings Claims that the MSRT reading module assesses reading ability of test-takers. Therefore, the findings showed the improvement of construct validity of the reading MSRT module in the Iranian context and rejection of the first and second null research hypotheses.

The judgmental phase of the present study could not support the findings of the quantitative phase (FA). In the qualitative phase of the study, the majority of the EFL expert judges and the test-takers agreed on the skills to be assessed by the MSRT reading module. In other words, the MSRT reading skills were similarly identified by the expert judges and the test-takers to be assessing heterogeneous list of skills. However the skills were not confirmed when Factor Analysis was run. The results of the Factor Analysis did not show the items in the MSRT reading module to be loaded on different components. Since the qualitative phase of the study, subjective one, confirms that MSRT reading module assess EFL learners' reading ability in Iranian context, but the quantitative phase, objective one could not confirm different skills for the construct of the MSRT reading module was not known to be unitary in the context of Iranian EFL learners. So, the answering of these questions are:

1) Do the greatest of the Iranian EFL testing experts believe that different items on MSRT reading section would actually measure the reading ability of the applicants?

- YES

2) Do the greatest of the Iranian EFL students believe that different items on MSRT reading section would actually measure their reading ability?

- YES

3) Does the result of explanatory factor analysis confirm that MSRT reading parts assess the reading skills in Iranian context?

- NO

References

Alderson, J. C. (1990a).Testing reading comprehension skills. *Reading in a foreign language, 6*(2), 425-438.

Alderson, J. C. (2000). *Assessing Reading Cambridge*: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511732935

Anderson, N. J. (1999). *Exposing Second Language Reading, Issues and Strategies*. Cambridge: University Press.

Allan, D. (2005). *Oxford Placement Test*. Retrived Agust 30, 2012, from www.waterstones.com EFL Examination practice tests.

Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford:Oxford University Press.

Bachman, L. F., Davidson, F., Ryan, K., & Choi, I. C. (1995). An investigation into the comparability of two tests of English as a foreign language: The Cambridge TOEFL comparability study: Final Report. Cambridge University Press.

Barati, H. (2005). *Test Taking Strategies and the Assessment of Reading Skills*: An Approach to Construct Validation; Unpublished Ph. D. Thesis, University of Bristol.

Brown, H. D. (2004). *Language Assessment Principals and Classroom Practices*: San Francisco State University: Longman.

Canale, M. (1998). *From Communicative Competence to Communicative Language Pedagogy*. In J. C. Richards & R. Schmidt (Eds.), Language and Communication. London: Longman.

Carroll, P. (1988). Three Components of background knowledge in Redaing Comprehension. *Language Learnin, 33*(2), 183-207. http://dx.doi.org/10.1111/j.1467-1770.1983.tb00534.x

Cook, V. (1991). *Second Language Learning and Language Teaching*, London: Edward Arnold.

Cronbach, L. J. (1988). Coefficient alpha and the international structure of tests. *Psychometrika, 16*, 292-334.

Farhady, H. Ja'farpur, A, & Birjandi, P. (2001). *Testing Language Skills: From the theory to practice*. Tehran. Samt.

Messick, S. (1998). Validity. In R. Linn (Ed.), *Educational Measurement* (pp. 13-103). New York: American Council on Education and Macmillan.

Tavakoli, E., & Barati, H. (2011). *Investigating the Construct Validity of the FCE Reading Paper in Iranian EFL Context, Journal of Language Teaching and Research, 2*(1), 239-247. http://dx.doi.org/10.4304/jltr.2.1.239-247

Weir, C. J. (1983). *Identifying the Language needs of overseas students in tertiary education in the United Kingdom. Ph.D. thesis*, Institute of Education, University of London.

Weir, C. J., & Urquhart, A. H. (1998). *Reading in a second language: Process, Product and Practice*, Longman, New York.

Appendix

There are some appendices for complete the proofs.

