

Shortened Version of the Token Test: Normative data for Spanish-speaking pediatric population

L. Olabarrieta-Landa^a, D. Rivera^b, A. Rodríguez-Lorenzana^c, S. Pohlenz Amador^d,
C.E. García-Guerrero^e, A. Padilla-López^f, M. Sánchez-SanSegundo^g, J. Velázquez-Cardoso^h,
J.P. Díaz Maranteⁱ, R.A. Caparros-Gonzalez^j, I. Romero-García^k, J. Valencia Vásquez^l,
C. García de la Cadena^m, J.M. Muñoz Mancillaⁿ, B.V. Rabago Barajas^o, R. Barranco Casimiro^p,
A. Galvao-Carmona^q, P. Martín-Lobo^r, M. Saracosti Schwartzman^s and J.C. Arango-Lasprilla^{b,t,*}

^a*Faculty of Psychology and Education, University of Deusto, Bilbao, Spain*

^b*BioCruces Health Research Institute, Cruces University Hospital, Barakaldo, Spain*

^c*Escuela de Psicología, Universidad de Las Américas, Quito, Ecuador*

^d*Escuela de Ciencias Psicológicas, Universidad Nacional Autónoma de Honduras, Tegucigalpa, Honduras*

^e*Mindpedia Centro de Psicología Avanzada, Monterrey, México*

^f*Laboratorio de Psicofisiología, Facultad de Ciencias Humanas, Universidad Autónoma de Baja California, Mexicali, México*

^g*Department of Health Psychology, University of Alicante, Spain*

^h*Instituto Nacional de Neurología y Neurocirugía, MVS, Ciudad de México, México*

ⁱ*Facultad de Ciencias Médicas Finlay-Albarrán, Havana, Cuba*

^j*CIMCYC-The Mind, Brain and Behaviour Research Centre, Universidad de Granada, Granada, Spain*

^k*Universidad Interamericana de Puerto Rico, Recinto de San Germán, Puerto Rico*

^l*Universidad Católica San Pablo, Arequipa, Perú*

^m*Departamento de Psicología, Universidad del Valle de Guatemala, Guatemala City, Guatemala*

ⁿ*Universidad Autónoma de Asunción (UAA), Asunción, Paraguay*

^o*Departamento de Neurociencias, Universidad de Guadalajara (CUCS), Guadalajara, México*

^p*Research Center CERNEP, Almeria University, Almería, Spain*

^q*Department of Psychology, Universidad Loyola Andalucía, Sevilla, Spain*

^r*Universidad Internacional de la Rioja (UNIR), Logroño, Spain*

^s*Universidad de la Frontera, Temuco, Chile*

^t*IKERBASQUE, Basque Foundation for Science, Bilbao, Spain*

Abstract.

OBJECTIVE: To generate normative data for the Shortened Version of the Token Test in Spanish-speaking pediatric populations.

METHOD: The sample consisted of 4,373 healthy children from nine countries in Latin America (Chile, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Puerto Rico) and Spain. Each participant was administered the Shortened Version of the Token Test as part of a larger neuropsychological battery. Shortened Version of the Token Test total scores were normed using multiple linear regressions and standard deviations of residual values. Age, age², sex, and mean level of parental education (MLPE) were included as predictors in the analyses.

*Address for correspondence: Juan Carlos Arango Lasprilla, Ph.D., BioCruces Health Research Institute, Cruces University Hospital, IKERBASQUE, Basque Foundation for Science, Plaza

de Cruces s/n. 48903, Barakaldo, Bizkaia, Spain. Tel.: +34 946006000/Ext. 7963; E-mail: jcalasprilla@gmail.com.

RESULTS: The final multiple linear regression models showed main effects for age in all countries, such that score increased linearly as a function of age. In addition, age² had a significant effect in all countries, except Guatemala and Puerto Rico. Models showed that children whose parent(s) had a MLPE >12 years obtained higher score compared to children whose parent(s) had a MLPE ≤12 years in Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, and Spain. The child's sex did not have an effect in the Shortened Version of the Token Test total score for any of the countries.

CONCLUSIONS: This is the largest Spanish-speaking pediatric normative study in the world, and it will allow neuropsychologists from these countries to have a more accurate interpretation of the Shortened Version of the Token Test when used in pediatric populations.

