

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/280603405>

Design and validation of a Photographic Expressive Persian Grammar Test for children aged 4–6 years

Article in *Child Language Teaching and Therapy* · July 2015

DOI: 10.1177/0265659015595445

CITATIONS

8

READS

148

4 authors, including:



Fatemeh Haresabadi
Mashhad University of Medical Sciences

18 PUBLICATIONS 40 CITATIONS

[SEE PROFILE](#)



Abbas Ebadi
Baqiyatallah University of Medical Sciences

529 PUBLICATIONS 2,631 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Resilience of Patients With Chronic Physical Diseases: A Systematic Review and Meta-Analysis [View project](#)



Nurse Education [View project](#)

Design and validation of a Photographic Expressive Persian Grammar Test for children aged 4–6 years

Child Language Teaching and Therapy

1–12

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0265659015595445

clt.sagepub.com



Fatemeh Haresabadi

Pediatric Neurorehabilitation Research Center, Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

Abbas Ebadi

Behavioral Sciences Research Center (BSRC), Nursing Faculty, Baqiyatallah University of Medical Sciences, Tehran, Iran

Tahereh Sima Shirazi

Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

Mehdi Dastjerdi Kazemi

Research Institute of Exceptional Children, Research Institute of Education, Tehran, Iran

Abstract

Syntax has a high importance among linguistic parameters, and syntax-related problems are the most common in language disorders. Therefore, the present study aimed to design a Photographic Expressive Persian Grammar Test for Iranian children in the age group of 4–6 years and to determine its validity and reliability. First, the target morphosyntactic structures among Persian-speaking 4–6-year-old children were extracted, and items related to each structure were designed. After both content and face validity were determined and modifications were applied, the initial version of the test was performed on 100 children. The final version of the test was performed on 400 eligible, typically-developing children selected using the random (cluster) method. Psychometric properties investigated in this study include construct validity (convergent validity and age and gender discriminative validity) and reliability (test–retest, inter-rater, and internal consistency). Content validity of each item was between 0.8 and 1, and content validity of the whole scale was 0.86. Exploration of construct validity suggested that age and gender affect the test scores. Convergent validity was found to be significant. Results of test–retest, and inter-rater reliability were significantly correlated and the test was found to have high internal consistency.

Corresponding author:

Abbas Ebadi, Behavioral Sciences Research Center (BSRC) and Nursing Faculty, Baqiyatallah University of Medical Sciences, Tehran, Iran.

Emails: ebadi1347@yahoo.com

The Photographic Expressive Persian Grammar Test with 40 items is the first reliable and valid scale that exclusively and accurately evaluates morphosyntactic characteristics of 4–6-year-old Persian-speaking children.

Keywords

Children, language assessment, Persian, syntax

I Introduction

Children begin to learn to talk with short, one-word phrases and gradually, by gaining competence in the grammar of the language, develop the ability to produce and understand longer phrases (Hoof, 2009). Thus, children achieve communication, which is one of the most important and major needs of human life (Miremadi, 2010). Many children, however, face challenges in acquiring language skills for different reasons and have language disorders. These clients form a heterogeneous group in terms of severity, cause, language symptoms, and clinical prognosis (Bishop, 2006). One of the reasons for the heterogeneity of this group is the variety of disorders in different developmental parameters of language (syntax, semantics, and pragmatics) or a combination of them (Paul and Norbury, 2012). The literature has shown that a grammar deficiency is one of the most typical language problems of children with hearing loss, specific language impairment, and Down syndrome. So, it has long been the focus of research in studies on children with language disorders (Golpour et al., 2006; Maleki et al., 2012; Paul and Norbury, 2012; Zarifian et al., 2013).

To examine the linguistic profile of these children, we need linguistic evaluations. Most clinical researchers and professionals collect speech samples and evaluate the expressive language properties and problems of children based on their analyses of these speech samples (Paul and Norbury, 2012). Although sampling of a child's speech is recognized as the best existing indicator of language performance, particularly regarding syntax and morphology, there are problems associated with this method. It is suggested that the primary obstacle in using this method is its time-consuming nature, and the secondary obstacle is the lack of standard methods for data collection and subsequent interpretations (Rew and Irwin, 1985). Therefore, it is proposed that using standard instruments can be a suitable alternative in order to avoid these kind of limitations and to identify children with language disorders promptly (Clark, 2003).