Appendix 1. Total variance explained for students

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.927	11.221	11.221	3.927	11.221	11.221
2	3.040	8.686	19.907	3.040	8.686	19.907
3	2.566	7.330	27.237	2.566	7.330	27.237
4	2.319	6.626	33.863	2.319	6.626	33.863
5	2.074	5.926	39.789	2.074	5.926	39.789
6	2.024	5.783	45.572	2.024	5.783	45.572
7	1.880	5.370	50.943	1.880	5.370	50.943
8	1.601	4.575	55.517	1.601	4.575	55.517
9	1.472	4.205	59.722	1.472	4.205	59.722
10	1.335	3.815	63.537	1.335	3.815	63.537
11	1.296	3.702	67.239	1.296	3.702	67.239
12	1.215	3.472	70.711	1.215	3.472	70.711
13	1.167	3.334	74.044	1.167	3.334	74.044
14	1.068	3.052	77.096	1.068	3.052	77.096
15	.909	2.598	79.695			
16	.880	2.513	82.208			
17	.712	2.033	84.241			
18	.691	1.974	86.214			
19	.616	1.759	87.973			
20	.581	1.660	89.634			
21	.506	1.446	91.079			
22	.492	1.406	92.485			
23	.463	1.324	93.809			
24	.362	1.033	94.842			
25	.298	.851	95.693			
26	.282	.805	96.498			
27	.237	.678	97.176			
28	.204	.582	97.758			
29	.178	.508	98.267			
30	.151	.432	98.699			
31	.143	.409	99.108			
32	.121	.347	99.454			
33	.089	.253	99.707			
34	.052	.149	99.856			
35	.050	.144	100.000			

Extraction Method: Principal Component Analysis.

Appendix 2. Total variance explained for masters

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.815	13.757	13.757	4.815	13.757	13.757
2	3.585	10.244	24.001	3.585	10.244	24.001
3	3.455	9.871	33.872	3.455	9.871	33.872
4	3.208	9.165	43.037	3.208	9.165	43.037
5	2.493	7.122	50.159	2.493	7.122	50.159
6	2.437	6.962	57.121	2.437	6.962	57.121
7	2.081	5.946	63.067	2.081	5.946	63.067
8	1.976	5.646	68.713	1.976	5.646	68.713
9	1.644	4.698	73.411	1.644	4.698	73.411
10	1.507	4.306	77.717	1.507	4.306	77.717
11	1.301	3.718	81.435	1.301	3.718	81.435
12	1.103	3.15	84.585	1.103	3.15	84.585
13	1.068	3.051	87.636	1.068	3.051	87.636
14	0.946	2.704	90.34			
15	0.858	2.452	92.792			
16	0.612	1.749	94.541			
17	0.524	1.497	96.037			
18	0.441	1.259	97.296			
19	0.317	0.906	98.202			
20	0.231	0.661	98.863			
21	0.179	0.512	99.375			
22	0.143	0.407	99.782			
23	0.076	0.218	100			
24	7.52E-16	2.15E-15	100			
25	4.15E-16	1.19E-15	100			
26	2.61E-16	7.46E-16	100			
27	1.80E-16	5.14E-16	100			
28	9.89E-17	2.83E-16	100			
29	1.08E-17	3.10E-17	100			
30	-9.53E-17	-2.72E-16	100			
31	-1.17E-16	-3.34E-16	100			
32	-2.81E-16	-8.03E-16	100			
33	-3.82E-16	-1.09E-15	100			
34	-5.02E-16	-1.43E-15	100			
35	-5.35E-16	-1.53E-15	100			

Extraction Method: Principal Component Analysis.

Appendix 3. Comparison of experts' and students' agreement

Factors	Masters	Students
Q1	Agreement	Agreement
Q2	Agreement	Agreement
Q3	Disagreement	Disagreement
Q4	Agreement	Disagreement
Q5	Disagreement	Disagreement
Q6	Agreement	Disagreement
Q7	Agreement	Disagreement
Q8	Agreement	Disagreement
Q9	Agreement	Disagreement
Q10	Agreement	Disagreement
Q11	Disagreement	Disagreement
Q12	Disagreement	Disagreement
Q13	Agreement	Disagreement
Q14	Agreement	Disagreement
Q15	Agreement	Disagreement
Q16	Agreement	Disagreement
Q17	Agreement	Agreement
Q18	Agreement	Disagreement
Q19	Agreement	Disagreement
Q20	Agreemen	Disagreement
Q21	Disagreement	Disagreement
Q22	Disagreement	Disagreement
Q23	Agreement	Agreement
Q24	Agreement	Disagreement
Q25	Agreement	Agreement
Q26	Agreement	Disagreement
Q27	Agreement	Agreement
Q28	Agreement	Agreement
Q29	Agreement	Disagreement
Q30	Agreement	Agreement
Q31	Disagreement	Agreement
Q32	Agreement	Disagreement
Q33	Agreement	Disagreement
Q34	Agreement	Disagreement
Q35	Agreement	Disagreement

Copyrights