Research Paper
Developing the KHANA Test to Evaluate Reading Skills in Persian-speaking Students: A Preliminary Study

Talieh Zarifian[1] , **Atieh Ashtari[1*]** , **Reza Nilipour[1]** , **Shahin Nematzadeh[2]** , **Narges Bayat[1,3]**

1. Department of Speech Therapy, School of Rehabilitation Sciences, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
2. Department of Linguistics, Faculty of Literature, Al Zahra University, Tehran, Iran.
3. Department of Health Sciences, School of Rehabilitation Sciences, University of Ottawa, Ottawa, Canada.

ABSTRACT

Objectives: Reading impairments are the most prevalent problem at school. It is vital to have a valid and reliable test to assess the student's reading skills. This study aims to develop the KHANA test as a valid and reliable test to evaluate reading skills in Persian-speaking students and study the accuracy, rate, and reading comprehension in different grades and genders

Methods: This study included 87 students from second to seventh grade (75 typically developing students [TDS] and 12 students with reading impairments [SRI]). First, two preliminary parallel forms of student books, A and B, were prepared, both including 12 reading passages organized based on increasing difficulty as the student moved from one passage to the next. All the passages contain five comprehension questions. Then, the content and face validity, descriptive statistics, construct validity, and reliability, including test re-test, inter-rater, and the correlation between passages A and B were measured.

Results: Based on the results, all the passages and questions were valid. The passages can make significant differences in reading skills between different grades, while gender differences are limited. However, according to the test re-test analysis, the inter-rater reliability, and the correlation between the two passages, KHANA is a reliable tool and can discriminate SRI from TDS.

Discussion: According to the results, KHANA seems a valid and reliable test to assess reading skills in Persian-speaking students from second to seventh grade.

* Corresponding Author:
Atieh Ashtari, *PhD.*
Address: Department of Speech Therapy, School of Rehabilitation Sciences, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
Tel: +98 (21) 222180043
E-mail: at.ashtari@uswr.ac.ir

Highlights

• KHANA seems to be a valid and reliable test to assess reading skills (rate, accuracy, and reading comprehension) in Persian-speaking students from second to seventh grade.

• KHANA can discriminate students with reading impairments (SRI) from typically developing students.

Plain Language Summary

Some students have difficulty in reading. This difficulty can be in the rate, accuracy, or comprehension of reading. By identifying students with reading impairments (SRI) early, they can receive appropriate treatment that prevents problems later in life. Therefore, we developed the KHANA test to assess reading skills in Persian-speaking students from second to seventh grade. We implemented this test on students with and without reading impairments, and according to the results, KHANA is a suitable tool to assess reading skills among students.

Introduction

Being literate improves individuals' quality of life since it allows them to pursue their studies and careers [1]. These days, the mainstream of human communication is via reading and writing. The technological progress also convinced individuals to use electronic devices more, which has become more crucial [2].

Reading is a complex cognitive ability and is necessary for language learning and communication [3]. This ability includes three skills, accuracy, rate, and reading comprehension. Accuracy of reading refers to the number of words that are correctly produced while reading aloud. The reading rate is related to the reading time, and reading comprehension is related to the number of questions that the readers answer correctly [4]. Any problems associated with reading will inevitably lead to difficult written communication and knowledge acquisition [5].

Reading disorder is the most prevalent type of learning disorder and is estimated to occur in approximately 5% to 12% of school-age children [6]. Students with reading impairments (SRI) have problems with accurate or fluent word recognition and spelling, as well as decoding, despite enough instruction, normal intelligence, and intact sensory abilities [7, 8]. Difficulties in educational skills, such as listening, reading, writing, mathematics, and problem-solving can affect children's rate, accuracy, and comprehension skills, as well as academic performance and social communication [9]. SRIs face negative consequences, such as challenges in continuing their education and choosing a suitable job [1]. Therefore, it affects their social participation and quality of life [10].

Over the past years, a considerable number of studies have assessed children's ability to learn reading and the challenges that they may face in learning this skill. These researches shed light on the nature of reading impairments as well as the assessment methods [3].

Tests are available to assess reading impairments, such as Woodcock reading mastery tests-revised (WRMTR) [11], the Gray oral reading test (GORT) [12], the qualitative reading inventory-5 (QRI-5) [13], the reading and dyslexia test (NEMA) [14], the diagnostic reading test [15]. These tests represent different aspects of the problem as well as the strengths and weaknesses that can be used to design and implement effective treatment plans [3]. WRMTR is one such test that evaluates reading ability in individuals from kindergarten to adulthood. It contains various subtests, such as letter and word identification, and word and passage comprehension [11]. GORT is another reading test that evaluates reading abilities (i.e. accuracy, rate, and reading comprehension) using silent and oral reading. GORT-5 is the fifth edition and a norm-referenced of this test. It has norms for people aged between 6 to 23 years. It is a valid and reliable test that contains 16 story passages followed by 5 comprehension questions [12]. Another test to assess reading skills is QRI-5, which provides information about conditions under which students can detect words and understand text successfully or unsuccessfully. This inventory contains graded word lists and numerous passages designed to assess the oral and silent reading and listening ability of students from the preprimary school through the high school levels [13].