I Types of tests in language assessment

Formal tests designed for language assessment can be divided into two categories: those that assess several different aspects of language and those that assess a single aspect of language in more depth. Assessments in the first category examine language in the five areas of morphology, syntax, semantics, phonology, and pragmatics. Such tests might assess language only at the level of comprehension or expression, although some cover both. The Clinical Evaluation of Language Fundamentals (CELF; Semel and Wiig, 1980) for individuals between 6 and 16 years of age includes both comprehension and expressive levels that assess higher-order semantic, grammatical, and verbal memory abilities (Kaderavek, 2014). The Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999) is another test from this category that provides a precise picture of language processing skills and structural knowledge across the age range 3 years 9 months to 21 years 11 months. The 15 sub-tests included in this battery measure comprehension, expression, and retrieval skills in four structural categories: lexical/semantic, syntactic, pragmatic and supralinguistic (Hersen, 2004). These tests are extensive, expensive and time-consuming to

administer, taking 30 to 60 minutes to test a participant, and needing more than one session to complete.

Tests of the second category are those that exclusively address one area of language, such as grammar (morphology and syntax), in order to study that particular area in more detail. Some grammar-related tests assess both expression and comprehension; examples include the Northwestern Syntax Screening Test (NSST; Lee, 1969), which evaluates expressive language and comprehension of children aged from 3 years to 7 years 11 months, and the Rice/Wexler Test for Early Grammatical Impairment (TEGI; Rice and Wexler, 2001), a test for identifying, screening, and diagnosing grammatical problems in children aged 3 to 8 years (Davis, 2010). A few tests specifically evaluate expressive grammar including the Test for Examining Expressive Morphology (TEEM; Shipley et al., 1983) and the Structured Photographic Expressive Language Test – 3rd edition (SPELT-3; Perona et al., 2005).

TEEM has been validated for use with children aged between 3 years and 7 years 11 months, evaluates use of expressive morphemes and takes 7 minutes to administer. It assesses present progressive, plurals, possessives, past tenses, third-person singulars and derived adjective (Paul and Norbury, 2012). SPELT-3 is applicable to children of 4 to 9 years of age, is standardized on 1,580 typically-developing English-speaking children and is available in both English and Spanish. This test contains 54 colour photographs from everyday situations and activities. These photographs are used along with a series of eliciting sentences in order to elicit certain morphosyntactic structures from child speech. The structures evaluated include 11 morphological and 8 syntactic structures. The time required for administering the test is 15–25 minutes and the obtained information is analysed in a short time (Dawson et al., 2003).

2 Persian Grammar

Persian is an Indo-European language largely spoken in Iran, Afghanistan, and Tajikistan. This language is very rich in terms of morphology, and many variations are observed in its word order. In Persian, there are no gender and state agreements (Bakhtiar et al., 2013). The study of grammatical structures related to Persian-speaking children (Jalilevand et al., 2011; Jalilevand and Ebrahimpour, 2013; Meshkatoddini, 2004) and comparison of its results with the grammatical structures of English-speaking children (Brown, 1973; Clark, 2003; Ervin and Miller, 1963) demonstrated that no significant difference exists in terms of the appearance time of these structures in both languages. In Persian six inflectional morphemes are used for verb inflection and verbal clitics (Jalilevand, 2012), whereas in English, only the third person singular has inflectional morphemes. In English, personal pronouns of third person singular are ‘she’ (female) and ‘he’ (male) (Brown, 1973). In Persian, however, there is only /ʔu/ for third person singular, and the language does not distinguish genders. Past-tense-forming morphemes in Persian are /-id/, /d/, /-a:d/ and /-t/, whereas in English, there is only ‘ed’ (Jalilevand, 2012).

The variety of question words and verb tenses in Persian appears to be more extensive than in English. While in English plural morphemes and third person singular morpheme change according to the phonetic qualities of the word ending (Brown, 1973). This variation does not exist for the plural morpheme in Persian (Jalilevand, 2012). In English, information questions require inverting of the order of the auxiliary and subject and a rising intonation, but in Persian, only a rising intonation is used to form questions (Jalilevand, 2012). In Persian, personal pronouns are also used as possessive pronouns, whereas in English, in addition to this feature, there are pronouns that are specifically possessive (Jalilevand and Ebrahimpour, 2013). We may conclude that some structures are more complex in one language than in others. Thus, the presence of such differences and complexities make it impossible to use a test specific to one language for other languages.

