

Validity and Reliability of a Language Development Scale for Persian-speaking Children Aged 2-6 Years

Fatemeh Sheibani¹, PhD; Zahra Sadat Ghoreishi¹, PhD; Reza Nilipour¹, PhD; Abbas Pourshahbaz², PhD; Sakineh Mohammad Zamani¹, MSc

¹Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran; ²Department of Psychology, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

Correspondence:
Reza Nilipour, PhD;
University of Social Welfare and Rehabilitation Sciences, Kodakyar Ave., Daneshjoo Blvd., Evin, Postal code: 19857-13834, Tehran, Iran
Tel: +98 21 22180043
Fax: +98 21 22180109
Email: gsheibani@yahoo.com
Received: 28 August 2018
Revised: 10 October 2018
Accepted: 14 October 2018

What's Known

- Language assessment tools currently available are either parent-report instruments or only evaluate a specific aspect of language. These tools do not cover children under 4 years old.
- There are no comprehensive and validated developmental language screening tests for Persian-speaking children in Iran.

What's New

- For the first time in Iran, a valid and reliable language development scale is developed to assess receptive and expressive language skills of children aged 2-6 years.
- The Persian Language Development Scale (PLDS) utilizes real objects and toys to assess young children; most importantly, children under 4 years old.

Abstract

Background: Language disorders may affect receptive and/or expressive language skills. The use of a validated and reliable assessment tool is essential to assess these skills in children. The present study aimed to develop a valid and reliable language development instrument for Persian-speaking children aged 2-6 years.

Methods: The present cross-sectional study was conducted during 2016-2017 in three main Iranian cities, namely Mashhad, Tehran, and Isfahan. The target population was children between the ages of 2 and 6 in various kindergartens and schools. The Persian Language Development Scale (PLDS) was developed by incorporating linguistic characteristics of the Persian language and Iranian culture. Following a number of iterations, including a pilot study of 36 children, the final version of the PLDS tool was used to assess the receptive and expressive language skills of 460 children. The reliability and validity of the PLDS tool were examined.

Results: The content validity ratio (CRV) of the PLDS tool was 0.85. The tool could differentiate children by age, but not by sex. The test-retest reliability, with 10 days interval, showed a significant correlation between the coefficients of receptive (0.96) and expressive (0.93) scales. The intraclass correlation coefficient (ICC) for receptive and expressive scales was 0.93 and 0.98, respectively. The internal consistency, using the KR-21, for the receptive and expressive scales was 0.88 and 0.92, respectively.

Conclusion: A language development scale has been developed to assess receptive and expressive language skills in Iranian children aged 2-6 years. The validity and reliability of the tool were confirmed.

Please cite this article as: Sheibani F, Ghoreishi ZS, Nilipour R, Pourshahbaz A, Mohammad Zamani S. Validity and Reliability of a Language Development Scale for Persian-speaking Children Aged 2-6 Years. *Iran J Med Sci.* 2020;45(4):259-268. doi: 10.30476/ijms.2020.72538.0.

Keywords • Language development • Iran • Child • Language disorders

Introduction

Language development during early childhood is the most important phase of cognitive development in children.¹ A common problem among children is language delay as a result of cognitive impairment, lack of sensory-motor skills, and neurological, emotional, or pervasive developmental disorders; some of unknown etiology.^{2, 3} Language delay refers to deficiencies in

verbal and non-verbal communications that could distort grammar, phonology, semantics, and pragmatics.⁴

Language disorder among preschool children in the United States is a common problem.⁵ Similarly, a recent study conducted in Arak (Iran) reported that the prevalence of articulation, voice, various speech disorders, and stuttering in primary school children was 8%, 3.5%, 11.9%, and 1%, respectively.⁶ A study of 6-year-old children in Shiraz (Iran) reported that the prevalence of specific language impairment (SLI) was 2.7%.⁷ A speech therapy clinic in Tehran (Iran) reported that the prevalence of speech and language delay, articulation and voice disorders, and stuttering was 40.3%, 6.21%, and 2.21%, respectively.⁸ These statistics on language disorders in the Iranian children are indicative of potential behavioral and social risks, which eventually imposes a substantial financial burden on society.⁹

Various language assessment tools have been identified to screen and assess the social and behavioral developmental needs of preschool children in English-speaking countries.¹⁰ Some of these tools were used to identify problems related to language comprehension or language expression skills. The second edition of Clinical Evaluation of Language Fundamentals (CELF preschool-2), designed for children aged 3-6 years, addresses both language comprehension and expression skills to assess semantics, grammar, and verbal capabilities.¹¹ The New Reynell Developmental

Language Scales (NRDLS) is designed for children between 2-7 years old and provides diagnostic information about a child's production and understanding of spoken language.¹²