Keywords: Shortened Version of the Token Test, neuropsychology, Spanish-speaking populations, pediatric population

1. Introduction

The evaluation of children's linguistic abilities is crucial for the detection of possible language disorders (e.g., dyslexia, dysgraphia, specific language impairment). These disorders are especially common during infancy, with prevalence rates between 2 and 25% (Law, Boyle, Harris, Harkness, & Nye, 2000; McLeod & Harrison, 2009; Tomblin et al., 1997). Early detection allows for professional intervention to prevent the potential negative effects caused by these disorders. As indicated by Bishop & Edmundson (1987), the longer these disorders persist, the greater the possibility of continued difficulty in these and related areas. For example, studies have demonstrated that children with these disorders have a greater possibility of reading difficulty (Schuele, 2004), mental health issues (e.g., depression, anxiety; Beitchman et al., 2001; Conti-Ramsden, & Botting, 2008; Sundheim & Voeller, 2004), and behavioral problems (Beitchman et al., 2001; Willinger et al., 2003) when compared to children without a history of language disorders. These challenges in turn lead to difficulties with school (Schuele, 2004) and in social relationships, especially bullying or exclusion by peers (Hughes, 2014; Snowling, Bishop, Stothard, Chipchase, & Kaplan, 2006).

One of the most utilized tests to detect language problems, specifically related to receptive language, is the Token Test (Gallardo, Guárdia, Villaseñor, & McNeil, 2011; Malloy-Diniz et al., 2007; Wassenberg et al., 2008). The test was originally developed in 1962 by De Renzi and Vignolo with the purpose of detecting language deficits in adults with aphasia, however, it has also been applied to children and adolescents given the utility of the test (Malloy-Diniz et al., 2007; Paquier et al., 2007). The original test consists of 62 commands that increase in difficulty, and uses blue, white, red, green, and yellow rectangular and circular tokens (De Renzi & Vignolo, 1962).

The Token Test is widely used in many Latin American countries and Spain. Indeed, across multiple studies describing the state of neuropsychology in different countries, the Token Test is one of the 20 most utilized neuropsychological tests by professionals in Spain (Olabarrieta-Landa, et al., 2016), Colombia (Arango-Lasprilla et al., 2015), Mexico (Fonseca-Aguilar et al., 2015), Argentina (Fernandez, Ferreres, Morlett-Paredes, Rivera, & Arango-Lasprilla, 2016), and other Latin American countries (Arango-Lasprilla, Stevens, Morlett-Paredes, Ardila, & Rivera, 2016). This widespread acceptance may be due to its simplicity, easy administration, and free access.

Due to its popularity, the original test has been adapted in several ways, including the revised version (McNeil & Prescott, 1978), the short form of 16 items (Spellacy & Spreen, 1969), as well as computerized versions (McNeil et al., 2015; Turkyilmaz & Belgin, 2012). However, the only modified and shortened version developed by the original authors of the test is Shortened Version of the Token Test created in 1978 (De Renzi & Faglioni, 1978). This test includes 36 commands, an introductory section that contains seven items of minimal difficulty (e.g., "Touch the circle"), and modification of the stimuli with square tokens replacing rectangular tokens and the color black replacing blue (Renzi & Faglioni, 1978). These changes were put in place based on clinician observations that individuals with brain damage had difficulty discriminating the color blue from green (Renzi & Faglioni, 1978). As a result, the stimuli now consist of 20 tokens, 10 circles and 10 squares in black, white, red, yellow, and green (Renzi & Faglioni, 1978).

Two major motives led to the creation of this version. First, there was a desire to offer the scientific community a conventional form of the test that would serve as a reference when discussing the results of different studies, since, as discussed; there are a variety of versions (e.g. different stimuli, different scoring

calculations, the reduction in the number of commands, etc.). Second, test developers also wanted to offer a version that would reduce fatigue in the person being evaluated without losing the ability to discriminate between people with and without aphasia, and the detection of changes in performance that occur in children, especially as they increase in age (Renzi & Faglioni, 1978). For these reasons, Renzi and Faglioni (1978) recommend the use of the short form of 36 commands.

With regard to normative data, despite the success of the test in Latin American countries and Spain, few studies offer normative data for these populations, especially for the version recommended by the authors of the test. The only studies are those conducted by Peña-Casanova et al. (2009) and Aranciva et al. (2012), both of which are part of the NEURONORMA project and offer normative data for the Spanish adult and young adult population, but do not offer normative data for children and adolescents. Therefore, the purpose of the present study is to determine normative data for the child/youth population from nine Latin American countries (Chile, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Puerto Rico) and Spain.

2. Method

2.1. Participants

The sample consisted of 4,373 healthy children who were recruited from Chile, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, and Spain. Participants were selected according to the following criteria: a) were between 6 and 17 years of age, b) were born and currently lived in a country where the study was conducted, c) Spanish as primary language, d) an $IQ \geq 80$ on the Test of Non-Verbal Intelligence (TONI-2, Brown, Sherbenou, & Johnsen, 2009), and e) a score < 19 on the Children's Depression Inventory (CDI, Kovacs, 1992).

Children with history of neurologic or psychiatric disorders as reported by the participant's parent(s), were excluded due to its effects on cognitive performance. Participants in the study were from public or private schools, and signed an informed consent. Socio-demographic and participant characteristics for each of the countries' samples have been reported elsewhere (Rivera & Arango-Lasprilla, 2017). Ethics Committee approval was obtained for the study in each country.