Although several tests exist to assess reading skills in English, they cannot be used in Persian due to different characteristics, such as orthographic and script styles.

Most Persian reading tests do not evaluate all aspects of reading ability and are not up-to-date because they are designed based on the previous version of the course books taught in schools. Also, a great number of Persian reading tests are left unpublished and are not available to use in clinics and research projects [4, 16, 17].

NEMA is the only Persian standardized test to diagnose dyslexia in students from first to fifth grades [14]. This test does not evaluate the rate and accuracy of reading in the passage and the guidelines for performing the tasks are not clear. The diagnostic reading test is another Persian tool to screen and diagnose reading impairments as well as assess spelling skills in second-grade students. No cut-off scores were calculated for this test and the passages designed for evaluation are very easy for students. Moreover, the test was designed based on the old version of the course books (significant changes were observed in Persian course books in recent years) [15].

As mentioned above, reading impairments hurt children's lives because they deter students from succeeding at school. Early diagnosis of this impairment results in early treatment, which is more effective than treatment in later years.

While various tests exist to assess reading in the world [12, 18, 19], Persian tests are limited and out of date. However, among these few tests, only one of them has been published (NEMA) [14], and the rest are merely available as final reports in the research and educational departments. None of these Persian reading tests are designed based on the current course books; therefore, this study aims to develop a valid and reliable KHANA test to evaluate reading skills in Persian-speaking students and study accuracy, rate, and reading comprehension in different grades and genders. This study reports the preliminary stage of developing a comprehensive test to evaluate reading ability in Persian students.

Materials and Methods

This study was a methodological study that had two phases. We developed the test and measured its psychometric properties in the first phase. In the second phase, 87 students were included and assessed using the designed test.

First, we prepared two preliminary parallel forms of the student books (A and B), each book includes 12 reading passages that were organized based on increasing difficulty as the student moves from one passage to the next. Any passage contains five comprehension questions.

These two parallel forms provide the opportunity to determine the reliability of reading skills assessment as well as to evaluate the student's performance before and after treatment. This draft was based on course books from the second to seventh grades to assess reading skills (i.e. accuracy, rate, and reading comprehension). We were inspired by reading tests, such as WRMTR [11], GORT-5 [12], QRI-5 [13], NEMA [14], the diagnostic reading test [15], and Shafiei et al.'s test [4] in designing KHANA.

Then, we measured the content and face validity. Also, we determined the descriptive statistics, construct validity, and reliability, including test re-test, inter-rater, and the correlation between passages A and B.

In the current study, 75 typically developing students (TDS) (46 girls and 29 boys) who studied from the second to seventh grade were selected using a convenience sampling method. Twelve SRI (6 girls and 6 boys) also participated. The inclusion criteria for TDS included the second to seventh-grade student in Tehran City or Karaj City, Iran speaking Persian as a dominant language at home, having the normal hearing ability and a non-verbal intelligence quotient (IQ) of more than 85 based on the student health report available at school, having a good reading skill, no language deficit or speech sound disorders based on the examiner's informal evaluation, and no reading or writing problems based on the teachers' report. The exclusion criteria included having symptoms of speech sound disorders, visual problems, symptoms of sensory deficits (even wearing glasses or hearing aids), and symptoms of psychiatric disorders, such as autism spectrum disorders, psycho-physical delays, evident oro-motor deficits, and a history of recurrent middle ear infections, epilepsy, convulsion, syncope, and brain damages based on their medical health report, asking their parents and teachers. All the criteria were true for the SRI, except for having good reading skills. They had problems in reading based on their teachers' reports and or a speech therapist's diagnosis. Table 1 presents the demographic information of the participants.