The formal education on speech and language therapy in Iran has a relatively long history.¹³ However, the number of essential standardized clinical linguistic measures to identify and assess developmental language disorders is limited. In recent years, some studies have focused on the adaptation or development of language assessment tools for children, namely Test of Language Development-Primary, 3rd edition (TOLD-P, 3rd),¹⁴ MacArthur-Bates Communicative Development Inventory (MCDI),¹⁵ Newsha Developmental Scale,¹⁶ Sentence Repetition Test,¹⁷ Story Retelling Test,¹⁸ Persian Syntax Comprehension Test,¹⁹ Photographic Expressive Persian Grammar Test (PEGT),²⁰ and Persian Developmental Sentence Scoring (PDSS).²¹ An overview of some child language assessment tools deployed in Iran is described in table 1. These assessment tools, particularly those used in Iran, have certain limitations. For instance, the MCDI and Newsha Developmental Scale are parent-report instruments,^{22, 23} and the Sentence Repetition and Story Retelling tests lack sufficient sample size.^{17, 18} Besides, these tools only cover part of the clinical needs and do not provide a comprehensive outcome since the manifestation of a language disorder varies during the periods of growth. Furthermore, other tools such as the Persian Syntax Comprehension Test, PDSS,

Table 1: Characteristic of some child language assessment tools deployed in Iran

Name	Year	Language aspect	Age	Sample size	Psychometrics features
Test of Language Development-Primary, 3rd edition ¹⁴	2002	Syntax, semantic and phonology comprehension and expression	4-8 years	1,235	Internal consistency=0.55 Criterion validity=0.6
McArthur-Bates Communicative Development Inventory ¹⁵	2009	Vocabulary comprehension and expression, use of gestures and phrases	8-16 months	115	Min. stability coefficient=0.43 Max. stability coefficient=0.98
Newsha Developmental Scale ¹⁶	2009	Hearing, receptive and expressive language	Birth to 6 years	593	Test-retest reliability>0.95 Inter-rater reliability>0.95 CVI between 0.8 and 1
Sentence Repetition Test ¹⁷	2011	Grammatical potency	2.5-4 years	72	CVI=80% ICC=0.95
Story Retelling Test ¹⁸	2012	Expression, narrative	6-7 years	72	CVI=0.89 ICC=83% Cronbach's α =0.77
Persian Syntax Comprehension Test ¹⁹	2015	Syntax comprehension	4-6 years	536	CVI=0.81 Test-retest reliability=0.56 Internal consistency=0.89
Photographic Expressive Persian Grammar Test ²⁰	2016	Grammar expression	4-6 years	500	CVI=0.86 Inter-rater reliability=0.98 Test-retest reliability=0.98 KR-21=0.82
Persian Developmental Sentence Scoring ²¹ (24)	2016	Morpho-syntactic abilities	30-65 months	115	Convergent validity=0.97 Internal consistency=0.86 Cronbach's α =0.79

and PEGT only evaluate a specific aspect of language¹⁹⁻²¹ and thus cannot be considered as a comprehensive language test. The TOLD-P does not cover children under 4 years old¹⁴ and the NRDLS only requires participation for 6 months.²⁴

Clearly, there was a need for an effective assessment tool to avert physicians from wasting precious time on unrelated or unimportant issues. To the best of our knowledge, there are no comprehensive and validated developmental language screening tests for Persian-speaking children in Iran. Considering the high prevalence and the long-term consequences of early language delay, the present study aimed to develop a valid and reliable language development instrument for Persian-speaking children aged 2-6 years.

Materials and Methods

The present cross-sectional study was conducted during 2016-2017 in three main Iranian cities, namely Mashhad, Tehran, and Isfahan. The target population was children aged 2-6 years in various kindergartens and schools. The random stratified sampling method was used to recruit children from different areas (north, central, and south) of these cities based on their age and the education level of their parents. The inclusion criteria were: between 2-6 years old, speaking the Persian language, good general health (confirmed by teachers or parents), and full-term (born between 37-42 weeks gestation) based on medical records. The exclusion criteria were: serious visual, hearing, speech, and language disorders (confirmed by a speech-language pathologist); articulatory abnormalities; and neurological, psychological, or mental disorders. Accordingly, a total of 460 children were recruited from Mashhad (n=160), Tehran (n=150), and Isfahan (n=150). The children were assigned to ten age groups with 6 months age increment. As part of a pilot test for validity assessment, 40 additional children were recruited from Mashhad and grouped into ten age groups (four children per group).

The study was approved by the Ethics Committee of the University of Social Welfare and Rehabilitation Sciences, Tehran, Iran (number: IR.USWR.REC.1394.173). Written informed consent was obtained from the parents and anonymity of the children was guaranteed.