2.2. Instrument administration

The Shortened Version of the Token Test consists of 20 plastic tokens; ten circles and ten squares. Five circles and squares are large, and five circles and squares are small. The 36-item test is organized in six parts, which refer to the verbal orders that the examinee must understand and complete. If following each item of part 1–5 the participant responds incorrectly, the examiner returns the tokens to their original order, and repeats the item. One point is credited for a correct performance on the first presentation and 0.5 point if the performance is correct only on the second presentation. In the last part, however, there is only one possibility to respond, getting 1 point if correct, and 0 if incorrect (De Renzi & Faglioni, 1978). Appendix 1 shows the Spanish Shortened Version of the Token Test used in this study.

2.3. Statistical analyses

Detailed statistical analyses used to generate the normative data for the Shortened Version of the Token Test are described in Rivera & Arango-Lasprilla (2017). In summary, the scores were standardized using multiple linear regression analyses by means of a four-step procedure. First, the Shortened Version of the Token Test total score was computed by means of the final multiple regression models. The full regression models included as predictors: age, age², sex, and mean level of parental education (MLPE). Age was centered (=calendar age – mean age in the sample by country) before computing the quadratic age term to avoid multicollinearity (Aiken & West, 1991). Sex was coded as male = 1 and female = 0. The MLPE variable was coded as 1 if the participant's parent(s) had >12 years of education or 0 if participant's parent(s) had ≤12 years of education. If predictor variables were not statistically significant in the multivariate model with an alpha of 0.05, the non-significant variables were removed and the model was run again. A final regression model was conducted: $\hat{y}_i = B_0 + B_1 \cdot (Age - \bar{x}_{Age \text{ by country}})_i + B_2 \cdot (Age - \bar{x}_{Age \text{ by country}})_i^2 + B_3 \cdot Sex_i + B_4 \cdot MLPE_i$. Second, residual scores were calculated based on the final model ($e_i = y_i - \hat{y}_i$). Third, Residuals were standardized using the residual Standard Deviation (SD_e) value provided by the regression model: $z_i = e_i / SD_e$. Fourth, standardized residuals were converted to percentile values using the standard normal cumulative distribution function. This four-

step process was applied to Shortened Version of the Token Test total score separately for each country.

For all multiple linear regression models, the following assumptions were evaluated: a) multicollinearity by the values of the Variance Inflation Factor (VIF), which must not exceed 10, and the collinearity tolerance values, which must not exceed the value of 1 (Kutner, Nachtsheim, Neter, & Li, 2005), and b) the existence of influential values by calculating the Cook's distance. The maximum Cook's distance value was related to a $F(p, n - p)$ distribution. Influential values are considered when percentile value is equal or higher than 50 (Cook, 1977; Kutner et al., 2005). All analyzes were performed using SPSS version 23 (IBM Corp., Armonk, NY).

3. Results

The final multivariate linear regression models for the ten country-specific Shortened Version of the Token Test total scores were significant (see Table 1). In all countries, the Shortened Version of the Token Test total score increased linearly as a function of age. The Shortened Version of the Token Test total score for Chile, Cuba, Ecuador, Honduras, Mexico, Paraguay, Peru, and Spain was also affected by a quadratic age effect. Children from all countries except for Chile and Cuba who had parent(s) with a MLPE >12 years obtained higher scores than children who had parent(s) with a MLPE ≤12 years. The child's sex did not affect the Shortened Version of the Token Test total score for any country. The amount of variance these predictors explained in Shortened Version of the Token Test total scores ranged from 22.1% (in Puerto Rico) to 52.8% (in Cuba). The assumptions of multiple linear regression analysis were met for all final models. There was not multicollinearity (the VIF values were below 10; VIF ≤1.060; collinearity tolerance values did not exceed the value of 1) or influential cases (the maximum Cook's distance value was 0.216 in a $F_{(3,200)}$ distribution which correspond to percentile 11).

3.1. Normative procedure

Norms (e.g., a percentile score) for the Shortened Version of the Token Test total score by country were established using the four-step procedure described in the statistical analysis section. An example will be provided to facilitate an improved understanding of

the procedure used to obtain the percentile associated with a score on this test. Let's assume we need to find the percentile score for an 8-year-old Mexican girl who scored a 33 on the Shortened Version of the Token Test and whose parent(s) have a mean of 14 years of education (MLPE).