3 Language assessments in Persian

A normative and valid test for evaluating Persian parameters in Iran is the Test of Language Development – Primary 3rd edition (TOLD-P3; Hasanzadeh and Minaei, 2000). This test contains 9 subtests and evaluates three language areas (syntax, semantics, and phonology) in two levels of comprehension and expression, and is designed for examining linguistic competence of children aged 4 to 8 years (Hasanzadeh and Minaii, 2002). Niusha's Development Test (Jafari and Asad-Malayeri, 2012) is another valid instrument that is employed in Iran for evaluating a child's development in the five areas of hearing, expressive language, comprehension language, speech, cognition, and social and motor skills. This test spans the age group of newborns to 72 months of age (Jafari and Asad-Malayeri, 2012). In both TOLD-P3 and Niusha, few items assess a child's expressive grammar due to the tests' multidimensionality, results might not represent an accurate profile of the child's grammatical status and may mislead clinical and research decisions.

Some tests such as the Story Retelling Test for the Assessment of Language Structure in Children (Jafari et al., 2012), Narrative Assessment Protocol (NAP; Ghasemi et al., 2012) and Sentence Repetition Test for Assessing Grammar in Children (Hasanati et al., 2011) are not frequently used because of the limited age range, the small sample size and a lack of the full psychometric evaluations (Kazemi et al., 2015).

It is indisputable that the morphological and syntactical rules are different in each language. As a result, findings of other studies about other languages cannot be generalized to Persian. Therefore, due to the shortage of suitable instruments in Iran, evaluation and clinical and research activities have been performed using informal and non-normative sundry tests. Although informal evaluations, such as collecting samples from children's speech, are therapist oriented and require a significant degree of clinical expertise and experience, too much reliance on these evaluations may lead to inaccurate research findings and clinical prognoses (Shipley and McAfee, 2009).

4 Aims of the study

Consequently, by taking into account the exclusive characteristics of Persian, this study aims to report the development and evaluation of a test where all items reflect Persian grammar properties for children aged 4 to 6 years because it is necessary to investigate the characteristics of this linguistic area accurately and exclusively (Angell, 2009). Schwab (1980) suggests that the development of measures falls into three basic stages. Stage 1 is item development, or the generation of individual items. Stage 2 is scale development, or the manner in which items are combined to form scales. Stage 3 is scale evaluation, or the psychometric examination of the new measure (Hinkin, 1995). In this study the target grammatical structures for assessment were determined through a qualitative study by the researcher (item development), and test items were developed and refined (scale development). Validity and reliability have traditionally been regarded as the basic criteria that any language test should satisfy (Walt and Steyn, 2008). In this study, several measures of validity and reliability were examined. First, face and content validity were calculated, and after the final test version was administered, its construct validity and reliability were assessed (scale evaluation).

II Methodology

1 Participants

Six hundred and twenty typically-developing Persian-speaking children aged 4–6 years participated across the various stages of the study. Sampling was carried out using the cluster method. The city of Tehran was first divided into three geographical areas of north, center, and south, from

which some nursery schools were selected randomly for sampling. Based on children's medical records, examiner's observations, parents and nursery school teachers' anecdotes, all children were identified as being healthy at the time of research and without any language disorders (as evaluated by TOLD-P3) or speech problems, such as phonological disorders (as evaluated by the Phonetic Information Test; Ghasisin et al., 2013) and stuttering (as detected by perceptual assessment of the researcher). The ethical considerations were approved by University of Social Welfare and Rehabilitation's Ethics Committee.

One hundred children participated in Step 2 of the study in which item analysis was carried out; 20 children participated in the analysis of face validity; data was then collected from 400 children (210 boys and 199 girls) to establish the psychometric properties of the test; and 100 children participated in the final stage of the study in which convergent validity was determined. As shown in Table 2 below for the psychometric analysis, children were categorized into four age groups: 48–54 months (group 1), 55–60 months (group 2), 61–66 months (group 3), and 67–72 months (group 4).

2 Test design: Psychometric properties

The test was designed and its validity and reliability were examined through the following steps.