Receptive and Expressive Scales

The initial part of the design of the language development scale was conducted in accordance with a previous study.²⁵ In-line

with the NRDLS¹² and CELF¹¹ tools, both the language comprehension and expression skills were considered and the sub-scales were extracted based on the developmental needs of children. Attention was also paid to the choice of the linguistic construct (nouns preceded verbs²⁸) which is beneficial in clinical diagnosis.^{26, 27} By doing so, the initial sub-scales for both scales were the receptive and expressive vocabulary of nouns and verbs. The two-word utterances²⁸ was then assessed in the sub-scales for receptive and expressive phrases. The receptive and expressive simple sentences were based on a child's ability to understand and produce an argument structure in a simple sentence.^{28, 29} To assess the understanding and production of present and past tense verbs,²⁸ receptive and expressive sections of verb morphology were considered. An average Iranian child speaks the Persian language from the age of 3 years,^{28, 30} based on which the sub-scale for receptive pronouns was designed. Syntactic complexity was evaluated by receptive and expressive complex sentences.²⁸ In the final section, children had to draw a conclusion from the contexts and clues³¹ and make grammaticality judgments to determine their metalinguistic awareness.³²

Considering the linguistic and cultural issues related to each item, the "item pool" design process was used to define each sub-scale of the receptive and expressive scales. Based on credible literature sources and the opinion of experts, a pool of linguistic items was developed in accordance with their lexical, morphological, and syntactic features. In addition, the typical linguistic characteristics of Persian-speaking children of the same age group were used. In such a design, a pool of common and frequently used nouns is required for the sub-scales of both receptive and expressive vocabulary.^{28, 33-36} The comprehension and expression of the designed two-word utterance was assessed using elements such as verbs, names of household items, toys, animals, and certain tasks.^{28, 34, 35} The vocabulary size increases by adding nouns and a few verbs. To do so, several verbs were selected from various credible literature sources.^{27, 28, 34-36} In order to be able to connect objects, a child requires lexical knowledge and the ability to use a language, based on which it can gradually process verbs and prepositions. Accordingly, the ability of a child to build a sentence by using simple words was assessed.^{28, 34, 35} Emergence of past tense verbal morphology included paired sentences with past and present tense verbs as well as receptive pronouns (singular pronoun) that appeared earlier.^{28, 30} The length and complexity of a sentence (e.g., phrases,

clauses) increase as a child grows. Syntactic complexity was evaluated toward the end of the receptive and expressive scales by using complex sentences such as relative clauses and questions.²⁸ The final part of the scales included inference items in terms of practicality (receptive scale) and grammaticality judgment (expressive scale).³⁷

Based on the above-mentioned process, a preliminary version of the Persian Language Development Scale (PLDS-1) was developed. The draft was presented to a panel of experts with research or clinical experience (11 speech-language pathologists, 2 clinical linguists, and 3 methodologists) for critical review and recommendations. To confirm the content validity of PLDS-1, feedback from the panel was obtained through face-to-face meetings or email communication with textual explanations. The outcome of the panel review resulted in introducing toys, objects, and images (colorful cartoon figures). These items had to be in line with the Iranian culture, and attractive as well as familiar to the children. The images were further examined by the research team members and five speech therapists to confirm various aspects such as resolution, clearness, and sequence. Accordingly, the second version of the Persian Language Development Scale (PLDS-2) was developed and was used in a pilot study of 40 children from Mashhad. The outcome of the pilot study led to the development of the third version of the Persian Language Development Scale (PLDS-3). This version was then used to conduct the present cross-sectional study of 460 children.

Assessment Process

A team of four speech and language pathologists performed the test. To prevent the assessor's own dispositions, the team was fully trained on the process and each member was given textual instructions for optimal implementation. The children were assessed individually in a quiet room at their current kindergartens or schools. The duration of the assessment varied from 30 to 45 minutes. A video camera (Sony HDR-PJ410, Japan) was used to record the voice and image of the participants. The camera was positioned at 3 m distance from the children to avoid distraction. Each item of the scale was scored based on a correct (1) or incorrect (0) response. The total score for the receptive and expressive scales were 72 and 64 points, respectively. The practice items were not included in the scoring system. The test would continue until a section was completed, provided the child

remained attentive and cooperative. The test was terminated if a child failed to successfully complete a section.

