

## **An Analysis of the Reading Section in Ferdowsi Persian Proficiency Test for Non-native Speakers**

Mohsen Roudmajani <sup>1</sup>, Dr.Behzad Ghonsooli <sup>2\*</sup>, Dr.Ehsan Ghabol <sup>3</sup>,  
Dr. Mohammad Jafar Yahaghi <sup>4</sup>

<sup>1</sup> Ph.D. Candidate of Persian Language and Literature, Ferdowsi University of Mashhad, Mashhad, Iran

<sup>2</sup> Professor, Department of English Language and Literature, Ferdowsi University of Mashhad, Mashhad, Iran

<sup>3</sup> Assistant Professor, Department of Persian Language and Literature, Ferdowsi University of Mashhad, Mashhad, Iran

<sup>4</sup> Professor, Department of Persian Language and Literature, Ferdowsi University of Mashhad, Mashhad, Iran

(Received: April 10, 2018 Accepted: November 26, 2018)

---

### **Extended Abstract**

#### **1. Introduction**

People are not born with reading ability and it is a new phenomenon in human history. You can still find some people who cannot read. But these days weakness in reading ability can cause serious problems for us. Also Reading ability and academic life are indispensable and many of the students activities are interweaved with reading. Books, essays, instructions, research reports are just a few examples of the activities in academic environment which need reading ability.

Because of the importance of reading ability most of Persian proficiency tests include a reading section. Ferdowsi University's Persian proficiency test which is designed to test the language proficiency of non-Persian students, include reading assessment. This test is official test of Ferdowsi University and it is held twice a year in Iran and some cities of Iraq. In this research we try to study the reading part of this test. We actually try to answer two questions:

1- What is the reliability of the reading section of this test?

2- What is the relation between test items and test takers abilities?

At the end we try to give some recommendation to improve this test's reliability. We will do this by trying to define the reading skill's construct and we also use language task framework which is a powerful tool to study reading activities in academic environment.

---

\* Corresponding Author Email: ghonsooly@ferdowsi.um.ac.ir

## 2. Review of Literature

The history of testing can be divided into three periods: the pre-academic period, the psychometric period; the structuralism and the psychology-sociology of language (Spolsky, 1977, cited in Davis, 2008). The first period goes back to more than two thousand years ago, when the empires used the test to select new members. The second period is related to the time when psychometrists devised new statistical tools for measuring test reliability which coincided with the constructivist flow in linguistics and led to the development of a discrete point measurement approach. However, in the late twentieth century, we see another approach, called integrated measurement, which emphasized language assessment in context, and, unlike the structuralists, did not consider language division into smaller and discrete points as efficient methods of testing.

In the second half of the twentieth century, with the development of scientific tools, companies emerged that produced mass-scale testing, thereby transforming testing into an industry (Spolsky, 2008). One of these companies is ETS. In 1964, on the basis of a discrete point assessment approach, the company designed the first proficiency test and little by little made changes to its test structure. However, in a project called TOEFL 2000, ETS decided to make major changes to this proficiency test. Reading Skills Assessment was also a part of the TOEFL 2000 project, which resulted in *Assessing Second Language Academic Reading from a Communicative Perspective: Relevance for TOEFL 2000* (Hudson, 1996) and the Reading Skills Framework (Enright et al., 2000).

Attempts have been made in Farsi to design a sufficiency test (Mousavi, 1999; Ghonsooli, 2010; Mamghmani, 2010; Jalili, 2011; Golpour, 2015). These studies devote part of their test to reading skills. Although what distinguishes the present study from all the research done is the precise definition of reading skill constructs, providing a model for measuring this skill, as well as a task-based view of the test design.

## 3. Method

First in order to understand what reading ability is we tried to define its constructs. Based on the latest model we found that reading consists of lower level processes and higher level processes. Lower level processes include word recognition, syntax parsing and meaning proposition encoding which are mainly automatic. On the other hand, the higher level processes consist of text model, situational model and executive control mechanism. In order to understand a text, the reader use higher level processes and lower level processes interchangeably.