Procedure

In phase one, an expert panel, including 13 experts (8 speech therapists with experience in reading impairments, 4 linguistics with experience in psycholinguistics and clinical linguistics, and 1 methodologist) received the first draft of the 24 passages and evaluated the content validity based on being well-formed, coherent, having interesting text, having no bias in the text, being suitable for the students, as well as having relevant questions

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

641

Table 1. Demographic information of the participants

Grade	No.	Gender	TDS (%)		SRI (%)	
			Cumulative Frequency	Relative Frequency	Cumulative Frequency	Relative Frequency
2	15	Male	5.3	36.4	8.3	25.0
		Female	9.3	63.64	25.0	75.0
3	14	Male	8	54.5	16.7	66.7
		Female	6.7	45.5	8.3	33.3
4	17	Male	5.3	23.5	0	0
		Female	17.3	76.5	0	0
5	13	Male	8	54.5	8.3	50
		Female	6.7	45.5	8.3	50
6	15	Male	6.7	38.5	8.3	50
		Female	10.7	61.5	8.3	50
7	13	Male	5.3	33.3	8.3	100
		Female	10.7	66.7	0	0
Total	87	Male	38.7	38.7	50	50
		Female	61.3	61.3	50	50

TDS: Typically developing students; SRI: Students with reading impairments.

to the passage. To measure the face validity, 10 teachers with at least five years of teaching experience at primary schools received the passages and questions and were asked to classify passages for each grade, determine the difficulty of the sentences, and comment on whether the reading comprehension questions were appropriate. After reviewing the results, the final version of the test was prepared for the pilot test, in which 10 TDS (4 girls and 6 boys) participated.

A protocol was designed for the examiners that contained information about the test. This information was about how to record the participants' voices while reading the passages, how to give feedback, or when to stop the test (if the participants had more than 10 mistakes in two consecutive passages, the test was stopped. This protocol was taught to the examiners who were final-year undergraduate students and speech therapist graduates. They were completely aware of the test and its aims, as well as its guidelines. The research team (the first and second authors) approved the examiners' eligibility after piloting a group of students. In the second phase, the trained and eligible examiners took the test from the students living in Tehran and Karaj cities for seven months.

The gathered data in the second phase were analyzed using SPSS software, version 24. Since the data did not follow a normal distribution (based on Kolmogorov-Smirnov and Shapiro-Wilk tests), non-parametrical tests were used. Descriptive statistics were used to determine the Mean±SD, minimum and maximum accuracy, rate, and reading comprehension in different grades. In addition to the content validity and face validity, construct validity to determine whether the test can differentiate rate, accuracy, and comprehension skills between the different grades, the TDS group and the SRI group as well as the two genders was measured. To evaluate the reliability, we measured the test re-test reliability (using Spearman's correlation coefficient), the inter-rater reliability (using the intraclass correlation coefficient [ICC]), and the correlation between passages A and B. To evaluate the performance of the two genders and children with and without reading impairments, the Mann-Whitney U test was used. Also, the Kruskal-Wallis's test was used to determine accuracy, rate, and reading comprehension between the different grades.

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

Results

In the current study, 87 students participated, including 75 TDS and 12 SRI. The results of the statistical analysis are described in the following.

The content validity of the A and B student books, each including 12 passages and five questions dedicated to them, was measured qualitatively based on the experts' opinions. Most passages were scored 90 and more and a few of them were scored 70. Those passages which scored 70 were reviewed by the research team. Then, they were sent back to the experts, and they approved the passages.

The face validity of the test was measured by asking 10 teachers to classify the passages and questions suitable for different grades and to determine the sentence difficulty percentage on a visual scale analysis. Based on the teachers' opinions, the stories and questions were sorted according to their difficulty level. Finally, 24 passages (12 for each A and B) were sorted in terms of difficulty.

Discriminant validity of the rate, accuracy, and comprehension in both passages were applied to determine the construct validity of the test. Table 2 and Table 3 present these results. Based on the results presented in Table 2, for passage A, a significant difference was observed between all the grades in the rate of reading. For accuracy, only passages 3 to 8 and for reading comprehension, only passages 2 and 8 demonstrated significant differences.

According to the validity results of discriminating the passages of group B, passages 1 to 9 were significantly different for the rate of reading, passages 4 to 7 and 10 for accuracy, and passages 3 and 11 for reading comprehension (Table 3).

The test's ability was measured to discriminate reading skills between TDS and SRI. According to the results, rate, accuracy, and comprehension skills were significantly better in TDS. Appendix 1 presents the results. Moreover, in detecting differences between boys and girls, no gender differences were found in the mean rate scores. In accuracy, girls had significantly higher scores than boys in passages A1, A3, A4, A7, and B3. Girls also showed better performance in comprehension of passages A3 and B5. Appendix 2 presents the results.

