

Original articles

Validity evidence of the Test of Narrative Language (TNL) adapted to Brazilian Portuguese

Gladis dos Santos¹<https://orcid.org/0000-0002-0323-2144>

¹ Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil.

Conflict of interests: Nonexistent



ABSTRACT

Purpose: to furnish psychometric evidence of the Test of Narrative Language version, examining a) the equivalence and analysis of internal consistency; b) performance differentiation per age; c) convergent validity; and d) dimensionality analysis.

Methods: the adapted test version was applied to 68 children with typical development, aged 5 to 6 years and 11 months, from private and public schools. The data were analyzed for the reliability and validity of the instrument.

Results: in the three narrative comprehension items, the highest scores were obtained when the story was told with five pictures in a logical sequence, followed by the one with a single picture, and lastly, the one with no pictures. In the three narrative production tasks, the best performance was in the story told along with a single picture, followed by the five pictures in a logical sequence, and lastly, the one with no pictures. There was no significant difference in the scores of all private-school children and in those of the 6-year-old public- and private-school children. The single-factor structure better explained the instrument dimensionality.

Conclusion: the reliability and validity evidence demonstrated its indication, and the potential of the results must be taken into account for future standardized construction to assess oral narrative.

Keywords: Language Development; Child Language; Language Tests; Narration; Speech, Language and Hearing Sciences

Received on: August 29, 2021
Accepted on: December 23, 2021

Corresponding address:

Gladis dos Santos
Coordenação do Curso de Fonoaudiologia
– Sala 9E11
Faculdade de Medicina / UFRJ
Rua Professor Rodolpho Paulo Rocco, 255
Cidade Universitária da Universidade
Federal do Rio de Janeiro
CEP: 21941-617 - Rio de Janeiro,
Rio de Janeiro, Brasil
E-mail: gladisdosantos@medicina.ufrj.br

INTRODUCTION

Social relations are based on communication. When children discover the symbolic function of language, they begin developing narratives, broadening their communicative competence. The narratives evolve and become increasingly complex until the child becomes a true narrator, which takes place around the age of 4 years old¹. To this end, the oral language in use in the child's social interactive environments must encompass a variety of styles. The diversity of communicative practices is a crucial factor to develop narratives² – both reports and stories –, which integrate the linguistic subsystems and cognitive skills.

The oral narrative has a classic definition: “A method to recall past experiences, corresponding a verbal sequence of clauses to the sequence of events that actually happened”³. The content of the narrative, through related events, is organized with global interconnections, macrostructures, related to coherence^{4,5}. On the other hand, the local connections, microstructures, related to cohesion, refer to syntax-related linguistic resources^{4,5}.

Difficulties in narrative production or comprehension may cause social and academic problems. A child's oral narrative skill will ground their interpretation of the first readings and written compositions⁶. The relationships between oral narrative and written stories are already perceived in the first years at schools^{7,8}. The child depends on successfully learning to read and write in order to progress academically. Therefore, the development of oral narrative production and comprehension skills must be followed up through various tasks and contexts that are effective enough to better guide pedagogical interventions^{8,9} as soon as preschool – which is when they begin to learn to read and write.

There is a known scarcity of standardized and up-to-date instruments to assess the macro- and microstructural components, comprehension, and production of children's oral narratives in Brazilian Portuguese¹⁰. The absence of such instruments makes it more difficult to identify impaired oral narrative skills and consequently the intervention that can be integrated into the children's school routine, or, in more severe cases, into the clinical work of a speech-language-hearing therapist. One way to make such assessment, which is very well accepted by children for being fun and pleasant, is story-telling and retelling – as proposed in the Test of Narrative Language (TNL)¹¹, developed to

assess the acquisition of oral narrative skills in children 5 years to 11 years and 11 months old.

The applicability of the TNL has been observed in the literature for various purposes. Its importance stands out in the identification and analysis of oral narrative difficulties in children diagnosed with language development disorder¹² and bilingual children^{13,14}. The TNL has been employed to document the oral narrative evolution in children who participate in intervention programs, aiming at their development¹⁵. The test has also been approached as a measure of oral language to verify the effectiveness of using another instrument to screen the risk of language disorders¹⁶. Moreover, the TNL, in the version adapted to Brazilian Portuguese, has been used to investigate the performance of children with typical development¹⁷.