Step 1: Designing the first version of the test (determining test items). Test items were selected using several sources of information: an extensive review of resources related to Persian grammar based on traditional linguistic perspective (Anvari and Givi, 2012; Elwell-Sutton, 1992; Givi and Anvari, 2012; Jones, 2013; Kent et al., 2011; Khanlari, 2014; Mace, 2015; Mahootian, 2003; Meshkatoddini, 2003; Miremadi, 2010; Perry, 2005; Windfuhr, 1979; Yousef and Torabi, 2013) and of literatures of morphosyntactic development of Persian-speaking children (Jalilevand and Ebrahimipour, 2013; Jalilevand et al., 2011; Meshkatoddini, 2004); by modeling SPELT-3 (Dawson et al., 2003); morphosyntactic analysis of the transcribed samples of spontaneous speech from 30 Persian speaking children aged 4–6 years; and interview of professionals in the fields of linguistics and language and speech pathology. The target grammatical structures for the first version of the test were extracted and the items related to each structure were designed. The target grammatical structures were selected based on the four criteria of use frequency in Persian, development, importance, picturability, and suitability for age group (4–6 years). All items of this test evaluate morphosyntactic characteristics of Persian speaking children aged 4–6 years exclusively. Each test item included an eliciting sentence such as 'the boy is not drawing, why...?', a target sentence such as 'because he has no pencil', and a photograph. The part of the target sentence that contained the target structure was considered as the target response and it was underlined. At this stage, the number of items designed for each structure varied between 1 and 9 items. For each item, it was attempted to design the eliciting sentence in such a way that the child would express only the target sentence, and the chances of expressing other sentences is minimized. Once eliciting and target sentences related to each item were determined, their corresponding photographs were designed. The photographs were designed in such a way as to be accurately illustrative of the eliciting and target sentences.

Content Validity of the selected items was evaluated. In order to calculate the quantitative Content Validity Index (CVI) for each Item (I-CVI) and Scale (S-CVI), 10 experts from the field of children's morphosyntactic development (speech and language therapists and linguistics) were asked to rate the test items according to a four-point Likert scale (1 irrelevant, 2 somewhat relevant, 3 very relevant, 4 highly relevant). For each item, two types of CVI were considered, including: relevancy of the target sentence to the eliciting sentence and relevancy of the photograph to the target sentence. A CVI equal or greater than 0.8 was considered as adequate content validity for

both I-CVI and S-CVI (Polit et al., 2007). I-CVI was calculated by dividing the number of experts who had a rating of either 3 or 4 by the total number of experts (Polit et al., 2007). According to the results of I-CVI, one of the items was eliminated (CVI < 0.8) (see Table 1). After calculating I-CVI, the S-CVI was also calculated by dividing the number of items that the experts rated 3 or 4 by the total number of items.

In order to examine the face validity of items, the test was administered to 20 typically-developing Persian-speaking children aged 4–6 years. Items that more than 30% of the participants did not answer were identified as difficult to understand and were deleted, and those that were not answered by less than 30% were revised and modified.

Step 2: Item analysis of the first version of the test. The first version of the test contained 74 items, and was administered to 100 children. Children's scores for the test items were analysed based on the three parameters of difficulty coefficient, discrimination coefficient, and Spearman coefficient (determining the correlation of each item with the total test score). In this study, difficulty coefficient of 0.15–0.85, correlation coefficient of higher than 0.2, and discrimination coefficient of higher than 0.5 were considered as acceptable levels for retaining items (Anastasi, 1998). Items for which three parameters or only two parameters were at an acceptable level were retained in the second version of the test. The second version of the test consisted of 52 items (see Table 1).

Step 3: Evaluating validity and reliability

- **Construct Validity:** Data obtained from administering the second version of the test on 400 participants underwent further item analysis, as described for Step 2, and 12 further items were eliminated (see Table 1). Therefore, the final version of the test contained 40 items, and the grammatical structures included were question–word interrogative sentences; exclamations, conditional, and yes–no interrogative sentences; negation; copulas; passive form; causal structure; regular and irregular past tense, pronouns (personal, demonstrative, exclamatory); prepositions; causal conjunctions; coordinating; comparative and superlative adjective; relative clauses (subject, adverb, and complement); bound subjects; genitive case; tense, mood, and aspect of verb. The participants' responses to these 40 items were analysed to establish the test's psychometric properties. In order to determine the construct validity, the test's ability to discriminate between different age and gender groups was examined.
- **Convergent Validity:** Additionally, in order to determine the convergent validity of the construct validity, the final version of the Photographic Expressive Persian Grammar Test (40 items) and the grammatical complement subtest of TOLD-P3, in which 24 incomplete sentences are presented to the child for completion, were administered to 100 Persian-speaking children of 4–6 years of age and the degree of correlation between the results obtained from both tests was calculated.
- **Test–retest Reliability:** In order to examine test–retest reliability, the Photographic Expressive Persian Grammar Test was administered twice, with an interval of 3 weeks, to 25 children, randomly selected from the 400 participants.
- **Inter-rater Reliability:** The correlation between the scores obtained by two examiners (both with a bachelor's degree in speech and language therapy) was calculated for 15 children selected randomly from the research population.
- **Internal Consistency:** In order to determine internal consistency, the correlation between items was calculated by using Kuder–Richardson 21.