According to Table 4, the test re-test analysis indicated a significant correlation between the reading skills in

Table 2. Discriminant validity of rate, accuracy, and comprehension in passage A

Passages	No.	Rate			Accuracy			Comprehension		
		Mean±SD	X^2 (df)	P	Mean±SD	X^2 (df)	P	Mean±SD	X^2 (df)	P
A1	75	91.81±28.77	5	0.001	20.33±0.91	5	0.534	4.52±074	5	0.542
A2	75	114.51±30.28	5	0.001	42.10±1.25	5	0.269	4.56±0.59	5	0.002
A3	75	121.60±33.29	5	0.001	60.00±2.67	5	0.021	4.27±0.94	5	0.064
A4	75	90.60±28.73	5	0.001	76.67±3.05	5	0.030	4.56±0.68	5	0.175
A5	74	96.96±32.15	5	0.001	107.85±3.95	5	0.020	4.35±0.80	5	0.414
A6	71	90.49±33.96	5	0.001	117.73±4.46	5	0.001	3.35±1.28	5	0.085
A7	63	95.79±32.10	5	0.001	140.28±5.68	5	0.001	4.14±1.30	5	0.155
A8	49	91.40±27.32	5	0.001	125.75±5.02	5	0.001	3.08±1.59	5	0.003
A9	46	101.85±26.89	5	0.010	133.71±3.28	5	0.396	4.02±5.57	5	0.683
A10	39	89.90±23.26	4	0.014	177.20±4.60	4	0.198	2.82±1.18	4	0.197
A11	34	118.59±28.79	4	0.041	266.08±4.30	4	0.259	4.41±0.92	4	0.445
A12	24	118.30±24.20	3	0.020	270.70±20.09	3	0.374	3.66±1.37	3	0.400

Note: Discriminant validity of rate, accuracy, and comprehension in passage A based on Mann–Whitney U.

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

643

Table 3. Discriminant validity of rate, accuracy, and comprehension in B passages

Passages	No.	Rate			Accuracy			Comprehension		
		Mean±SD	X² (df)	P	Mean±SD	X² (df)	P	Mean±SD	X² (df)	P
B1	74	106.49±30.70	5	0.001	22.29±0.94	5	0.655	4.77±0.56	5	0.437
B2	74	92.54±36.52	5	0.001	49.14±2.21	5	0.078	4.21±0.96	5	0.551
B3	74	112.27±34.29	5	0.001	65.48±2.57	5	0.721	4.54±0.72	5	0.057
B4	74	90.18±34.06	5	0.001	78.06±3.42	5	0.001	4.41±0.75	5	0.130
B5	72	94.72±35.07	5	0.001	123.00±4.81	5	0.001	3.70±1.21	5	0.151
B6	64	91.25±33.16	5	0.001	128.40±5.82	5	0.001	3.07±1.31	5	0.060
B7	50	99.63±27.37	5	0.021	130.44±4.33	5	0.002	3.98±1.20	5	0.145
B8	42	93.17±26.12	5	0.008	131.85±4.11	5	0.064	2.61±1.36	5	0.270
B9	34	112.05±26.79	4	0.056	152.55±3.41	4	0.427	4.08±1.05	4	0.189
B10	31	91.82±26.45	4	0.061	178.32±3.53	4	0.026	3.06±1.28	4	0.263
B11	27	115.48±33.55	4	0.093	198.62±4.19	4	0.180	2.74±1.45	4	0.021
B12	22	124.63±28.51	3	0.157	298.36±4.21	3	0.132	4.13±1.28	3	0.435

Iranian **R**ehabilitation **J**ournal

most passages A and B. The correlation coefficient results in Table 5 showed a significant correlation between the independent measurement of the accuracy of passages A and B by two evaluators. Table 6 also indicates that all passages A and B are significantly correlated in terms of rate, accuracy, and comprehension, except for passage 8 (comprehension), passage 9 (comprehension), passage 11 (comprehension), and passage 12 (accuracy).

Discussion

As a basic element for learning, reading is a crucial skill that opens doors to education, employment, and well-being [1]. Reading impairments hurt children; therefore, it is essential to identify SRI [20]. Since a limited number of published Persian tests exist (only one is available, NEMA), most of them are outdated and unavailable [14, 15], this study was conducted to develop a valid and reliable test for assessing reading skills in Persian-speaking students. KHANA contains two preliminary parallel forms of student books (A and B) that can be used to determine the reliability of reading skills assessment and evaluate the performance of the SRI before and after the treatment (however, the latter in the preliminary study was not performed). Also, accuracy, rate, and reading comprehension were measured in different grades and genders.

Based on the content and face validity results, most passages and questions were valid. Those with the lower scores at this stage were reviewed and edited based on the experts' and teachers' opinions.

In construct validity, reading skills were assessed between the students from the second to seventh grade, between TDS and SRI, as well as between the two genders using KHANA. According to the results, the rate of reading was significantly different between all the grades. Based on Ehri, students have slow reading speed in the first three years of school, and this rate increases in higher grades as their reading skills improve [21]. The results are also consistent with some Persian studies indicating that the rate of reading improves from the first to seventh grades [4, 16, 17].