Receptive Scale

The 72-item receptive scale included eight sections, namely vocabulary, phrase, verbs, simple sentence, morphology, pronouns, complex sentence, and inference. A warm-up session (5 items) was also included to accommodate young and shy children. The receptive vocabulary section included 10 test items, each associated with an object. Upon naming an object by the assessor, the child had to point to the corresponding object. The receptive phrase section included 10 tests and 2 practice items. Upon asking to perform a certain task, the child had to understand the words and prepositions associated with two objects. The receptive verbs section (10 tests and 4 practice items) covered the comprehension of transitive verbs using objects and figures. The 10-item receptive simple sentence section aimed to assess the child's understanding of simple sentences using objects and images. The receptive morphology section included 6 tests and 2 practice items to evaluate the child's comprehension of the contrast between the past tense and present tense using an image associated with a task. The receptive pronouns section (6 tests and 2 practice items) was image-based and evaluated the child's ability to differentiate non-reflexive from reflexive pronouns. The section on receptive complex sentence (10 test and 2 practice items) evaluated the comprehension of relative and non-canonical clauses using an image associated with a task. The final section, inference, included 10 test items. Based on an image, the child had to use its knowledge and inference skills to identify the depicted person.

Expressive Scale

The 64-item expressive scale included seven sections, namely vocabulary, phrase, verbs, simple sentence, verb morphology, complex sentences, and grammaticality judgment. Similar to the previous section, a warm-up session (5 items) was included. The expressive vocabulary section included 10 test items to assess the child's ability to name a familiar object. The expressive phrase section (10 test and 2 practice items) required the child to produce at least two nouns or a prepositional phrase. The section on expressive verbs (10 tests and 4 practice items) assessed the ability of a child to respond to any form of an appropriate target verb. The expressive simple sentence section included 10 tests and 4 practice items that

required the inclusion of a pronominal (agent) and a nominal (patient) in the child's response. The expressive verb morphology section (6 test and 2 practice items) assessed the child's ability to correctly produce verbs in the present tense and past tense. The section on expressive complex sentences included 10 tests and 3 practice items. In this section, the child had to be able to describe an image by producing questions and sentences that contain relative and non-canonical clauses. The final expressive grammaticality judgment section (3 practice and 8 test items) assessed the child's ability to identify a grammatically correct sentence.

Statistical Analysis

The data were analyzed using SPSS software

version 18.0 (Chicago, IL, USA) and Mplus software version 7. The Kolmogorov-Smirnov test was used to confirm the normal distribution of the samples. Analysis of variance (ANOVA) was used to measure the differences between the means. The Pearson correlation coefficient was used to measure the inter-rater agreement. $P < 0.05$ was considered statistically significant.

The internal consistency of the instrument was evaluated using the Kuder-Richardson Formula 21 (KR-21). The test-retest reliability of the instrument was measured using the intraclass correlation coefficient (ICC). Confirmatory factor analysis (CFA) was assessed with the Tucker-Lewis index ($TLI \geq 0.95$), comparative fit index ($CFI \geq 90$), and root mean square error of approximation ($RMSEA < 0.08$).

Table 2: Measurement model for each section of the reception scale

Receptive items	Factor loading (SE)	Receptive items	Factor loading (SE)
Vocabulary		6	0.997 (0.135)
1	1.000 (0.000)	7	1.232 (0.156)
2	1.204 (0.080)	8	0.899 (0.120)
3	0.672 (0.071)	9	0.754 (0.100)
4	1.040 (0.069)	10	0.596 (0.092)
5	1.225 (0.074)	Morphology	
6	1.143 (0.076)	1	1.000 (0.000)
7	1.083 (0.051)	2	0.450 (0.078)
8	0.980 (0.073)	3	1.149 (0.083)
9	0.997 (0.080)	4	0.625 (0.071)
10	0.672 (0.071)	5	0.619 (0.075)
Phrase		6	1.455 (0.109)
1	1.000 (0.000)	Inferencing	
2	.849 (0.099)	1	1.000 (0.000)
3	886 (0.064)	2	1.036 (0.076)
4	1.133 (0.091)	3	0.922 (0.069)
5	1.368 (0.091)	4	0.837 (0.060)
6	.798 (0.076)	5	0.676 (0.059)
7	1.056 (0.080)	6	0.824 (0.056)
8	1.206 (0.082)	7	0.514 (0.065)
9	1.027 (0.085)	8	0.944 (0.059)
10	1.147 (0.087)	9	0.200 (0.095)
Verbs		10	0.488 (0.079)
1	1.000 (0.000)	Pronouns	
2	0.965 (0.074)	1	1.000 (0.000)
3	0.622 (0.093)	2	1.147 (0.031)
4	0.994 (0.056)	3	0.696 (0.056)
5	0.524 (0.059)	4	1.025 (0.032)
6	0.823 (0.047)	5	0.989 (0.053)
7	0.491 (0.063)	6	0.880 (0.045)
8	0.824 (0.069)	Complex sentence	
9	0.751 (0.054)	1	1.000 (0.000)
10	0.815 (0.043)	2	1.163 (0.067)
Simple sentence		3	1.108 (0.072)
1	1.000 (0.000)	4	0.482 (0.056)
2	1.187 (0.172)	5	0.280 (0.070)
3	1.292 (0.175)	6	0.443 (0.060)
4	0.969 (0.128)	7	1.130 (0.066)
5	0.863 (0.120)	8	0.892(0.055)

