Original Article

Reconstructing, Investigating the Reliability and Validity and Scoring the Stanford Diagnostic Reading Test

Bahram Saleh Sedghpour[1], PhD; Sara Daazi
Shahid Rajaei Teacher Training University, Tehran, Iran

Aim: The aim of the present study was to reconstruct determining validity, and score The Stanford Diagnostic Reading Test fourth edition (SDRT4) in the sixth grade students.

Methods: The population of the study was all sixth grades of the 19 educational districts from Tehran, 571 students (255 boys and 316 girls) were selected by using a random multi-cluster sampling. The data were analyzed. The techniques were item analysis (difficulty index, discriminative index, and loop techniques). Validity, translation validity, content validity, and construct validity (factorial analysis), and reliability (Kuder-Richardson)

Results: the exploratory factor analysis determined five factors: declarative knowledge, inferential knowledge, procedural knowledge and visualization knowledge. The reliability of the Stanford diagnostic Reading Test's subtests by computing the Kuder-Richardson coefficient were 0.778, 0.732 and 0.748 for comprehension subtest, vocabulary subtest and scanning subtest in order.

Conclusion: By considering the results of present study, SDRT4 has good reliability and validity and can appropriately diagnose the reading disabled students in the sixth grade.

Ker words: diagnostic reading test, reading disability, reconstruction, scoring, validity, reliability

Submitted: 22 Nov 2011
Accepted: 10 Feb 2012

Introduction

Reading skill is necessary both to success in school and for the economic survival in the technological world. (1) The ability to read is combined to cognitive skills (mental skill involved in learning, thinking and information processing) to a large extent which requires decoding the words and understanding printed materials that mean the capability of understanding what has been read. (2)

For many people reading is successful but this is not the same for 15 to 20% of the entire population. (3) This group of people who have difficulties in the speed and accuracy of word decoding and understanding are diagnosed as people with reading disorders. (4) Reading disorder is the most common type of learning disorders that is common in 80% of those whose learning disorders are diagnosed.(5) In DSM-IV-TR (American Psychiatry Association, 2000) reading disorder defined as:

Reading progress, that is measured by standard individual reading tests of fluency or understanding. Basically the measured level of intelligence is lower than the chronological age and the age-appropriate education.

This skill deficiency should be considerably interfered with the academic achievement or activities of daily living requiring reading skills.

If a Sensory defect is present, there should be more reading problems than the expected reading problems due to that single defect.

Many researchers and scientists agreed on the etiology of reading disorder to be considered as a neuro-cognitive disorder with a genetic origin. (6)The results of Wilkins et al (2001) investigation about the role of deficiencies in temporal and visual processing (7), Austin (2008) work on the role of deficiencies in Magnocellular Visual System(8), and Fawcett and Nicolson (2008) investigation about the role of the Procedural timing defects (9) show the above agreement. On the other hand, different studies have confirmed the disorder is hereditary; for example Shaywitz and his colleagues' studies (2001) showed that 23% to 65% children, whose one parent has reading disorder, will

1- All Correspondance to : Bahram Saleh Sedghpour; Email:<bahramsaleh@gmail.com>

be affected with this disorder. The association of reading disorder in identical twins (84-100%) is considerably higher than non identical ones(20-35%).(11) Almost 30 to 50 percent of all the children that one of their parents has reading disorder, will have this disorder. Relative risk of this occurrence is almost four to eight times more than what is seen in the control group. (12)

Unfortunately, reading disorder will not disappear over time. The longitudinal studies showed that this is a consistent disorder and cannot be considered as a transient developmental lag. This means that all the children diagnosed with reading disorder will probably have difficulties in reading as adults, too.(4)

Shaywitz (2003) has concluded that all group of readers including good and/or bad readers will develop their reading skills over time, but their skill gap still remains and will not disappear.(13)

Previous studies on reading skills showed that children with reading disorder are confronting so many problems such as: self hesitancy, low self confidence, loneliness , humiliated by peers,(14) emotional disorders like anxiety, depression, lack of self-esteem, Social Incompatibility, (15) self-contradictory and Inappropriate Social behavior and cognition.(16)

Considering the negative consequences of weakness in reading skill, It is very important to distinguish the children with this disorder and provide them adequate treatment on time. There are many studies about the advantages of timely diagnosis and treatment. Denton and Mathes (2003) said that early intervention can make a strong base for successful academic achievements of the student in future (17) Torgesen (2002a,b and 2004) has confirmed the positive effect of early intervention.(18-20)