Also a framework was introduced to analyze the language tasks in an academic environment. This framework helps us to understand and simulate language tasks. It has three different parts called situation, text content and test features. This framework is designed based on communicative perspectives in language learning and assessment. The situation explains the context of the task in which it is performed. Its participants, purpose, content, setting and register. The text content include grammatical, functional and discourse features of the text and test features section describes the factors such as type of questions, type of response and scoring procedures.

The data in this research was gathered by a test which was held in Ferdowsi International center for teaching Persian to nonnative speakers. This test was performed to see to what extent test takers, have the ability to cope with academic activities. The cut score for this test is 60 out of 100. The test has five section including reading, listening, writing, speaking and structures.

90 test takers took part in the test. 33 of them were female and 57 were male. Most of the participants were Iraqi (63) and the rest of them were from Indonesia, Lebanon, Syria, Australia, South Korea, Pakistan, and Yemen. 38 of the examinees were studying in different field of engineering, 17 of them were studying humanities, 13 sciences and 7 of them were students of medical sciences. The field of studying for the rest of the examinees was not known.

The Persian proficiency test was held in Mordad 25<sup>th</sup>, 1996 at 8 o'clock and it took about 3 and half an hour. As the reading section was the second part of the test it started at half past eight. This part of the test took 60 minutes and the students answered 25 questions. Students were in 10 different rooms and there were 9 students in each class.

After the test was performed the papers were marked the data was ready for Rasch analysis. The step was done with the help of WINSTEPS software.

#### **4. Results and Discussion**

The analysis showed that all the questions except item number 21 has appropriate infit and outfit amount. It shows that the test is unidimensional and it is possible to be analyzed by Rasch model. The outfit for item number 21 is 1.53 which is more than the standard amount. This question is multiple choice and it is also so simple. It seems that the item is ambiguous. The results also showed that the test reliability is 76%. We know that reading test reliability should be more than 90%. So the test needs to improve its reliability.

Also the Write map showed that the test need some more difficult items because for the 14 top students in the exam there are only two items. And also

there are just 4 items appropriate for the 36 top students of the test. The exam cannot differentiate between intermediate and advanced students.

## 5. Conclusion

The results showed that the test reliability is not high so the test designers should improve their reliability for example by adding more question and items. On the other hand the item- examinee map showed that the test assess elementary or advanced students reading ability it means that this test is not appropriate to assess advanced student's reading skill.

The test includes three texts and 25 items. Compared to TOEFL exam which has about 40 items, the test needs more questions. In order to improve its reliability, more items can be added to the test. But considering type of information requested the test covers different levels. Which is a good aspect of this test. Also each text has about 500 words which is somehow appropriate for assessing academic reading skill.

The results showed that the test is easy. The reason for this is the matching type of the questions. For fourteen questions of the test the students should locate and cycle the answers. We know that usually these type of questions are not so difficult. So in order to improve the test difficulty other matching types such as integrative and productive questions should be added.

From the reading purpose perspective, the test questions mainly follow two purposes, reading to find information and reading for general information. The test needs items with other types of purposes such as reading to learn and reading to integrate information.

**Keywords:** Reading skill, Persian proficiency test, task base assessment, Proficiency base assessment, test reliability

## References (In Persian)

- Ghaffar Samar, R., Shirazizadeh, M. & Kiani, M. (2013). بسترها، چشم‌اندازها، کاربردها و چالش‌های مطالعه واژگان در متون دانشگاهی، ضرورت توجهی بیشتر به زبان فارسی و زبان‌آموزان فارسی‌زبان [Grounds, perspectives, practices and problems of studying vocabulary in academic texts, the need for paying more attention to Persian language and Persian language learners]. *Domahnameye jostarhaye zabani*, 25, 153-181.
- Gholpour, L. (2015). طراحی و اعتبارسنجی آزمون بسندگی زبان فارسی بر چهار مهارت زبانی [Designing and validating Persian proficiency test based on four language skills (Doctoral dissertation). Pyam-e-Noor University, Tehran, Iran.
- Ghonsooli, B. (2010). طراحی و اعتبارسنجی آزمون بسندگی زبان [Designing and validating Persian proficiency test]. *Pajoheshhaye zabanhaye khareji* 57, 115-129.