Based on the accuracy results, this reading skill improved from the second to seventh grades in both passages A and B, which is consistent with Ehri's results on reading development from alphabetical to orthographic reading [21]. It seems the student becomes more capable of using reading rules in higher grades. Therefore, we can expect to have higher accuracy scores in older students. It is also consistent with studies conducted by Shafiei et al., Aziziyan et al. and Jabbari et al. [4, 16, 17]. In the present study, some passages, such as pas-

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

Table 4. Test re-test reliability between rate, accuracy, and comprehension

Passages	Rate	Accuracy	Comprehension
A1	0.72**	0.25	0.35*
A2	0.74**	0.55**	0.82**
A3	0.81**	0.64**	0.61**
A4	0.86**	0.88**	0.54**
A5	0.93**	0.69**	0.41*
A6	0.77**	0.65**	0.49*
A7	0.81**	0.68**	0.67**
A8	0.82**	0.69**	0.56*
A9	0.76**	0.55*	0.15
A10	0.70**	0.74**	0.73**
A11	0.82**	0.49*	0.45
A12	0.68*	0.90**	0.70*
B1	0.76*	0.46*	0.83**
B2	0.91**	0.70**	0.74**
B3	0.79**	0.58**	0.54**
B4	0.89**	0.18	0.38
B5	0.94**	0.72**	0.38
B6	0.92**	0.58**	0.59**
B7	0.92**	0.73**	0.65*
B8	0.92**	0.73**	0.65*
B9	0.87**	0.57*	0.85**
B10	0.87**	0.29	0.22
B11	0.47	0.43	0.50*
B12	0.77*	0.60*	0.80*

P<0.001, *P<0.05.   **Iranian **R**ehabilitation **J**ournal

sages A1 and A2, could not make a significant difference since they both evaluate the basic level of reading. These results confirm the construct validity of the test.

Only a limited number of passages can make significant differences between the reading comprehension skills at different grades. Assessing reading comprehension is a challenge in most reading tests in foreign and Persian tests. This problem is evident in NEMA [14] and the diagnostic reading test [15]. Also, a great number of changes were made in the comprehension section of different versions of Gort [12]. It seems that the reason for this challenge is that the questions evaluate objective issues and the students did not have to deduct the whole passage for a correct answer.

In comparing the female and male students' reading rates, no significant differences were observed between the two genders; however, based on Shafiei et al., Aziziyan et al. and Shirazi, girls demonstrated better

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

645

Table 5. Inter-rater reliability of accuracy of passages A and B by two evaluators

Passages	1	2	3	4	5	6	7	8	9	10	11	12
A	0.90**	0.90**	0.92**	0.87**	0.81*	0.87**	0.88**	0.80*	0.67**	0.80**	0.62**	0.17
B	0.86**	0.34*	0.89**	0.79**	0.71**	0.22	0.83**	0.71**	0.31	0.57*	0.76**	0.70*

**P<0.001, *P<0.05.   Iranian Rehabilitation Journal

performance than boys [4, 16, 22]. Reading accuracy in a few numbers of the passages (such as A1, A3, and A7) was significantly better in girls. According to Shafiei et al. and Aziziyan et al., girls are better at accuracy than boys [4, 16]. Girls also had significantly higher scores in reading comprehension only in passages A3 and B3. The difference between the results may be due to the greater number of female students who participated in the current study.

We also compared the SRI and TDS reading skills. According to the results, TDS performed significantly better compared to SRI; therefore, KHANA can discriminate between these two groups. WRMTR [11] and GORT-5 [12] are other tests that have diagnostic capability.

According to evaluation of the passages A and B, the results showed a highly significant correlation between the reading skills in most passages A and B. However, it is preferred to have parallel passages that are correlated. Four passages A and B were not highly correlated; there-

fore, it is necessary to revise these texts (passages 8, 9, 11, 12).

Moreover, according to the test re-test reliability results, an acceptable correlation was observed between the two evaluators in most administrations; therefore, KHANA is a reliable tool.

Conclusion

In conclusion, the development of the KHANA test facilitates assessing reading skills in Persian-speaking students. The study successfully created two preliminary parallel forms of student books (A and B) to evaluate the accuracy, rate, and comprehension in different grades and genders. The validity assessment, including content, face, and construct validity, demonstrates that the passages and questions are well-constructed and measure the intended skills effectively. Based on the test re-test analysis, the inter-rater reliability, and the correlation between the two passages, KHANA is a reliable tool and

Table 6. Correlation between the reading skills between passages A and B

Passages	Rate	Accuracy	Comprehension
1	0.64**	0.59**	0.63**
2	0.74**	0.69**	0.50**
3	0.79**	0.75**	0.61**
4	0.74**	0.66**	0.43**
5	0.79**	0.55**	0.45**
6	0.66**	0.36**	0.30*
7	0.66**	0.52**	0.45**
8	0.83**	0.55**	0.35*
9	0.86**	0.63**	0.22
10	0.88**	0.57**	0.54**
11	0.84**	0.66**	0.34
12	0.88**	0.26	0.72**

*P<0.05, **P<0.001.   Iranian Rehabilitation Journal

is capable of discriminating between SRI and TDS. In summary, the KHANA test emerges as a valuable and reliable instrument for evaluating reading skills in Persian-speaking students, with the potential to contribute significantly to educational and clinical settings.