Table 1. Item reduction procedure at different stages of test design.

Stage	Number of items	Number of eliminated items	Number of remaining items
Content validity	75	1	74
Face validity	74	–	74
First item analysis	74	22	52
Second item analysis	52	12	40

3 Test administration and scoring

Examiners administering, scoring, and interpreting the Photographic Expressive Persian Grammar Test had a thorough understanding of child language development, particularly morphology and syntax. They had training in evaluation procedures.

The test was administered in a well-lit room free from auditory and visual distractions. Table height was appropriate for the child to see the stimulus pictures easily. During the test administration, the examiner placed the photographic stimulus book in front of the child, and the response form in front of himself or herself, not easily visible to the child, and then presented the eliciting sentences.

Every test question had one correct and one wrong answer in the response form. If the child chose the answer indicated on the response form, a circle was drawn around the answer. Otherwise, the child's answer (correct or wrong) was transcribed. For some items, the aim was to evaluate the expression of grammatical morphemes, if the child expressed that morpheme in combination with a word other than the word targeted by the test; his or her answer would be recorded and considered as a correct answer. Concerning other grammatical structures, such as conjunctions, which are variegated in Persian, some can be used interchangeably; if the answer a child provided was not the one anticipated, but similar to the target answer in terms of meaning and of the same structure, it was considered a correct answer. If the child gave no answer, the examiner could once use the prompt sentence without referring to the target structure.

If a child did not answer even after the prompt sentence was repeated, the examiner put a dash (–) in the last column of the response form. The child's correct answers were indicated by a check mark (✓) and wrong answers by a cross mark (✗) in the last column on the response form. Finally, the examiner added up all correct items to calculate the raw score. The time required for administering the test is 10–15 minutes.

4 Data analysis

Statistical tests used included the Spearman correlation coefficient (item analysis), intra-class correlation coefficient (test–retest reliability, inter-rater reliability), Pearson's correlation coefficient (convergent validity), Mann–Whitney (gender discriminative validity), and Kruskal–Wallis (age discriminative validity).

III Results

1 Content validity

The content validity of each item was between 0.8 and 1, and the content validity of the scale was 0.86 both for relevancy of the target sentence to the eliciting sentence and relevancy of the photograph to the target sentence.

2 Face validity

No items were identified as difficult to understand, and only 12 items were revised and modified because less than 30% of participants did not answer those. Therefore, face validity of the items was considered to be good.

3 Psychometric properties of the final version of the test

Descriptive statistics of the 400 children in each age group in terms of gender is presented in Table 2. There was a significant difference between the mean of scores of girls and boys ($p < 0.0001$, $Z = -3.73$). As Table 3 shows, the girls' mean score is higher than that of the boys. The test discriminated for age with significant differences between the mean scores of the four age groups ($p < 0.0001$, $df = 3$). As shown in Table 4, children's scores increase with their age (6-month interval). To determine convergent validity, scores obtained from Photographic Expressive Persian Grammar Test and those obtained from the subtest of TOLD-P3 were significantly correlated ($r = 0.50$, $p < 0.0001$).

Table 2. Descriptive statistics of participants by gender in four age groups ($n = 400$) (percentages are given in parentheses).

Age group	Female	Male	Total
48–54 months	53 (54.1)	45 (45.9)	98 (24.5)
55–60 months	49 (48.5)	52 (51.5)	101 (25.2)
61–66 months	49 (47.6)	54 (52.4)	103 (25.8)
67–72 months	48 (49.0)	50 (51.0)	98 (24.5)
Total	199 (49.8)	201 (50.2)	400 (100)

Table 3. Mann–Whitney test to compare mean scores of Photographic Expressive Grammar Test by gender (maximum Score = 40).

Gender	Range	Mean (SD)	Z	Significance level
Male	5–36	24.28 (6.42)	-3.73	$p < 0.0001$
Female	10–40	26.64 (5.84)		

Notes. $Z = -3.73$; $p < 0.0001$.