It has been estimated that appropriate interventions in the class can decrease the risk of reading disorder from 25% to 6% in student population. Meanwhile studies suggested that brain functioning , appeared in the MRI studies of different people with reading disorder, will improve in the individuals who receive and follow treatments to the extent that their brain functioning would be similar to the normal readers.(21)

Regarding to the considerable number of Iranian students with learning disorders especially reading disorder (4-12%) (22), There is an obvious need for a sufficient tool to diagnose the reading disorder in students. Because prompt identification of this group of students can provide enough time and space for their administration, teachers and education professionals to develop appropriate educational programs and/or change in the existing education plan when necessary.

One of the worldwide most common tests to identify the children with reading disorder is Stanford Diagnostic Reading Test. The main aim of the present study was to reconstruct, investigate the reliability and validity and scoring this test for the sixth grade Iranian students. So the main questions of the study are:

1- How will the Stanford Diagnostic Reading Test questions be analyzed?
2- Is Stanford Diagnostic Reading Test reliable enough to identify the students with reading disorder studying in sixth grade?
3- Is Stanford Diagnostic Reading Test valid to identify the students with reading disorder studying in sixth grade?
4- How is the scoring of Stanford Diagnostic Reading Test?

Method

This is a descriptive study aimed to describe the characteristics of a situation or a subject in a subjective, real and systematic way. On the other hand, this implementation will be only useful for a better understanding of the current situation or to facilitate the decision making process.(23)

The research population consists of all the students studying in sixth grade of normal guidance schools in Tehran province during school year 2010-2011. The sampling method was multistage randomized clustering so that 4 regions were selected randomly among 22 regions of Tehran, 3 schools randomly in each region and 2 classes randomly in each school. 26 students in the chosen classes were selected randomly to be tested.(table 1)

Table 1. Students in the chosen classes

School region	1st school		2nd school		3rd school	
	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2
1	26	26	26	26	26	26
3	26	26	26	26	26	26
8	26	26	26	26	26	26
14	26	26	26	26	26	26

Sample size were calculated using the formula ($n=(s^2\times z^2)/d^2$) to estimate the sample size from the population in 95% confidence interval to be 624 students. (24) In the end 571 questionnaires have

been analyzed and 53 incomplete ones were excluded.

Stanford Diagnostic Reading Test has been used in this study. This test has three subscales; Comprehension, Vocabulary and scanning that respectively has 34, 24 and 22 questions. The participants should choose the right answer among four options for each question. There are nine fun, practical and informative texts in the Comprehension subscale that the students should mark the right answer of the multiple choice questions related to those texts. The item scoring will be 0 for the wrong and 1 to the right answers. So the highest score in this subscale is 34 and the lowest is 0. In the Vocabulary subscale, the participants should find the synonym to the underlined word among 4 given options. This subscale scoring is similar to the previous one so the score range of this subscale will be between 24 and 0. There are two texts in the Scanning subscale that have some questions proceeding to it. The participants must determine which question belongs to which paragraph then read that paragraph and find the right answer. In this part all of the text will not be read and the scoring is exactly the same as previous and ranged between 22 and 0.

Results

The aim of test question analysis is evaluating all the questions and determining the amount of their accuracy and shortcomings that shows their strong and weak points.(25) To respond to the first research question, all the test questions were analyzed using difficulty index, coefficient of determination and Loop method. The difficulty index (DI) is the percent of all students that marked the correct answer to the questions. The DIs between 0.3-0.7 are the most informative ones about the participants' differences. Coefficient of Determination (CD) shows the question potential of distinguishing between strong and weak participants. It is obvious that the greater the CD is, the potential determination is more and vice versa.(25)

Loop method will measure the total reliability of the test and the internal coordination of the questions. To do this, first the total reliability has to be determined and then its changes will be measured in case of excluding each question.

If excluding a question will increase the reliability coefficient, this question is not coordinated with the others and if omitting an item intend to decrease of reliability coefficient, that item is well coordinated with the others. But if excluding a question does not change the reliability coefficient, shows that although this question is not significant in coordination to the other questions but it can be considered as an item for practicing of the respondents.

Based on the analysis of the Comprehension subscale questions, items 54, 52, 51, 50, 49, 43, 42, 41, 37, 36, 35, 32, 31, 28, 21, 19, 18, 10 ,9 were excluded, In Vocabulary subscale the items 29,21, 19, 18, 14, 11 were excluded and in the Scanning subscale, questions 19, 18, 17, 16, 15, 9, 8, 1 were excluded because the omitted items increased the reliability coefficient of the other questions in their subscale.