The steps to obtain the percentile for this score are: First, find Mexico in Table 1, which provides the final regression models by country for the Shortened Version of the Token Test total score. Use the B weights to create an equation that will allow you to obtain the predicted Shortened Version of the Token Test total score for this child using the coding provided in the statistical analysis section. The corresponding B weights are multiplied by the centered age (= calendar age – mean age in the Mexican sample which is equal to 11.4 years), centered age² (=calendar age – mean age in the Mexican sample which is equal to 11.4 years)², MLPE code based on the 12 years of education threshold. See Rivera & Arango-Lasprilla (2017) to figure out the mean age of each country's sample. Then the result is added to the constant generated by the model in order to calculate the predicted value. Child's sex was not a significant predictor, and therefore is not included in this model.

In the case of the Mexican girl, the predicted Shortened Version of the Token Test total score would be calculated using the following equation: $\hat{y}_i = 32.768 + [0.367 \cdot (Age_i - 11.4)] + [-0.040 \cdot (Age_i - 11.4)^2] + (0.587 \cdot MLPE)$. The girl's age is 8. The MLPE (14 years) is split into either ≤12 years (and assigned a 0) or more than 12 years (and assigned a 1) in the model. Since the parent(s) of the hypothetical child in the example have 14 years of education, the MLPE value is 1. Thus, the predicted value equation is: $\hat{y}_i = 32.768 + [0.367 \cdot (8 - 11.4)] + [-0.040 \cdot (8 - 11.4)^2] + (0.587 \cdot 1) = 32.768 + (-1.247) + (-0.462) + 0.587 = 31.646$.

Second, in order to calculate the residual value (indicated with an e_i in the equation), we subtract the actual Shortened Version of the Token Test score (she scored 33) from the predicted value we just calculated ($e_i = y_i - \hat{y}_i$). In this case, it would be $e_i = 33 - 31.646 = 1.354$.

Third, consult the SD_e column in Table 1 to obtain the country-specific SD_e (residual) value. For Mexico, it is 2.156. Using this value, we can transform the residual value to a standardized z score using the equation $z_i = e_i / SD_e$. In this case, we have $1.354 / 2.156 = 0.628$. This is the standardized z score for an 8-year-old Mexican girl who scored a

Table 1
Final multiple linear regression models for Shortened Version of the Token Test

Country	B	Std. Error	<i>t</i>	Sig.	<i>R</i> ²	<i>SDe</i> (residual)
Chile						
Constant	33.105	0.191	173.720	< 0.001	0.274	2.478
Age	0.402	0.037	10.976	< 0.001		
Age ²	-0.060	0.012	-4.998	< 0.001		
Cuba						
Constant	34.815	0.103	338.347	< 0.001	0.528	1.326
Age	0.396	0.020	20.088	< 0.001		
Age ²	-0.027	0.006	-4.171	< 0.001		
Ecuador						
Constant	32.136	0.300	107.249	< 0.001	0.247	2.306
Age	0.352	0.039	8.996	< 0.001		
Age ²	-0.033	0.013	-2.529	0.012		
MLPE	1.142	0.307	3.715	< 0.001		
Guatemala						
Constant	30.803	0.193	159.803	< 0.001	0.222	2.352
Age	0.460	0.067	6.892	< 0.001		
MLPE	0.829	0.383	2.163	0.032		
Honduras						
Constant	32.565	0.195	166.672	< 0.001	0.283	1.931
Age	0.369	0.036	10.366	< 0.001		
Age ²	-0.033	0.011	-2.879	0.004		
MLPE	0.462	0.230	2.010	0.045		
Mexico						
Constant	32.768	0.136	240.351	< 0.001	0.288	2.156
Age	0.367	0.020	17.957	< 0.001		
Age ²	-0.040	0.007	-5.876	< 0.001		
MLPE	0.587	0.143	4.088	< 0.001		
Paraguay						
Constant	32.940	0.285	115.461	< 0.001	0.277	2.386
Age	0.394	0.040	9.873	< 0.001		
Age ²	-0.038	0.014	-2.777	0.006		
MLPE	0.717	0.290	2.471	0.014		
Peru						
Constant	33.119	0.269	123.164	< 0.001	0.294	2.771
Age	0.394	0.048	8.246	< 0.001		
Age ²	-0.090	0.016	-5.747	< 0.001		
MLPE	0.643	0.311	2.065	0.040		
Puerto Rico						
Constant	30.202	0.462	65.352	< 0.001	0.221	3.339
Age	0.340	0.068	5.004	< 0.001		
MLPE	3.174	0.542	5.857	< 0.001		
Spain						
Constant	33.605	0.119	282.378	< 0.001	0.315	1.861
Age	0.368	0.018	20.605	< 0.001		
Age ²	-0.034	0.006	-5.982	< 0.001		
MLPE	0.311	0.124	2.504	0.012		

Note. MLPE: Mean level of parental education.

33 on the Shortened Version of the Token Test who has parent(s) with a MLPE of 14 years.