Limitations and recommendations

This study had limitations. The COVID-19 pandemic affects the sample size, especially SRI. Also, the number of girls participating in this study was more than boys. The other limitation was about matching the TDS and SRI groups. They merely matched their grade levels. The preferred way is to have two groups matched in terms of the language testing results or literacy achievements; however, no standard Persian test is found to evaluate language or literacy in students. Therefore, the examiners used informal tests. It is suggested to use a random sampling method with a larger sample size and calculate the specificity and sensitivity, as well as standard scores for the reading skills in each grade in future studies. Furthermore, according to the low correlation between passages A and B and comprehension questions, it is recommended to revise them.

Ethical Considerations

Compliance with ethical guidelines

The study was approved by the Ethics Committee of the University of Social Welfare and Rehabilitation Sciences (Code: IR.USWR.REC.1398.114). All participants and their parents were informed about the test and the procedure and signed the informed consent form.

Authors' contributions

Conceptualization, methodology and supervision: Talieh Zarifian, Atieh Ashtari, Reza Nilipour, Shahin Nematzadeh; Data collection: Talieh Zarifian, Atieh Ashtari, Narges Bayat; Data analysis: Talieh Zarifian and Atieh Ashtari; Investigation and writing: All authors.

Conflict of interest

The authors declared no conflict of interest.

References

[1] Snowling MJ, Hulme C. Annual research review: The nature and classification of reading disorders--a commentary on proposals for DSM-5. Journal of Child Psychology and Psychiatry, and Allied Disciplines. 2012; 53(5):593-607. [DOI:10.1111/j.1469-7610.2011.02495.x] [PMID]

[2] Poe MT. A History of Communications: Media and society from the evolution of speech to the internet. New York: Cambridge University Press; 2010. [Link]

[3] Fletcher JM, Foorman BR, Boudousquie A, Barnes MA, Schatschneider C, Francis DJ. Assessment of reading and learning disabilities a research-based intervention-oriented approach. Journal of School Psychology. 2002; 40(1):27-63. [DOI:10.1016/S0022-4405(01)00093-0]

[4] Shafiei B, Tavakol S, Alinia L, Maracy MR, Sedaghati L, Foroughi R. [Developing a screening inventory reading test (IRT) for the Isfahanian students of the first to fifth grade (Persian)]. Bimonthly Audiology. 2009; 17(2):53-60. [Link]

[5] MacCullagh L, Bosanquet A, Badcock NA. University students with dyslexia: A qualitative exploratory study of learning practices, challenges and strategies. Dyslexia. 2017; 23(1):3-23. [DOI:10.1002/dys.1544] [PMID]

[6] Sedaghati L, Foroughi R, Shafiei B, Maracy MR. [Prevalence of dyslexia in first to fifth grade elementary students Isfahan, Iran (Persian)]. Bimonthly Audiology. 2010; 19(1):94-101. [Link]

[7] Peterson RL, Pennington BF. Developmental dyslexia. Annual Review of Clinical Psychology. 2015; 11:283-307. [DOI:10.1146/annurev-clinpsy-032814-112842] [PMID]

[8] Colenbrander D, Ricketts J, Breadmore HL. Early Identification of dyslexia: Understanding the issues. Language, Speech, and Hearing Services in Schools. 2018; 49(4):817-28. [DOI:10.1044/2018_LSHSS-DYSLC-18-0007] [PMID]

[9] Alexander PA, Judy JE. The interaction of domain-specific and strategic knowledge in academic performance. Review of Educational Research. 1988; 58(4):375-404. [DOI:10.3102/00346543058004375]

[10] Riva S, Antonietti A. The application of the ICF CY model in specific learning difficulties: A case study. Psychology of Language and Communication. 2010; 14(2):37-58. [DOI:10.2478/v10057-010-0009-2]

[11] Eaves RC. Woodcock reading mastery tests-revised (WRMTR). Diagnostique. 1990; 15(1-4):277-97. [DOI:10.1177/15345084890151-425]

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

647

[12] Hall AH, Tannebaum RP. Test review: J. L. Wiederholt & B. R. Bryant. (2012). Gray oral reading tests-fifth edition (GORT-5). Austin, TX: Pro-Ed. Journal of Psychoeducational Assessment. 2013; 31(5):516-20. [DOI:10.1177/0734282912468578]