Table 4. Kruskal–Wallis test to compare mean scores of Photographic Expressive Grammar Test by age (maximum Score = 40).

Age groups	Range	Mean (SD)
48–54 months	5–35	20.84 (5.76)
55–60 months	10–35	24.55 (6.22)
61–66 months	10–40	27.11 (5.12)
67–72 months	16–38	29.27 (4.44)

Notes. Total mean (SD) = 25.46 (6.24); $X^2 = 102.7$; $p < 0.0001$.

4 Test reliability

Results of Kuder–Richardson Test 21 on internal consistency showed that there is a high correlation between test items ($r = 0.82$). Test–retest reliability was evaluated and a significant correlation was found between participants' scores at two different assessment times ($r = 0.91$, $p < 0.0001$). Results related to inter-rater reliability indicated that there was a significant correlation between that the scores of two different examiners ($r = 0.98$, $p < 0.0001$).

IV Discussion

One of the most important features of the Photographic Expressive Persian Grammar Test is its exclusiveness in evaluating morphosyntactic characteristics of Persian-speaking children compared to other instruments existing in Iran. Another feature of this test is that similar to the SPELT-3, it is designed according to an evoked-spontaneous procedure in which instances of everyday life are illustrated through color photographs with the aim of prompting particular spoken responses. As this method can be used for analysing certain linguistic structures such as grammatical morphemes, which do not usually appear in free, open-ended speech, it is an effective clinical instrument (Rew and Irwin, 1985). The results suggest that the Photographic Expressive Persian Grammar Test is a reliable and valid tool. Moreover, the time required for administering this test is 10–15 minutes and, unlike the method of speech sampling, analysis of data is obtained within a short period time. The time required for administering this test is shorter than for SPELT-3 and longer than for TEEM. Since TEEM only addresses morphological skills, but the Photographic Expressive Persian Grammar Test evaluates both morphological and syntactic skills, this difference in the time required for administering the two tests is rational. While both SPELT-3 and the Photographic Expressive Persian Grammar Test evaluate children's morphological and syntactic skills, SPELT-3 has 54 items whereas the Photographic Expressive Persian Grammar Test contains 40 items; as a result, administering SPELT-3 takes longer (Perona et al., 2005).

In the present study, high correlation between test items suggests that the test measures a single variable; in other words, it evaluates only children's morphosyntactic properties. In addition, strong inter-rater reliability was found between the scoring of two examiners. Therefore, the participant's score is not influenced by examiners' bias. Test–retest reliability of 0.91 was also found; this indicates the consistency and reliability of test scores through time. Thus, if this test is administered several times to a participant, his or her score will be the same. In SPELT-3, inter-rater reliability was reported 0.97 to 0.99, and test–retest reliability was reported 0.94. Therefore, results obtained in this study were similar to those of the SPELT-3. Results of inter-rater reliability of the Photographic Expressive Persian Grammar Test also matched with those of the TEEM (0.94) (Paul and Norbury, 2012), and the TEGI had a high correlation on the retest study but a higher test–retest reliability compared with the NSST (0.69 for the expressive portion) (Davis, 2010).

In determining content validity of test, experts' scores showed that these test items can well represent and capture the morphosyntactic qualities of 4–6-year-old children. In order to determine the convergent validity of SPELT-3, correlation between the results of this test and those of another standard test in this field (TEEM) and the Test of Language Development: Grammatical Understanding Subtest – Primary 3rd edition on 42 children was determined. The degree of correlation reported was 0.82 between SPELT-3 and TEEM and 0.51 with TOLD-GU (Perona et al., April 2005). Convergent validity was also reported for TEEM, 0.87 (Paul and Norbury, 2012) and TEGI, 0.48 (Davis, 2010). In the present study, convergence between the scores of the Photographic Expressive Grammar Test and those of the subtest of TOLD-P3 was examined and the correlation between these two tests was found to be 0.50. Overlapping of scores between these two tests indicates that the performance of the Photographic Expressive Persian Grammar Test reflects a profile

of the morphosyntactic properties of children's expressive language. Since the items related to the subtest of TOLD-P3 were only designed based on the syntactic properties of Persian and children's morphosyntactic characteristics were ignored, this might account for the low correlation coefficient of the convergent validity of the test under study compared to the convergent validity of SPELT-3 and TEEM. Nonetheless, results of the convergent validity the Photographic Expressive Persian Grammar Test matched with those of the TEGI.