Results

The mean score for the receptive and expressive scales was 58.09 ± 7.82 and 51.07 ± 8.79 , respectively. The Pearson correlation coefficient between the scores of the scales was 0.92 ($P=0.0001$). The validity of the instrument was assessed using the content validity ratio (CVR). Based on the number of experts in the panel, $CVR \geq 0.47$ was representative of good content validity. As a result, of the 178 items, 24 items with $CVR < 0.47$ were omitted; the CVR of the remaining items was 0.85. The face validity of the instrument was assessed in the pilot study on 36 children. Different sections of the scales such as instructions, items, and target responses were investigated qualitatively. No items were deleted, but some elements were modified for a better understanding and clarity. The construct validity of the instrument was assessed using CFA with

weighted least squares means and variance. The RMSEA, with 90% confidence interval, for the receptive and expressive scales was 0.04 and 0.02, respectively. Both the CFI and TLI for the receptive and expressive scales were 0.85 and 0.95, respectively. The factor loading was higher than 0.4 for all items of the receptive and expressive scales (tables 2 and 3).

There was a significant difference between the mean scores of the receptive and expressive scales in the different age bands (figures 1 and 2). Post-hoc analysis of the receptive scale scores revealed significant differences between each age band, except for the age bands 1 versus 2 and 3 versus 4. Similarly, significant differences between the scores of each age band in the expressive scale were observed, except for the age bands 4 versus 5, 5 versus 6, 6 versus 7, 7 versus 8, and 8 versus 9. In contrast, there was no significant difference in the language

Table 3: Measurement model for each section of the expression scale

Expressive items	Factor loading (SE)	Expressive items	Factor loading (SE)
Vocabulary		3	0.973 (0.032)
1	1.000 (0.000)	4	1.009 (0.033)
2	1.009 (0.083)	5	0.929 (0.034)
3	0.891 (0.071)	6	0.983 (0.035)
4	0.973 (0.049)	7	1.036 (0.033)
5	0.706 (0.032)	8	0.755 (0.044)
6	0.548 (0.068)	9	1.062 (0.033)
7	0.996 (0.084)	10	0.842 (0.038)
8	1.022 (0.091)	Verb morphology	
9	0.922 (0.072)	1	1.000 (0.000)
10	0.926 (0.084)	2	0.695 (0.054)
Phrase		3	0.925 (0.061)
1	1.000 (0.000)	4	0.881 (0.067)
2	0.997 (0.033)	5	0.911 (0.060)
3	0.968 (0.028)	6	0.754 (0.055)
4	1.018 (0.025)	Simple sentence	
5	1.044 (0.026)	1	1.000 (0.000)
6	0.961 (0.034)	2	1.068 (0.024)
7	0.949 (0.032)	3	1.013 (0.024)
8	0.972 (0.035)	4	0.981 (0.027)
9	0.976 (0.031)	5	1.078 (0.031)
10	0.950 (0.033)	6	1.058 (0.030)
Verbs		7	1.050 (0.032)
1	1.000 (0.000)	8	0.828 (0.048)
2	1.162 (0.093)	9	0.863 (0.045)
3	1.138 (0.092)	10	0.866 (0.041)
4	1.242 (0.096)	Grammaticality judgment	
5	0.758 (0.091)	1	1.000 (0.000)
6	0.918 (0.076)	2	1.331 (0.097)
7	1.100 (0.103)	3	1.217 (0.088)
8	0.937 (0.078)	4	0.807 (0.083)
9	1.050 (0.098)	5	1.181 (0.103)
10	0.953 (0.077)	6	0.878 (0.083)
Complex sentences		7	1.099 (0.094)
1	1.000 (0.000)	8	1.094 (0.083)
2	0.923 (0.035)		

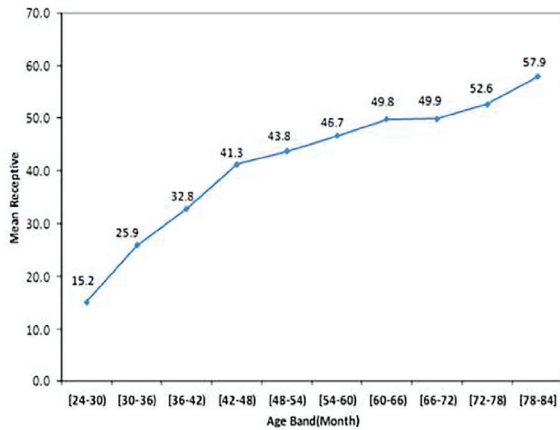


Figure 1: Mean scores of the receptive scale per age band are shown.