To respond to the second research question, the construct, content, translation and face validity were determined.

Translation validity: In order to assess the translation accuracy of the tool, we asked some professional translators[1]for back translation the tool we have translated to Persian, into English. Then we have asked some other professionals to compare and determine if the back translation and original text are concordant.

Face validity: We have asked some professionals to confirm the face validity and the form of the questions to be adequate for evaluating the reading skill of the students in sixth grade.

Content validity: Some professional opinions were asked to confirm the content validity and determine the adequacy and goodness of the test content regarding the subject matter.

We have used explorative factor analysis to identify the number of important and significant factors in the comprehension subscale. Varimax and Oblimin Rotations were used to determine if the factors are correlated of dependent to each other. In Oblimin method there was no rotation but the Varimax rotation existed so the explorative factor analysis were done using the the original items and the Varimax rotation. Kaiser-Meyer-Olkin (KMO) Measure was used to verify the sampling adequacy and Bartlett's test of sphericity has confirmed that there is a correlation among the test items in the population. The measured Kaiser-Meyer-Olkin in

1- Qasem zadeh, Dordi negad and Nejati, all PhD and faculty members of Foreign Language department in Shahid Rajayee University

this research was 0.828 that in comparison to the accepted amount (0.6) shows that the research sample was adequate to do the factor analysis (Table 2). The Bartlett's test results was significant in 0.99 confidence interval that confirms the enough correlation among the questions of Comprehension Subscale of Stanford Diagnostic Reading Test and it can be analyzed by factor analysis method,

Table 2. KMO and Bartlett's test of sphericity of the comprehension subscale of SDRT

Sig	df	χ^2 Bartlett's test of sphericity	KMO
P<0.001	561	2078	0.828

The first item called "Real Knowledge" which means the Knowledge about the world that can be represented as a true conscious knowing and encompasses questions 6, 8, 12, 2, 22, 1, 4, 20. The other factor "Deductive Knowledge" that is the knowledge related to the logical judgment based on the evidences, previous judgments and results is formed by questions 23, 30, 32, 27, 25, 13, 31, 17. Questions 9, 29 and 16 formed the third factor "Processing Knowledge" which is about how to do things and questions 5, 3, 7 are the components of the "Applied Knowledge" factor that is the knowledge to apply the principles, discoveries and theories in their related practical fields. The "Imaginary Knowledge" where a visual imagination will appear conformed of questions 34, 14 and 26. (Table 3).

Table 3. Factorial Loads of Comprehension Subscale of SDRT

Questions		Factors				
		RK	DK	PK	AK	IK
1-Real Knowledge	6. Earthworms breath through….	0.649				
	8. Why did Arman's father need a spider?	0.646				
	12. At first, Arman thought that his father brought a …	0.605				
	2. This article says that……	0.571				
	22. The word "Raid" in this story means…..	0.490				
	1. When it rains, earthworms come out of the soil so that they can …	0.479				
	4. Regarding to the article what do earthworms eat?	0.473				
	20. This article has been written because…..	0.356				
2-Deductive K.	30. What did Roman do to sweeten the snow flavor?		0.627			
	32. Which of these choices is an opinion in the article?		0.607			
	27. Why the teacher was worried about Farzad?		0/596			
	25. You need the tape to…..		0.592			
	13. Which of these choices were not present at the Poster?		0.480			
	31. The other good title for this article is……		0.461			
	17. The graphite is cooked to ……		0.449			
	23. Dingo went because…….		0.390			
3-Process	9. When this story happened?			0.748		
	29. These boxes show the important events of the story, mark the correct choice for the empty box?			-0.561		
	16. Which one is not an opinion?			0.493		
4-Applied	3. Earthworms are probably helpful more in…..				0.722	
	5. You can say earthworms are more skilled in ……				0.471	
	7. Perhaps the best way to know more about the earthworms is…..				0.456	
5-Imaginary	14. If you want to know more about riding you must ….					0.712
	26. The wood numbers help when someone…..					0.486
	34. You can say from this commercial that….					0.365

The confirmative factor analysis was used to evaluate the goodness of fit of the Comprehension subscale factor structure model by means of Amos software version 18. There were factorial confirmations based on both the type of the text and the comprehension method. The findings of the accepted model of the text were: χ^2 df (1.082), mean of estimated error squares (0.01), Goodness of fit index (0.949), comparative fit index (0.984), adaptive goodness of fit index (0.937), Normative goodness of fit index (0.825); and the findings acceptable for all kinds of comprehension were: χ^2/df (1.084), mean of estimated error squares (0.01), Goodness of fit index (0.95), comparative fit index (0.984), adaptive goodness of fit index (0.95), Normative goodness of fit index (0.829).