[13] Leslie L, Caldwell JS. Qualitative reading inventory. Indianapolis: Pearson Education; 2016. [Link]

[14] Moradi A, Hosaini M, Kormi Nouri R, Hassani J, Parhoon H. [Reliability and validity of reading and dyslexia test (NEMA) (Persian)]. Advances in Cognitive Sciences. 2016; 18(1):22-34. [Link]

[15] Sima-Shirazi T, Nili-pour R. [Development and standardization of reading diagnostic test (Persian)]. Archive of Rehabilitation. 1383; 5(1 and 2) :7-11. [Link]

[16] Aziziyan M, Abedi M. [Construction and standardization of the reading level diagnostic test for third grade primary school students (Persian)]. Iranian Journal of Psychiatry and Clinical Psychology. 2005; 11(4):379-87. [Link]

[17] Jabbari S, Khademi M. [Creating a reading and comprehension diagnostic test for elementary school students (Persian)]. Curriculum Studies Journal. 2014; 3(2):33-51. [DOI:10.22099/jcr.2014.2970]

[18] Neale MD. Neale analysis of reading ability. Adelaide: Australian Council for Educational Research; 1999. [Link]

[19] Cox MM. The relationship of speech and reading in an elementary school program. Communication Education. 1959; 8(3):211-8. [DOI:10.1080/03634525909377021]

[20] Nergård-Nilssen T, Eklund K. Evaluation of the psychometric properties of "the Norwegian screening test for dyslexia". Dyslexia. 2018; 24(3):250-62. [DOI:10.1002/dys.1577] [PMID]

[21] Ehri LC. Learning to read words: Theory, findings, and issues. Scientific Studies of reading. 2005; 9(2):167-88. [DOI:10.1207/s1532799xssr0902_4]

[22] Shirazi S. Phonological Awareness and its implications for Reading Acquisition. Iranian Rehabilitation Journal. 2006; 4(1):40-4. [Link]

648

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

Appendix 1. Discriminate validity of rate, accuracy, and comprehension between TDS and SRI

Passages	Reading Skills	Statistics			Power	η²
		U	Z	P		
A1	Rate	179.00	-2.68	0.008	0.54	0.35
	Accuracy	167.00	-4.04	0.001	0.99	0.21
	Comprehension	166.50	-4.22	0.001	1.00	0.30
A2	Rate	159.00	-3.84	0.001	0.97	0.15
	Accuracy	238.50	-3.07	0.002	0.95	0.13
	Comprehension	174.00	-4.09	0.001	1.00	0.33
A3	Rate	257.00	-3.18	0.001	0.89	0.11
	Accuracy	226.50	-3.10	0.001	0.98	0.16
	Comprehension	142.50	-4.22	0.001	1.00	0.30
A4	Rate	213.00	-2.66	0.005	0.69	0.70
	Accuracy	42.00	-4.80	0.001	1.00	0.27
	Comprehension	194.00	-3.21	0.001	0.99	0.22
A5	Rate	194.50	-2.78	0.024	0.69	0.07
	Accuracy	58.00	-2.36	0.018	1.00	0.31
	Comprehension	47.00	-4.51	0.001	1.00	0.33
A6	Rate	81.00	-0.69	0.485	0.13	0.01
	Accuracy	20.50	-1.19	0.232	0.88	0.12
	Comprehension	53.50	-1.50	0.132	0.26	0.02
A7	Rate	19.50	-0.65	0.516	0.07	0.003
	Accuracy	-	-	-	-	-
	Comprehension	10.50	-1.26	0.207	0.13	0.01
B1	Rate	195.50	-3.41	0.001	0.84	0.09
	Accuracy	224.00	-3.30	0.001	0.96	0.14
	Comprehension	182.00	-4.74	0.001	1.00	0.43
B2	Rate	213.00	-3.19	0.001	0.77	0.08
	Accuracy	109.00	-4.47	0.001	0.99	0.22
	Comprehension	114.00	-4.61	0.001	1.00	0.43
B3	Rate	228.50	-3.00	0.003	0.79	0.08
	Accuracy	143.50	-4.04	0.001	0.99	0.23
	Comprehension	87.00	-5.21	0.001	1.00	0.52

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

649

Passages	Reading Skills	Statistics			Power	η^2
		U	Z	P		
B4	Rate	202.00	-1.46	0.142	0.31	0.02
	Accuracy	102.00	-3.05	0.002	0.98	0.17
	Comprehension	107.00	-3.24	0.001	0.99	0.21
B5	Rate	190.50	-1.56	0.118	0.35	0.03
	Accuracy	88.50	-3.21	0.001	0.96	0.15
	Comprehension	70.00	-3.62	0.001	0.99	0.19
B6	Rate	76.50	-0.59	0.554	0.08	0.005
	Accuracy	39.50	-1.71	0.086	0.22	0.002
	Comprehension	46.50	-1.53	0.124	0.40	0.04
B7	Rate	48.00	-0.09	0.924	0.05	0.001
	Accuracy	-	-	-	-	-
	Comprehension	-	-	-	-	-

U: Mann-Whitney U test; Z: Z score.