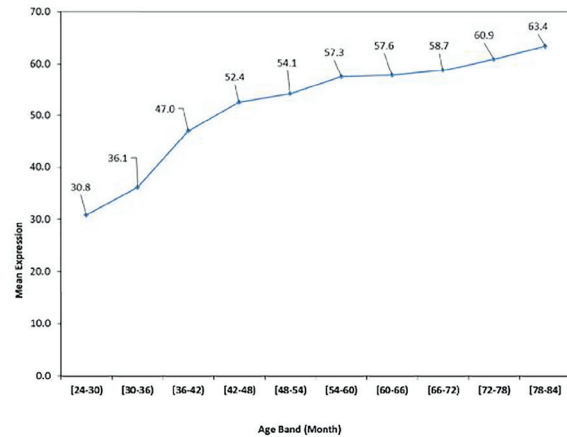


Figure 2: The figure shows Mean scores of the expression scale per age band.

performance score between boys and girls (table 4). However, girls generally had a higher score than boys.

The internal consistency (KR-21) for the receptive and expressive scales was 0.88 and 0.92, respectively. The results showed a high correlation between the items. The test-retest with 10 days interval showed a significant correlation between the coefficients of both scales (table 5). In terms of inter-rater reliability, the correlation between the grades given to the test by three raters was calculated independently: the receptive scale (ICC=0.93, P<0.001) and the expressive scale (ICC=0.98, P<0.001).

Discussion

The findings of the present study confirmed PLDS as a reliable and valid language development tool in children aged 2-6 years. The main feature of PLDS lies in its comprehensiveness to assess both language comprehension and expression skills. In addition, the use of real objects and toys before resorting to images encouraged the children to become more cooperative.

In line with the result of the PEGT,²⁰ the calculated CVR value for PLDS was 0.86; indicating a good degree of validity. Mohamadi

and colleagues also reported a good content validity for the Persian Syntax Comprehension Test and used CVR≥0.47 as a threshold for good validity.¹⁹ The content validity for the PDSS tool was not reported, but its convergent validity (correlation coefficient of 0.97) was good.²¹ Similarly, a good criterion validity (0.6) for the TOLD-P tool was reported.¹⁴

The results showed that PDLS could differentiate children according to their age. Considering the effect of age (ten age bands) on the language scores, the PDLS tool met the social and behavioral developmental needs of the children. We observed that the mean language score increased with an increase in the child's age. Similarly, the PEGT could also differentiate the ages of the children; reporting a significant difference in the mean score of four age bands.²⁰ The effect of age was also reported in the PDSS tool²¹ and the result of the ANOVA test showed a significant difference in the mean score between different age groups [F (5,109)=24.69]. The developers of the Newsha Developmental Scale²³ and TOLD-P¹⁴ also reported the effect considering different age groups. However, the effect of children's age and sex was not reported for the Persian Syntax Comprehension test.¹⁹ The developers of

Table 4: The mean score of receptive and expressive scales with respect to the sex of children

PLDS	Male (Mean±SD)	Female (Mean±SD)	P value*
Receptive	51.32±12.78	53.81±10.26	0.224
Expressive	41.09±15.19	44.25±12.69	0.196

*Student's t test, P<0.05

Table 5: Results of test-retest reliability with 10 days interval

PLDS	Test	Re-test (Mean±SD)	Correlation coefficient (Mean±SD)	P value
Receptive	58.09±7.83	61.11±6.38	0.96	0.0001
Expressive	51.70±8.80	59.3±7.87	0.93	0.0001

the PEGT tool, however, reported a significant difference in the mean score between boys and girls.²⁰ For the Newsha Developmental Scale, only 3 out of the 373 items showed a significant difference between boys and girls.²³ Note that the PLDS tool could not differentiate children based on sex, however, it did identify that girls had a better language performance than boys.

The results of the test-retest reliability showed a significant correlation between the coefficients of both scales; indicating high stability of the test and consistency of the PLDS score over a 10 days interval. Similarly, the PEGT and Newsha Developmental Scale tools had a high test-retest reliability of $r=0.91$ and $r>0.95$, respectively.^{20, 23} The result of the inter-rater reliability showed an excellent ICC and a high degree of agreement between the raters. The reported ICC for the PDSS instrument was 0.77.²¹ The results indicated that the score of the language performance was not influenced by the rater bias. The same inter-rater reliability result was reported for PEGT ($r=0.91$) and Newsha Developmental Scale $r>0.95$.²⁰ The internal consistency, based on KR-21, of the PLDS tool for the receptive and expressive scales was excellent. A high correlation between items in each scale indicated that the scales measured a single variable. Similarly, the reported internal consistency for PEGT revealed a high correlation ($r=0.82$).²⁰ For the PDSS tool, Cronbach's alpha of 0.79 and split-half coefficient of 0.86 was reported; indicating an acceptable internal consistency.²¹ The stability coefficient of the sub-scales of MCDI was between 0.43 to 0.98 and a mean internal consistency of 0.55 was reported for TOLD-P.^{14, 22}