Reliability

The reliability of this test was examined through the Kuder-Richardson Method of calculation which represented in table (4). This calculation was in response to the third research question.

Table 4. Reliability coefficients of the subscales of SDRT

Stanford Diagnostic Reading Test Subscales	Reliability Coefficient
Comprehension Subscale	0.778
Vocabulary Subscale	0.732
Scanning Subscale	0.748

Discussion

The main aim of this study was to reconstruct and to investigate the Reliability and Validity and Scoring the Stanford Diagnostic Reading Test for the sixth grade students. This test has three subscales that each has 24, 34 and 22 questions respectively. The test is applicable both individually and in the groups. The time required to implement each of the three subscales is 60, 30 and 25 minutes respectively that takes less for most of the students. The important timing issue is that this time table is restricted and must not be extended because it is a significant factor to diagnose the students who have reading disorder and is not able to answer all the questions in this time line.

Participants should answer to the multiple choice questions about nine fun, practical and informative texts in the comprehension subscale and should choose the right answer among four options for each question. The item scoring will be 0 for the wrong and 1 to the right answers. So the highest score in this subscale is 34 and the lowest is 0. In the Vocabulary subscale, the participants should find the synonym to the underlined word among 4 given options. This subscale scoring is similar to the previous one so the score range of this subscale will be between 24 and 0. There are two texts in the Scanning subscale that have some questions proceeding to it. The participants must determine which question belongs to which paragraph then read that paragraph and find the right answer. In this part all of the text will not be read and the scoring is exactly the same as previous and ranged between 22 and 0.

After scoring, the students' raw scores will be transformed to the T scores (Chart1).

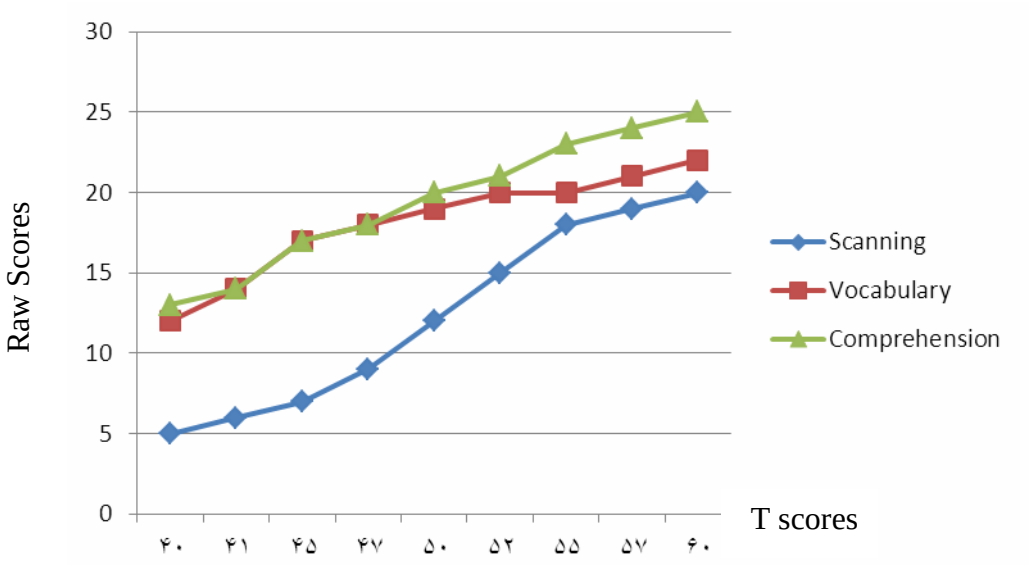

Chart 1. Profile of the SDRT subscales scores of the students in 6th grade.