The fourth and final step is to use the tables available in most statistical reference books (e.g., Strauss, Sherman, & Spreen 2006) to convert *z* scores to percentiles. In this example, the *z* score (probability) of 0.628 corresponds to the 73rd percentile.

3.2. User-friendly normative data

The four-step normative procedures explained above offers the clinician the ability to determine

an exact percentile for a child who has a specific score on the Shortened Version of the Token Test. However, this method can be prone to human error due to the number of required computations by hand. To enhance user-friendliness, the authors have completed these steps for a range of raw scores based on age, sex, and MLPE and created tables for clinicians to more easily obtain a percentile range/estimate associated with a given raw score on this test. These tables are available by country in the Appendix.

In order to obtain an approximate percentile for the above example (converting a raw score of 33 on

Shortened Version of the Token Test for a Mexican girl who is 8 years old whose parent(s) have 14 years of education) using the simplified normative tables provided in the Appendix, the following steps must be followed. First, identify the appropriate table ensuring the appropriate country. In this case, the table for the Shortened Version of the Token Test for Mexico can be found in Table A6. Second, select the correct section of the table. The table is divided based on MLPE (≤ 12 vs. more than 12 years of education). Since the parent(s) had 14 years of education, we will use the upper section of the table for >12 years of MLPE. Third, find the appropriate age of the child, in this case, 8 years old. Fourth, look in the 8 years' age column to find the approximate location of the raw score obtained on the test. Within the 8 years' column, the score of 33 obtained by this Mexican girl corresponds to an approximate percentile of 70.

The percentile obtained using this user-friendly table is slightly different than the hand-calculated, more accurate method (73rd vs. 70th) because the user-friendly table is based on a limited number of percentile values. Individual percentiles cannot be presented in these tables due to space limitations. If the exact score is not listed in the column, you must estimate the percentile value from the list of raw scores available.

4. Discussion

The purpose of this study was to obtain normative data for the Shortened Version of the Token Test for children and adolescents from nine Latin American countries (Chile, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Puerto Rico) and Spain. The final regression models for the Shortened Version of the Token Test explained between 22.1% and 52.8% of the variance.

There are only a few normative data studies of the Shortened Version of the Token Test that focus on children and adolescent populations. The majority of studies have normative data for adults (Aranciva et al., 2012; De Renzi & Faglioni, 1978; Peña-Casanova et al., 2009), to the exception of a study conducted by Malloy et al. (2007) who obtained normative data for children between the ages 7 and 10. The limited number of normative studies of the Token Test in children and adolescents focus mostly on the revised version of the Token Test (Gallardo et al., 2011; Kumar, Kumar, & Kumari, 2013). Since the number of subtest and commands varies between versions, it is difficult to

compare the results of the present study with those of these studies.

Age was significantly related to the total score of the Shortened Version of the Token Test, such that score increased linearly as age increased. This gradual increase in performance is consistent with studies of the Shortened Version of the Token Test normative data for children (Malloy-Diniz et al., 2007) and adults (Peña-Casanova et al., 2009). However, there are some normative studies performed in adults where the influence of age on performance was not found (Aranciva et al., 2012; De Renzi & Faglioni, 1978). According to Aranciva et al. (2012), this could be due to the ceiling effect observed in the adolescence. In the present study, it was also found a curvilinear relation between age and the Shortened Version of the Token Test total scores for all countries except Guatemala and Puerto Rico. It appears that, although performance improves with time, this usually stabilizes around the age of 11, where a ceiling effect is observed as described by Aranciva et al. (2012). This is consistent with the idea that children at the age of 11 achieve a performance comparable to that of adults in the Token Test (Rich, 1993).

Other variable associated with the Shortened Version of the Token Test total score was the parents' MLPE. The inclusion of MLPE as a predictive variable in normative data studies is fairly recent, which explains its absence in normative data studies of the Token Test. The first authors to use it as a predictive variable to generate normative data were Van der Elst, Hurks, Wassenberg, Meijjs, & Jolles (2011). In their normative study of verbal fluency tests, Van der Elst and colleagues found that children whose parent(s) had a higher educational level performed better on verbal fluency tests compared to children whose parent(s) had a lower educational level. Even though MLPE is not a common variable used in normative studies, this is a very important variable that needs to be considered in studies of child development. Research has shown that socioeconomic status, often measured by the educational level of the mother, influences the development of both receptive and expressive language of the child (Hoff, 2006). More specifically, studies have determined the central role of the mother's education as a predictor of child language development (Dollaghan et al. 1999; Hoff, 2003; Pancsofar & Vernon-Feagans, 2006), although the influence of the father's education has also been emphasized (Pancsofar & Vernon-Feagans, 2006; Pancsofar, Vernon-Feagans, & Family Life Project Investigators, 2010).