The main limitation of the present study was the lack of well-standardized developmental language assessment tools. As a result, we could not perform concurrent validity to correlate the PLDS scales with a gold standard. In addition, our findings were limited to the children in kindergarten and schools, excluding those at home or in private schools. It is recommended to evaluate the sensitivity and specificity of the PLDS in other populations, such as children with speech-language impairment or hearing loss. Future studies with larger sample size and the inclusion of other Iranian cities are recommended to allow standardization of the PLDS tool.

Conclusion

A language development scale is presented that assesses receptive and expressive language skills in children aged 2-6 years. The validity and

reliability of the tool were confirmed. The PLDS can be used by speech-language pathologists or as an instrument in diverse research projects. It can also support various clinical diagnoses associated with language disorders.

Conflict of Interest: None declared.

References

- 1 Prelock PA, Hutchins T, Glascoe FP. Speech-Language Impairment: How to Identify the Most Common and Least Diagnosed Disability of Childhood. *Medscape J Med.* 2008;10:136-. PubMed Central PMCID: PMCPMC 2491683.
- 2 Paul R. Language disorders from infancy through adolescence: Assessment & intervention. New York: Elsevier Health Sciences; 2007.
- 3 Owens RE, Metz DE, Haas A. Introduction to communication disorders. Boston: Allyn and Bacon Needham Heights, MA.; 2000.
- 4 Sices L, Feudtner C, McLaughlin J, Drotar D, Williams M. How do primary care physicians identify young children with developmental delays? A national survey. *J Dev Behav Pediatr.* 2003;24:409-17. doi: 10.1097/00004703-200312000-00002. PubMed PMID: 14671474.
- 5 Rossetti LM. Communication intervention: Birth to three. New Tech Park: Cengage Learning; 2001.
- 6 Yavari A, Fatehi F, Dalvand H, Valizadeh A, Moradzadeh R, Mirhoseini FS. Prevalence of Speech Disorders in Arak Primary School Students, 2014-2015. *Prevalence.* 2016;19:87-94. Persian.
- 7 Oryadi-Zanjani MM, Vahab M, Rasouli J, Mohammadi F, Firouzi-Parizi M. The Prevalence of Specific Language Impairment in 6-Year-Old Persian-Speaking Children in Shiraz City, Iran, 2015. *International Journal of School Health.* 2016;3:1-5. doi: 10.17795/intjsh-31841.
- 8 Mehri A, Zadeh SF, Mousavi SZ. Prevalence of speech and language disorders in patients referring to the Speech therapy clinic in Tehran Rehabilitation Faculty during 10 years. *J Rehab Med.* 2016;5:132-40. Persian.
- 9 Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res.* 1997;40:1245-60. doi: 10.1044/jslhr.4006.1245. PubMed PMID: 9430746; PubMed Central PMCID: PMCPMC5075245.