The TNL has been recently translated and transculturally adapted to Brazilian Portuguese, under individual and nontransferable authorization granted in 2015, formalized in a legal contract between the researcher and the PRO-ED, Inc. publisher¹⁸. Essential steps were taken to adapt the test¹⁸, which made it possible to appraise the item, semantic, and concept equivalence and carry out a pilot study to refine the instrument and prepare the evaluators¹⁹. Besides the test itself, the application reference material was maintained, and the manual was translated, providing the conditions for the test to be well applied by the evaluators, and its results, corrected and analyzed according to the original proposal. Based on the evidence of construct validity and reliability, the results of the Brazilian Portuguese version favored the application of the TNL to children who are learning to read and write.

The objective of this study is to furnish evidence on the psychometric qualities of reliability and validity of the TNL version for clinical use, examining a) the equivalence and analysis of the internal consistency of the instrument; b) performance differentiation per age; c) convergent validity; and d) analysis of instrument dimensionality.

METHODS

Ethical aspects

The study was conducted in compliance with the ethical precepts, as stated in the Resolution of the *Conselho Nacional de Saúde* (National Health Council) no. 466, of December 12, 2012, of the Ministry of Health. The research protocol was evaluated and approved

by the Ethics Committee of the Institute of Neurology Deolindo Couto, Brazil, under number 1,972,625.

The researcher held in-person meetings at the schools to present the proposal to the children's parents/guardians. Those who attended the meetings and authorized their children's participation signed an informed consent form developed for the study. They also answered a socioeconomic classification questionnaire, based on the Brazilian Economic Classification Criteria, which was developed by the *Associação Brasileira de Empresas de Pesquisa* (Brazilian Association of Research Companies – ABEP, 2015). It comprises 12 items on comfort, householder's educational attainment, and access to tap water and paved streets.

Study population

The children were recruited between August and December 2017. They were 5 years to 6 years and 11 months old and studied in four private and one public school in the municipality of Rio de Janeiro.

The publisher had permitted to apply the test only to a small number of participants (around 60). This restricted the application of the TNL to two age groups to obtain validity evidence. The age range was established based on the important opportunity of taking both educational and clinical measures and providing better oral language development in children with positive reflexes on their process of learning to read and write.

Inclusion and exclusion criteria

The teacher's observation of the children's academic performance compatible with their grade in school was considered, as well as the absence of any complaints or difficulties related to sensory, mental, neurological, neuropsychiatric, linguistic, and or cognitive changes. The children's cognitive-linguistic development was also assessed with phonological processing and nonverbal intelligence assessments and the Language Development Assessment (LDA)²⁰. Each child was individually submitted to these assessments in two 45-minute sessions, at a time and place convenient to the school.

The LDA, the assessment of phonological processing (lexical access, phonological awareness, and working memory), and the TNL were applied by the speech-language-hearing researcher. The nonverbal intelligence test-R2 was applied by psychologists

because its use is restricted to such professionals. They applied the test during the second session; the researcher was present because she was already acquainted with the children and the school setting. The children's results were as expected for their age; therefore, no participant was excluded.

The Test of Narrative Language (TNL)

The TNL assesses the knowledge of story structure, considering the elements of which narratives are made up. The instrument significantly correlates age and performance in both narrative comprehension and production, as well as other traditional oral language measures¹¹.

The transcultural adaptation of the TNL, made by other authors, found significant differences in the performance of public-school Brazilian Portuguese-speaking children per age²¹. No other validity evidence for the TNL was reported. The version made for the present study, aimed at investigating validity evidence¹⁸, was used to find the performance of public- and private-school children in the TNL.

The test comprises three oral narrative comprehension tasks and three oral narrative production tasks, to be done by the child. The three oral narrative comprehension tasks consist of three stories told to the child, presented as follows: a) the "McDonald's" story, with no pictures, b) "The Broken Boat" story, with five pictures in a logical sequence, and c) "The Dragon" story, with a single picture. The three oral narrative production tasks are done by the child as follows: a) retelling the "McDonald's" story, with no pictures, b) creating a story about a boy who was "Late to school", with five pictures in a logical sequence, and c) creating a story about "Extraterrestrials", with a single picture. The tasks are intercalated, and the child does one comprehension task and then one production task, as explained below.