Finally, the gender variable was not associated with performance on the Shortened Version of the Token Test. These results are consistent with normative studies of this version of the Token Test in children (Malloy-Diniz et al., 2007) and adults (Aranciva et al., 2012, Peña-Casanova et al., 2009), as well as in normative studies of the revised version of the Token Test in children (Gallardo et al., 2011). Several studies have demonstrated the absence of gender differences in children's language skills. For example, Ardila and Rosselli (1994) found no gender differences in the performance of boys and girls in a neuropsychological battery that included the Token Test. In addition, meta-analysis studies on gender differences in language skills between boys and girls have suggested that these differences are minimal and therefore non-existent (Hyde & Linn, 1988).

4.1. Limitations and future directions

The study has some limitations. First, only children whose mother tongue was Spanish were recruited. Therefore, the results of this study cannot be generalized to children whose mother tongue was different. Since many countries in Latin America and Spain are multilingual (Chamoreau, 2014; Garrido Medina, 2007), future studies should develop norms in the languages spoken in these countries.

Second, the results of this study cannot be generalized to other Latin American countries such as Argentina, Bolivia, Panama, or Venezuela, among others, and future studies should obtain normative data for the Shortened Version of the Token Test in these countries.

Third, despite the fact that in Chile, Mexico, Paraguay, Puerto Rico, and Spain the sample was collected from several areas of the country, for the rest of the countries of the study only children from a specific area were recruited. It would have been ideal to recruit children from different areas of these countries as well. Future studies should expand the data from this study with samples from different areas. Likewise, most of the children were recruited from urban areas, thus future studies should include more children from rural areas.

Finally, normative data were generated using healthy children and adolescents. Future studies should include clinical samples (e.g., children and adolescents with brain damage, epilepsy, etc.) in order to obtain proper cutoff points and calculate the sensitivity and specificity of the Shortened Version of the Token Test.

4.2. Implications and conclusions

The evaluation of language skills in children is of vital importance to detect possible language disorders, whose prevalence is relatively high during childhood. One of the most commonly used tests for detection of language problems is the Shortened Version of the Token Test, which is widely used in Latin America and Spain. The present study has developed normative data for the Shortened Version of the Token Test for nine countries in Latin America and Spain, so now, professionals have the norms to evaluate and diagnose receptive language problems using this test. These norms are expected to be useful for all those who work in the neuropsychological evaluation of children and adolescents in these countries.

Conflict of interest

None to report.

Supplementary material

The Appendix tables are available in the electronic version of this article: <http://dx.doi.org/10.3233/NRE-172244>.

References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Aranciva, F., Casals-coll, M., Sánchez-Benavides, G., Quintana, M., Manero, R. M., Rognoni, T., . . . & Peña-Casanova, J. (2012). Estudios normativos españoles en población adulta joven (Proyecto NEURONORMA jóvenes): Normas para el Boston Naming Test y el Token Test. *Neurología*, 27(7), 394-399.
- Arango-Lasprilla, J. C., Olabarrieta-Landa, L., Rivera, D., Olivera Plaza, S. L., De los Reyes Aragón, C. J., . . . & Quijano, M. C. (2015). *Situación actual de la neuropsicología en Colombia*. En J.C. Arango-Lasprilla & D. Rivera (Eds.), *Neuropsicología en Colombia: Datos Normativos, Estado Actual y Retos a Futuro* (pp 21-46). Manizales, Colombia: Editorial Universidad Autónoma de Manizales.
- Arango-Lasprilla, J. C., Stevens, L., Morlett-Paredes, A., Ardila, A., & Rivera, D. (2016). The profession of Neuropsychology in Latin America. *Applied Neuropsychology: Adult*, 24(4), 318-330. doi: 10.1080/23279095.2016.1185423
- Ardila, A., & Rosselli, M. (1994). Development of language, memory and visuospatial abilities in 5- to-12-year-old children using a neuropsychological battery. *Developmental Neuropsychology*, 10(2), 97-120.

- Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., . . . & Douglas, L. (2001). Fourteen-year follow-up of speech/language-impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(1), 75-82.
- Bishop, D., & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*(2), 156-173.
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2000). *Test de inteligencia no verbal TONI-2*. Madrid: TEA ediciones.
- Chamorean, C. (2014). Diversidad lingüística en México. *Amerindia - Langues du Mexique, 37*(1), 3-20.
- Conti-Ramsden, G., & Botting, N. (2008). Emotional health in adolescents with and without a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 49*(5), 516-525.
- Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics, 19*(1), 15-18. doi: 10.2307/1268249
- De Renzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token test. *Cortex, 14*(1), 41-9.
- De Renzi, E., & Vignolo, L. A. (1962). The token test: A sensitive test to detect receptive disturbances in aphasics. *Brain, 85*(4), 665-678. doi:10.1093/brain/85.4.665
- Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn D. N. & Kurs-Lasky, M. (1999) Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research, 42*(6), 1432-1443.
- Fernandez, A. L., Ferreres, A., Morlett-Paredes, A., Rivera, D., & Arango-Lasprilla, J. C. (2016). Past, present, and future of neuropsychology in Argentina. *The Clinical Neuropsychologist, 30*(8), 1154-1178.
- Fonseca-Aguilar, P., Olabarrieta-Landa, L., Rivera, D., Aguayo Arellis, A., Ortiz Jiménez, X. A., Rábago Barajas, B., . . . & Arango-Lasprilla, J. C. (2015). Situación actual de la práctica profesional de la neuropsicología en México". *Psicología desde El Caribe, 32*(3), 343-364. doi: 10.14482/psdc32.3.7896
- Gallardo, G., Guárdia, J., Villaseñor, T., & McNeil, M. R. (2011). Psychometric data for the Revised Token Test in normally developing Mexican children ages 4-12 years. *Archives of Clinical Neuropsychology, 26*(3), 225-234.
- Garrido Medina, J. C. (2007). Multilingüismo y lengua externa e interna en la política lingüística en España. *Anuario de estudios filológicos, 30*, 131-149.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368-1378.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(6), 55-88.
- Hughes, S. (2014). Bullying: What speech-language pathologists should know. *Language, speech, and hearing services in schools, 45*(1), 3-12. doi:10.1044/2013.LSHSS-13-0013
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*(1), 53-69.
- Kovacs, M. (1992). *Children Depression Inventory CDI (Manual)*. Toronto: Multihealth systems.
- Kumar, S., Kumar, P. & Kumari, P. (2013). Standardization of the Revised Token Test in Bangla. *Indian Journal of Applied Linguistics, 39*(1), 140-163.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models (5th ed.)*. New York: McGraw Hill.
- Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35*(2), 165-188.
- Malloy-Diniz, L. F., Bentes, R. C., Figueiredo, P. M., Brandao-Bretas, D., da Costa-Abrantes, S., Parizzi, A. M., Borges-Leite, W. & Salgado, J. V. (2007) Normalización de una batería de tests para evaluar las habilidades de comprensión del lenguaje, fluidez verbal y denominación en niños brasileños de 7 a 10 años: Resultados preliminares. *Revista Neurología, 44*(5): 275-280.
- McLeod, S., & Harrison, L. J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4-to 5-year-old children. *Journal of Speech, Language, and Hearing Research, 52*(5), 1213-1229.
- McNeil, M. R., Pratt, S. R., Szuminsky, N., Sung, J. E., Fossett, T. R., Fassbinder, W., & Lim, K. Y. (2015). Reliability and validity of the computerized Revised Token Test: Comparison of reading and listening versions in persons with and without aphasia. *Journal of Speech, Language, and Hearing Research, 58*(2), 311-324.
- McNeil, M. R., & Prescott, T. E. (1978). *Revised token test*. Austin, TX: PRO-ED, Inc.
- Olabarrieta Landa, L., Caracuel, A., Pérez-García, M., Panyavin, I., Morlett, A., & Arango-Lasprilla, J. C. (2016). The Profession of neuropsychology in Spain: Results of a national survey. *The Clinical Neuropsychologist, 30*(8), 1335-1355.
- Pancsofar, N. & Vernon-Feagans, L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology, 27*(6), 571-587.
- Pancsofar, N., Vernon-Feagans, L., & Family Life Project Investigators. (2010). Fathers' early contributions to children's language development in families from low-income rural communities. *Early Childhood Research Quarterly, 25*(4), 450-463.
- Paquier, P. F., van Mourik, M., van Dongen, H. R., Catsman-Berrevvoets, C., Creten, W. L., & van Borsel, J. (2007). Normative data of 300 Dutch-speaking children on the Token Test. *Aphasiology, 23*(4), 427-437.
- Peña-Casanova, J., Quiñones-Úbeda, S., Gramunt-Fombuena, N., Aguilar, M., Casas, L., Molinuevo, J. L., . . . & Martínez-Parra, C. (2009). Spanish Multicenter Normative Studies (NEURONORMA Project): Norms for Boston naming test and token test. *Archives of Clinical Neuropsychology, 24*(4), 343-354.
- Rich, J. B. (1993). *Pictorial and verbal implicit recognition memory in aging and Alzheimer's disease: A transfer-appropriate processing account*. Ph.D. Dissertation, University of Victoria, Victoria, B.C, Canada.
- Rivera, D., & Arango-Lasprilla, J. C. (2017). Methodology for the development of normative data for Spanish Speaking pediatric population. *NeuroRehabilitation, 41*(3), 581-592.
- Schuele, C. M. (2004). The impact of developmental speech and language impairments on the acquisition of literacy skills. *Mental Retardation and Developmental Disabilities Research Reviews, 10*(3), 176-183.
- Snowling, M. J., Bishop, D. V., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of