Students with reading disorders can be categorized based on their raw scores to these three categories:

1- Students with reading disorder: This group of students got 13 or less in Comprehension, 12 or less in vocabulary and 5 or less in Scanning.
2- Students to be considered: The group of students that must undergo the preventive interventions and procedures. The raw scores of this group are between 13-15 in comprehension, 12-14 in vocabulary and 5-7 in scanning.
3- Students to be controlled: The group of students that should periodically be studied and controlled. The raw scores of this group are between 15-20 in comprehension, 14-19 in vocabulary and 7-12 in scanning.

Overall, we can conclude that this test has adequate reliability and validity and its subscale are capable of distinguishing the students who have reading disorders.

References

1. Sofie, C.A. and Riccio C. A. A comparison of multiple methods for the identification of children with reading disabilities. Journal of Learning Disabilities. 2002; 35:234-244
2. Brenitz, Z. The origin of dyslexia: the asynchrony phenomenon. In: G. Reid, A. Fawcett, F. Manis and L. Siegel (eds), the Sage Dyslexia Handbook. 2008. London, Sage.
3. Paquette, P. H. and Tuttle C. G. Learning Disabilities, It happened to me, No.1. 2003.The Scarecrow press, Inc.
4. Siegel I.S. Perspectives on dyslexia. Paediatr Child Health. 2006; 11(9):581-587
5. Meyer, M. S. The ability achievement discrepancy: does it contribute to an understanding of learning disabilities? Educational Psychology Review. 2000. 12:315-337
6. Wilkins, A. J. Lewis, E. Smith, F, Rowland, E., Tweedie, W. Coloured overlays and their benefit for reading. Journal of Research in Reading. 2001; 24(1):4
7. Stein, J. The neurobiological basis of dyslexia.In: G. Reid, A.Fawcett, F.Manis and L. Siegel (eds), The sage dyslexia Handbook. 2008. London. Sage.
8. Fawcett, A. J. and Nicolson, R. Dyslexia and the cerebellum, In:G. Reid, A.Fawcett, F.Manis and L. Siegel (eds), The sage dyslexia Handbook. 2008. London. Sage.
9. Shaywitz, B. A. Shaywitz, S. A. Pugha, K. R. and et al. The neurobiology of dyslexia. Clinical Neuroscience Research. 2001; 1: 291-299
10. Pennington, B. F. Diagnosing Learning Disorders: A neuropsychological Framework. (2nd ed.) 2009.The Guilford press, New York, London
11. Shaywitz, S. Qvercoming Dyslexia: A New and Complete Science-Based Program for Reading Problems at any level. 2003. New York: Alfred A. Knopf
12. Tabrizi, Mostafa. Reading Disorders Treatment. 2009. FaraRavan, Tehran.
13. Baezzat, Fereshteh and Saleh Sedghpour, Bahram. The comparison of self-esteem factors among students with dyslexia and normal students in third to fifth grades of primary schools of Tehran. Motaleaate Ravanshenakhti. 2008. 4(4):65-80
14. Bryan, T. Burstein, K and Ergul, C. The Social Emotional side of Learning Disabilities: A Science based presentation of the state of Art. Learning disability Quarterly. 2004. 27(1):45-51
15. Denton, C. and Mathes, P. Intervention for struggling readers: Possibilities and challenges. In B Foorman (ed).Preventing and remediating reading difficulties bringing science to scale. Baltimore, MD: York Press. 2003.pp: 229-252
16. Torgensen, J. K. Lessons learned from intervention research in reading: A way to go before we rest. In R. Stanhope (ed). Learning and teaching reading. London: British Psychological Society.2002a. 1: 89-103
17. Torgensen, J. K. The prevention of reading difficulties. Journal of School Psychology.2002b. 40: 7-27
18. Torgensen, J. K. Lessons learned from research on interventions for students who have difficulty learning to read . In V.P.C. Mc Cardle (ed). The voice of evidence in reading research.2004. Baltimore, MD: Paul Brooks; pp:355-382
19. Shaywitz, S. E. and Shaywitz, B.A. Reading disability and the brain. Educational Leadership. 2004; 61(6):6-11
20. Tabrizi, Mostafa. Dictation Disorders Treatment. 2008. FaraRavan. Tehran
21. Sarmad, Zohre; Bazargan, Abbas and Hejazi, Elahe. Behavioral Sciences Research Methods. 2007. Agah, Tehran
22. Delavar, Ali. Theoritical and Practical Fundamentals of Research in Human Science and Social Science. 2009. Roshd. Tehran.
23. Seif, Ali Akbar. Educational Measurement , Assessment and Evaluation. (5th edition). 2008. Doran. Tehran