- Ghonsooli, B. (2013). *برون‌فکنی اندیشه در مهارت خواندن در زبان دوم* [Thinking aloud in second language teaching]. Mashhad, Iran: Tabaran University Press.
- Jalili, S. A. (2011). *آزمون مهارتی فارسی (آمفا) برپایه چهار مهارت اصلی زبان* [Persian proficiency test based on four main language skills] (Unpublished master's thesis). Allameh Tabatabaee University, Tehran, Iran.
- Mousavi, Sorayya (1999). *طراحی و اعتبارسنجی آزمون استاندارد زبان فارسی برپایه چهارمهارت زبانی* [Designing and validating standard Persian proficiency test based on four language skills]. (Unpublished master's thesis). Shiraz University, Shiraz, Iran.

### References (In English)

- Ary, D., Jacobs, L. C., Sorensen, C., & Walker, D. A. (2014). *Introduction to research in education*. Belmont: Wadsworth.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford, England: Oxford University Press.
- Bachman, L. (2004). *Statistical analysis for language Assessment*. New York, NY: Cambridge University Press.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Chapelle, C. A., Enright M. K., Jamieson, J. M. (2008). Test score interpretation and use. In *Building a validity argument for the test of English as a foreign language* (pp. 1-25). Edited by Chapelle, C. A., Enright M. K., Jamieson, J. M. New York: Routledge.
- Chen, J. Sheehan, K. M. (2015). *Analyzing and comparing reading stimulus materials across the TOEFL family of assessments (TOEFLiBT-26)*. Princeton, NJ: Educational Testing System.
- Cullen, P., French, A., & Jakeman, V. (2014). *The official Cambridge guide to IELTS*. China: Cambridge University Press.
- Enright, M. K., Grabe, W., Koda, K., Mosenthal, P., Mulkahy-Ernt, P., & Shedl, M. (2000). *TOEFL 2000 Reading Framework: A Working Paper*. (TOEFL monograph No. 17). Princeton, NJ: Educational Testing system.
- ETS. (2009). *The official guide to the TOEFL Test* (3<sup>rd</sup> ed.). New York, NY: McGraw-Hill
- ETS. (2011). Reliability and comparability of TOEFL IBT scores. *TOEFL Research Insight Series*. Vol.3.
- Grabe, W. & Stoller, F. L. (2013). *Teaching and researching reading* (2<sup>nd</sup> ed.). New York, NY: Routledge.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.
- Hudson, T. (1996). *Assessing second language academic reading from a communicative perspective: Relevance for TOEFL 2000*. (TOEFL monograph No. 17). Princeton, NJ: Educational Testing System.

- Hughes, A. (1989). *Testing for language teaching*. New York, NY: Cambridge University Press.
- Jamieson, J. Jones, S., Kirsch, I., Mosenthal, P. & Taylor, C. (2000). *TOEFL 2000 framework: A working paper* (TOEFL monograph No. 16). Princeton, NJ: Educational Testing System.
- Jamieson, J. M., Eigner, D., Grabe W., & Kunnan, A. J. (2008). *Frameworks for a new TOEFL*. In C. A. Chapelle, M. K. Enright, J. M. Jamieson (Eds.). *Building a validity argument for the test of English as a foreign language* (pp. 1-25). New York, NY: Routledge.
- Lado, R. (1961). *Language testing*. London, England: Longman.
- Linacre, J. M. (2009). *WINSTEPS Rasch Measurement* computer program. Chicago, IL: Winsteps.
- Liu, O. L., M. Schedl, J. Malloy, & N. Kong. (2009). *Does content knowledge affect TOEFL iBT reading performance? A confirmatory approach to differential item functioning* (ETS Research Rep. No. TOEFLiBT-09). Princeton, NJ: Educational Testing System.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessment. *Educational Researcher*, 23(2), 13-23.
- Mousavi, S. A. (2012). *An encyclopedic dictionary of language testing* (5<sup>th</sup> ed.). Tehran, Iran: Rahnama.
- Patricia, C., Gordon, A., Schedl, M., & Tang, L. (1996). *An analysis of the dimensionality of TOEFL reading comprehension items* (ETS Research Rep. TOEFL-RR-53). Princeton, NJ: Educational Testing system.