- 10 Ringwalt S. Developmental Screening and Assessment Instruments with an Emphasis on Social and Emotional Development for Young Children Ages Birth through Five. National Early Childhood Technical Assistance Center (NECTAC). 2008;ED505971.
- 11 Paslawski T. The Clinical Evaluation of Language Fundamentals, (CELF-4) A Review. *Canadian Journal of School Psychology*. 2005;20:129-34. doi: 10.1177/0829573506295465.
- 12 Edwards S, Letts C, Sinka I. The new Reynell developmental language scales. 2011. doi: 10.1037/t05236-000.
- 13 Nilipour R. Emerging issues in speech therapy in Iran. *Folia Phoniatr Logop*. 2002;54:65-8. doi: 10.1159/000057916. PubMed PMID: 12037418.
- 14 Hassanzade S, Minayi A. Adaption and language development test TOLD-P: 3 for children Tehrani Persian language (part I). *J Research on Exception Children*. 2002;1:34-119. Persian.
- 15 Kazemi Y, Nematzadeh S, Hajian T, Heidari M, Daneshpajouh T, Mirmoeini M. Finding the validity and reliability coefficient of McArthur-Bates communicative development inventory in Farsi. *Journal of Research in Rehabilitation Sciences*. 2008;3:45-51. Persian.
- 16 Jafari Z, Ashayeri H, Malayeri S, Alaedini F. Translation, reliability and validity of the Newsha Hearing, Language and Speech Development Scale among children in Iran. *Payesh (Health Monitor)*. 2009;8:271-8. Persian.
- 17 Hasanati F, Agharasouli Z, Mahmoudi Bakhtiyari B, Kamali M. Sentence repetition test for measurement of grammatical development in Farsi speaking children. *Bimonthly Audiology-Tehran University of Medical Sciences*. 2011;20:73-81. Persian.
- 18 Jafari S, Agharasouli Z, Modaresi Y, Kamali M. Developing a story retelling test for the assessment of language structure in Persian-speaking children. *Bimonthly Audiology-Tehran University of Medical Sciences*. 2012;21:51-61. Persian.
- 19 Mohamadi R, Alavije MR, Minayi A, Modaresi Y, Dastjerdi MK, Ghaderi M. Generation and content validation of a Persian syntax comprehension test. *Psychology of Language and Communication*. 2015;19:222-36. doi: 10.1515/plc-2015-0013.
- 20 Haresabadi F, Ebadi A, Shirazi TS, Dastjerdi Kazemi M. Design and validation of a Photographic Expressive Persian Grammar Test for children aged 4–6 years. *Child Language Teaching and Therapy*. 2016;32:193-204. doi: 10.1177/0265659015595445.
- 21 Jalilevand N, Kamali M, Modarresi Y, Kazemi Y. The Persian developmental sentence scoring as a clinical measure of morpho-syntax in children. *Med J Islam Repub Iran*. 2016;30:435-59. PubMed PMID: 28210600; PubMed Central PMCID: PMC5307620.
- 22 Kazemi Y, Nematzadeh S, Hajian T, Heidari M, Daneshpajouh T, Mirmoeini A. The validity and reliability coefficient of Persian translated McArthur-Bates Communicative Development Inventory. *Journal of Research in Rehabilitation Sciences*. 2008;4.
- 23 Zahra J, Hasan A, Saeed M, Farshid A. Reliability and Validity of “Hearing Test, Language and Speech Testing Test for Newsha” in Children Birth to 6 Years Old Persian Language. *Monitoring*. 2009;8:271-8.
- 24 Letts C, Edwards S, Schaefer B, Sinka I. The new Reynell developmental language scales: descriptive account and illustrative case study. *Child Language Teaching and Therapy*. 2014;30:103-16. doi: 10.1177/0265659013492784.
- 25 Malhotra NK. Questionnaire design and scale development. *The handbook of marketing research: Uses, misuses, and future advances*. 2006. p. 176-202.
- 26 Maleki Shahm Mahmood T, Soleymani Z, Faghihzade S. The study of language performances of Persian children with specific language impairment. *Bimonthly Audiology-Tehran University of Medical Sciences*. 2011;20:11-21. Persian.
- 27 Mohammad M. Examining expression of lexical categories in Farsi-speaking children with specific language impairment. *Speech and Language Pathology*. 2013;1:8-19.
- 28 Jalilevand N. *Speech & language development in Farsi speaking children*. Tehran: Danjeh; 2001. Persian.
- 29 Tomasello M. *Constructing a language*. Cambridge: Harvard university press; 2009.
- 30 Jalilevand N, Ebrahimipour M. Pronoun acquisition in Farsi-speaking children from 12 to 36 months. *Journal of Child Language Acquisition and Development*. 2013;1:1-9.
- 31 Letts C, Leinonen E. Comprehension of inferential meaning in language-impaired and language normal children. *Int J Lang Commun Disord*. 2001;36:307-28. PubMed PMID: 11491482.
- 32 Ghayoumi-Anaraki Z, Haresabadi F, Maleki Shahm Mahmood T, Ebadi A, Vakili V, Majidi Z. The Grammatical Deficits of Persian-Speaking Children with Specific Language Impairment. *Journal of Rehabilitation Sciences & Research*. 2017;4:102-8.

- 33 Bakhtiar M, Nilipour R, Weekes BS. Predictors of timed picture naming in Persian. *Behav Res Methods*. 2013;45:834-41. doi: 10.3758/s13428-012-0298-6. PubMed PMID: 23292568.
- 34 Nematzadeh S, Dadras M, Dastjerdi Kazemi M, Mansoorizadeh M. Persian core vocabulary based on Iranian children. Tehran: Madreseh. 2011. Persian.
- 35 Zarei Mahmood Abadi M, Zarifian T. Expressive Vocabulary Development of a 15-34 months Persian Speaking Child: A Longitudinal case study. *Rehab Med*. 2018;6:308-17. doi: 10.22037/jrm.2018.110984.1674. Persian.
- 36 Mehdipour N, Shirazi T, Nematzadeh S, editors. Development of expressive vocabulary and sentences in 18-24 months old Persian-speaking children. Tehran: First Conference of Iranain Linguistics Society; 2002. Persian.
- 37 McDonald JL. Grammaticality judgments in children: the role of age, working memory and phonological ability. *J Child Lang*. 2008;35:247-68. doi: 10.1017/S0305000907008367. PubMed PMID: 18416859.