Iranian Rehabilitation Journal

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

Appendix 2. Discriminant validity of rate, accuracy, and comprehension between girls and boys

Passages	Reading Skills	Statistics		
		U	Z	P
A1	Rate	839.00	-0.61	0.538
	Accuracy	675.50	-2.19	0.028
	Comprehension	721.50	-1.84	0.066
A2	Rate	745.50	-1.42	0.154
	Accuracy	712.50	-1.82	0.068
	Comprehension	866.50	-0.42	0.673
A3	Rate	752.00	-1.36	0.171
	Accuracy	622.00	-2.55	0.011
	Comprehension	695.00	-1.98	0.048
A4	Rate	715.00	-1.36	0.172
	Accuracy	627.50	-2.15	0.031
	Comprehension	741.50	-1.29	0.195
A5	Rate	710.00	-1.07	0.284
	Accuracy	650.00	-1.63	0.102
	Comprehension	729.50	-0.95	0.339
A6	Rate	614.50	-0.22	0.822
	Accuracy	555.50	-0.89	0.373
	Comprehension	607.00	-0.30	0.749
A7	Rate	423.00	-0.79	0.429
	Accuracy	324.50	-2.16	0.031
	Comprehension	463.00	-0.26	0.793
A8	Rate	216.50	-0.05	0.957
	Accuracy	215.00	-1.04	0.294
	Comprehension	250.00	-0.29	0.768
A9	Rate	213.50	-0.44	0.656
	Accuracy	181.00	-1.21	0.225
	Comprehension	196.00	-0.87	0.380
A10	Rate	152.00	-0.30	0.761
	Accuracy	154.00	-0.22	0.807
	Comprehension	140.00	-0.69	0.489

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

651

Passages	Reading Skills	Statistics		
		U	Z	P
A11	Rate	96.00	-0.90	0.364
	Accuracy	71.00	-1.85	0.063
	Comprehension	83.00	-1.64	0.100
A 12	Rate	49.00	-0.66	0.505
	Accuracy	52.50	-0.44	0.655
	Comprehension	53.00	-0.42	0.668
B1	Rate	848.00	-0.53	0.590
	Accuracy	869.50	-0.37	0.705
	Comprehension	783.00	-1.46	0.142
B2	Rate	838.00	-0.62	0.533
	Accuracy	829.50	-0.70	0.481
	Comprehension	898.50	-0.10	0.916
B3	Rate	868.00	-0.36	0.716
	Accuracy	675.00	-2.04	0.040
	Comprehension	847.00	-0.60	0.544
B4	Rate	785.00	-0.14	0.887
	Accuracy	673.50	-1.21	0.226
	Comprehension	784.50	-0.16	0.871
B5	Rate	733.00	-0.16	0.866
	Accuracy	653.00	-0.96	0.334
	Comprehension	559.50	-1.96	0.050
B6	Rate	495.00	-0.27	0.784
	Accuracy	422.50	-1.22	0.220
	Comprehension	452.50	-0.85	0.395
B7	Rate	260.00	-0.73	0.464
	Accuracy	269.50	-0.54	0.584
	Comprehension	296.00	-0.58	0.558
B8	Rate	163.50	-0.68	0.496
	Accuracy	172.00	-0.45	0.652
	Comprehension	132.50	-1.56	0.118

Passages	Reading Skills	Statistics		
		U	Z	P
B9	Rate	121.50	-0.18	0.854
	Accuracy	98.50	-1.03	0.229
	Comprehension	98.00	-1.12	0.263
B10	Rate	70.00	-1.26	0.207
	Accuracy	90.00	-0.39	0.694
	Comprehension	64.00	-1.59	0.111
B11	Rate	62.00	-0.74	0.457
	Accuracy	58.50	-0.93	0.351
	Comprehension	71.00	-0.27	0.786
B12	Rate	48.00	0.001	1.000
	Accuracy	39.50	-0.63	0.527
	Comprehension	44.50	-0.28	0.776

Abbreviations: U, Mann-Whitney U test; Z, Z score.

Iranian **R**ehabilitation **J**ournal

Zarifian T, et al. Developing the KHANA Test. IRJ. 2023; 21(4):639-654.

653

This Page Intentionally Left Blank