Investigating the effect of gender on participants' scores indicated a significant difference in scores of male and female genders such that girls achieved higher scores than boys. Numerous studies conducted on the effect of gender on linguistic skills of children support this finding that girls perform better than boys in terms of verbal abilities (lexical fluency, grammar, spelling, reading, vocabulary repertoire, and comprehension) (Halpern, 2013). Therefore, results from the present study correspond with these findings.

Use of morphosyntactic structures in expressive language is a construct that is developmental in nature, i.e. there is an expectation that, in the presence of normal development, as the child gets older and matures, use of his or her morphosyntactic structures should improve. Further, this development occurs more rapidly in earlier than later years, and there is more variability in the rate of development among younger children than in older children (Dawson et al., 2003). Accordingly, to determine the developmental nature of the Photographic Expressive Persian Grammar Test, the effect of age on the scores of children was studied in four groups. The results showed the high construct validity of the test in regard to age group discrimination. In other words, as expected, the average of test scores increased with children's age. Another consideration to be noted is that this difference is observed in age groups with a difference of six months of age.

V Conclusions

This study was conducted with the aim of design and validation of the Photographic Expressive Persian Grammar Test for children aged 4–6 years. This test is the first valid and reliable tool that exclusively evaluates the morphosyntactic properties of Persian-speaking children aged 4–6 years with its 40 items compared to other tests in Iran and identify their strengths and weaknesses. Also, this valid test can be used in future as a reference in studies in the field of Persian syntax and morphology.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

This work was supported by the Pediatric Neurorehabilitation Research Center, Tehran, Iran. This article is part of a PhD thesis approved in University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

References

- Anastasi A (1998) *Psychological testing*. New York: Macmillan.
- Angell CA (2009) *Language development and disorders: A case study approach*. Sudbury, MA: Jones and Bartlett.
- Anvari H and Givi HA (2012) *Persian Grammar 2*. Tehran: Fatemi.
- Bakhtiar M, Nilipour R, and Weekes BS (2013) Predictors of timed picture naming in Persian. *Behavior Research Methods* 45: 834–41.
- Bishop D (2006) What causes specific language impairment in children? *Current Directions in Psychological Science* 15: 217–21.

- Brown R (1973) *A first language: The early stages*. Cambridge: Harvard University Press.
- Carrow-Woolfolk E (1999) *Comprehensive Assessment of Spoken Language (CASL)*. Circle Pines, MN: American Guidance Service.
- Clark E (2003) *First language acquisition*. Cambridge: Cambridge University Press.
- Davis AS (2010) *Handbook of pediatric neuropsychology*. New York: Springer.
- Dawson JI, Stout CE, and Eyer JA (2003) *Structured Photographic Expressive Language Test-III Manual*. DeKalb, IL: Janelle.
- Elwell-Sutton LP (1992) *Elementary Persian grammar*. Cambridge: Cambridge University Press.
- Ervin SM and Miller WR (1963) Language development. In: Stevenson HW (eds) *Child psychology: The sixty-second yearbook of the National Society for the Study of Education: Part I*. Chicago, IL: University of Chicago Press, 108–43.
- Ghasisin L, Ahmadi T, Mostajeran F, Moazam M, and Derakhshande F (2013) Evaluating the reliability and validity of phonetic information test in normal 5–6-year-old children of Isfahan city. *Journal of Research in Rehabilitation Sciences* 9: 153–160.
- Givi HA and Anvari H (2012) *Persian grammar 1*. Tehran: Fatemi.
- Golpou L, Nilipour R, and Roshan BA (2006) Comparison between morphological and syntactic features of 4 to 5 years old in education severe to profound hearing impaired and normal children. *Audiology Journal* 15: 3–29.
- Halpern DF (2013) *Sex differences in cognitive abilities*. New York: Taylor & Francis.
- Hasanati F, Agharasouli Z, Mahmoudi Bakhtiari B, and Kamali M (2011) Sentence repetition test for measurement of grammatical development in Farsi-speaking children. *Audiology Journal* 20: 73–81.
- Hasanzadeh S and Minaii A (2002) Adaptation and normalization of Language Development Test TOLD-P:3 for Persian-speaking children of Tehran: Part 2. *Research on Exceptional Children* 1: 119–34.
- Hersen M (2004) *Comprehensive handbook of psychological assessment, intellectual and neuropsychological assessment*. Hoboken, NJ: John Wiley.
- Hinkin TR (1995) A review of scale development practices in the study of organizations. *Journal of Management* 21: 967–88.
- Hoff E (2013) *Language development*. Boston, MA: Cengage Learning.
- Jafari Z and Asad-Malayeri S (2012) The psychometric properties of Newsha Developmental Scale: An integrated test for Persian speaking children. *Iranian Journal of Pediatrics* 22: 28–34.
- Jalilevand N (2012) *Speech and language development in Farsi-speaking children*. Tehran: Danjeh.
- Jalilevand N and Ebrahimipour M (2013) Pronoun acquisition in Farsi-speaking children from 12 to 36 months. *Journal of Child Language Acquisition and Development* 1: 1–9.
- Jalilevand N, Ebrahimipour M, and Pourqarib J (2011) Mean length of utterance and grammatical morphemes in speech of two Farsi-speaking children. *Audiology Journal* 21: 96–108.
- Jones SW (2013) *A grammar of the Persian language*. Lenox, MA: HardPress.
- Kaderavek JN (2014) *Language disorders in children: Fundamental concepts of assessment and intervention*. San Francisco, CA, USA: Pearson.
- Kazemi Y, Stringer H, and Klee T (2015) Study of child language development and disorders in Iran: A systematic review of the literature. *Journal of Research in Medical Sciences* 20: 66–77.
- Kent RG, Emeneau MB, and Cammann S (2011) *Old Persian, grammar, texts, lexicon: American Oriental series*. Maryland Heights, MO: Literary Licensing, LLC.
- Khanlari P (2014) *Persian Grammar*. Tehran: Toos.
- Mace J (2015) *Persian Grammar: For reference and revision*. New York: Taylor & Francis.
- Mahootian S (2003) *Persian*. New York: Taylor & Francis.
- Maleki T, Soleimani Z, and Faghizadeh S (2012) Investigating linguistic performances in Persian-speaking children with specific language impairment. *Audiology Journal* 20: 12–21.
- Meshkatoddini M (2003) *Persian Grammar based on transformation theory*. Mashhad: Ferdowsi University of Mashhad.
- Meshkatoddini M (2004) Investigating the development of inflectional affixes and connective structures in developmental articulation of children. *Journal of Linguistics* 19: 10–26.
- Miremadi A (2010) *Persian syntax based on government and binding theory*. Tehran: SAMT.