- children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 47(8), 759-765.
- Spellacy, F. J., & Spreen, O. (1969). A short form of the Token Test. *Cortex*, 5(4), 390-397.
- Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. American Chemical Society.
- Sundheim, S., & Voeller, K. (2004). Psychiatric implications of language disorders and learning disabilities: Risks and management. *Journal of Child Neurology*, 19(10), 814-826.
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, & Hearing Research*, 40(6), 1245-1260.
- Turkyilmaz, M. D., & Belgin, E. (2012). Reliability, validity, and adaptation of computerized revised token test in normal subjects. *Journal Of International Advanced Otolaryngology*, 8(1), 103-112.
- Van der Elst, W., Hurks, P., Wassenberg, R., Meijs, C., & Jolles, J. (2011). Animal Verbal Fluency and Design Fluency in school-aged children: Effects of age, sex, and mean level of parental education, and regression-based normative data. *Journal of Clinical & Experimental Neuropsychology*, 33(9), 1005-1015. doi: 10.1080/13803395.2011.589509
- Wassenberg, R., Hurks, P. M., Hendriksen, J. M., Feron, F. M., Meijs, C. C., Vles, J. H., & Jolles, J. (2008). Age-related improvement in complex language comprehension: Results of a cross-sectional study with 361 children aged 5 to 15. *Journal of Clinical & Experimental Neuropsychology*, 30(4), 435-448. doi:10.1080/13803390701523091
- Willinger, U., Brunner, E., Diendorfer-Radner, G., Sams, J., Sirsch, U., & Eisenwort, B. (2003). Behaviour in children with language development disorders. *The Canadian Journal of Psychiatry*, 48(9), 607-614.

Appendix 1. Spanish Shortened Version of the Token Test

Token Test – Versión reducida

Cada respuesta correcta es un punto (1). En caso de repetir la instrucción y obtener una respuesta correcta, se otorgan 0.5 puntos durante las primeras 5 partes, y en caso de obtener una respuesta incorrecta 0 puntos. En la parte 6 no se permite la repetición de la instrucción. En esta sección se otorga un punto (1) por respuesta correcta y 0 puntos por respuesta incorrecta. Para obtener el total, sume las puntuaciones correctas.

Parte 1. Todas las fichas.	1. Toque un círculo	
	2. Toque un cuadrado	
	3. Toque una pieza amarilla	
	4. Toque una roja	
	5. Toque una negra	
	6. Toque una verde	
	7. Toque una blanca	
Total Parte 1		

Parte 2. Retire las fichas pequeñas.	8. Toque un cuadrado amarillo	
	9. Toque un círculo negro	
	10. Toque un círculo verde	
	11. Toque un cuadrado blanco	
Total Parte 2		

Parte 3. Coloque nuevamente las fichas pequeñas.	12. Toque un círculo blanco pequeño	
	13. Toque un cuadrado amarillo grande	
	14. Toque un cuadrado verde grande	
	15. Toque un círculo negro pequeño	
Total Parte 3		

Parte 4. Retire las fichas pequeñas.	16. Toque el círculo rojo y el cuadrado verde	
	17. Toque el cuadrado amarillo y el cuadrado negro	
	18. Toque el cuadrado blanco y el círculo verde	
	19. Toque el círculo blanco y el círculo rojo	
Total Parte 4		

Parte 5. Coloque nuevamente las fichas pequeñas.	20. Toque el círculo blanco grande y el cuadrado verde pequeño	
	21. Toque el círculo negro pequeño y el cuadrado amarillo grande	
	22. Toque el cuadrado verde grande y el cuadrado rojo grande	
	23. Toque el cuadrado blanco grande y el círculo verde pequeño	
Total Parte 5		

Parte 6. Retire las fichas pequeñas.	24. Ponga el círculo rojo encima del cuadrado verde	
	25. Toque el círculo negro con el cuadrado rojo	
	26. Toque el círculo negro y el cuadrado rojo	
	27. Toque el círculo negro o el cuadrado rojo	
	28. Ponga el cuadrado verde lejos del cuadrado amarillo	
	29. Si existe un círculo azul, toque el cuadrado rojo	
	30. Ponga el cuadrado verde junto al círculo rojo	
	31. Toque los cuadrados lentamente y los círculos deprisa	
	32. Ponga el círculo rojo entre el cuadrado amarillo y el cuadrado verde	
	33. Toque todos los círculos, menos el verde	
	34. Toque el círculo rojo. ¡No! El cuadrado blanco	
	35. En vez de tocar el cuadrado blanco, toque el círculo amarillo	
	36. Además del círculo amarillo, toque el círculo negro	
	Total Parte 6	

Total final: ____