- Paul R and Norbury CF (2012) *Language disorders from infancy through adolescence: Assessment and intervention*. St Louis, MO: Mosby.
- Perona K, Plante E, and Vance R (2005) Diagnostic accuracy of the Structured Photographic Expressive Language Test (SPELT-3). 3rd edition. *Language, Speech and Hearing Services in Schools* 36: 103–15.
- Perry JR (2005) *A Tajik Persian reference grammar*. Leiden-Boston: Brill.
- Polit DF, Beck CT, and Owen SV (2007) Focus on research methods is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing and Health* 30: 459–67.
- Qasemi M, Nakhshab M, Alineghad B, and Shafiei M (2012) Description of particular language structures in normal 4-year-olds' narrations and their sex-related differences according to Narrative Assessment Protocol (NAP). *Journal of Research in Rehabilitation Sciences* 8: 1–15.
- Rew CJ and Irwin DL (1985) Assessment of morphological production utilizing two language sampling procedures. *National Student Speech–Language–Hearing Association* 13: 10–19.
- Rice ML and Wexler K (2001) *Rice Wexler Test of Early Grammar Impairment Examiner's Manual*. San Antonio, TX: The Psychological Corporation.
- Schwab DP (1980) Construct validity in organizational behavior. *Research in Organizational Behavior* 2: 3–43.
- Semel E and Wiig EH (1980). *Clinical Evaluation of Language Fundamentals*. San Antonio, TX: The Psychological Corporation.
- Shipley K and McAfee J (2009) *Assessment in speech-language pathology: A resource manual*. Boston, MA: Cengage Learning.
- Walt JL and Steyn H (2008) The validation of language tests. *Stellenbosch Papers in Linguistics* 38: 191–204.
- Windfuhr GL (1979) *Persian Grammar: History and state of its study*. Germany: Mouton.
- Yousef S and Torabi H (2013) *Basic Persian: A grammar and workbook*. New York: Routledge.
- Zarifian T, Mohammadi R, and Mahmoudi-Bakhtiari MB (2013) Comparative study of the syntactic skills of hearing and hearing impaired Persian-speaking pre-school children. *Research in Rehabilitation Sciences* 8: 1–12.