Task one, named "McDonald's", does not have pictures. Initially, the researcher asked the child whether they had already eaten at McDonald's and what they liked to order there. If the child's answer was "no", the evaluator would explain that McDonald's is a fast-food restaurant and then waited for the child to name usual snacks, not giving them any score. Then, the child was told they would hear a story and after that, they would answer some questions and retell the story. Then, the child was asked 11 questions, some of which had more than one correct answer, totaling 15 points.

In task two, without pictures, the child was asked to retell the “McDonald’s” story, without any intervention. The score was given on the text, where keywords were written in bold, followed by a blank space to write down 1 point if the child had included it in the narrative. The story had 26 keywords in bold, totaling 26 points in task two.

Task three involved “The Broken Boat” story, read to the child while showing the five pictures in a logical sequence. The logical sequence remained exposed all the time the child was listening to the story and answering nine questions on its items and the characters’ mental state. Answering the characters’ names and actions and the story setting, for instance, indicated a coherent lexical-semantic knowledge of the story. The task totaled 11 points.

In task four, which involved the “Late to school” story, the researcher asked the child to tell a story based on five sequential pictures, showing them from left to right, without giving them any cues. Points were given to 18 essential items, expected for this narrative, besides zero to two points to macro- and micro-structure elements divided into events, grammar, and story. Events referred to the relationships temporally and causally established between them. The grammar score referred to the adequate use of verb tenses and sentence structures. Finally, the coherence, organization, and creativity of the story were observed. Task four totaled 30 points.

In task five, “The Dragon” story was told, using a single picture. The child was asked 10 questions about that picture, on items of the story and the characters’ mental state. Questions on names, places, and facts required objective answers about items present in the main events. One question elicited the character’s expressed mental state. One point was given to each correct answer, totaling 14 points.

Finally, in task six, a picture with extraterrestrials was shown. The child was asked to tell a complete story. The only intervention allowed was to say: “Tell me how this story begins”. The total score in this task was 34 points, distributed between the following elements of the resulting narrative, always according to the original protocol: a) Setting: where and when the story took place; b) Characters: identification of humans and extraterrestrials and the dialogue between them; c) Elements of the story: the problem, actions, and events to solve the problem, temporal and causal relationships, consequences, and ending; d) Vocabulary and grammar: adequate description of objects, personal

references, verb tenses, and sentence structures; e) Story: coherence, organization, and creativity.

The oral narrative comprehension and production tasks were separately scored. The raw score, age equivalent, percentage, and standard grade were calculated. The summed standard grades resulted in the compound score. This one in turn defined the narrative language skill index (NLSI). This index describes the narrative language skill as “very superior” (for scores equal to or above 130) and “very poor” (for scores equal to or below 70).

Data analysis procedures

The evidence of the psychometric qualities of the TNL version applied to Brazilian children was examined according to the procedures described below.

The reliability of the test was assessed by equivalence – i.e., the agreement between two evaluators – and internal consistency. Hence, for equivalence, it was necessary to count with the participation of two independent evaluators, trained by the researcher to give scores to the children’s answers, following the original instructions in the TNL manual. The evaluators did not meet the children or applied the test. The agreement between the two evaluators regarding the score of the children’s answers in each TNL subtest was calculated with the following coefficients: a) Pearson correlation for total TNL, comprehensive language, and expressive language; and b) Cohen’s Kappa for each test item. The analysis of the internal consistency of the instrument was calculated with Cronbach’s coefficient alpha.

Validity evidence was presented based on the children’s performance, considering their age and type of school, both in the separate scores for narrative comprehension and production and total TNL score. The analysis of variance (ANOVA) was conducted on each level of analysis, with the children’s age and type of school as factors.

The convergent validity for the 5-to-6-year age range was examined with the Pearson correlation analysis of the children’s performance in the TNL with the LDA.

The examination of the structure of the TNL – i.e., the analysis of the instrument dimensionality – was conducted with exploratory (EFA) and confirmatory factor analysis (CFA). Considering the two-dimensionality of the TNL (oral narrative production and comprehension), two CFA were conducted: one without specifying the number of factors and the other determining the distribution into two factors. The main

component analysis method was employed in all analyses. The measure of sampling adequacy was examined with the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity.

The maximum likelihood with robust (Huber-White) standard errors (MLR) was used for the CFA – whose purpose is to assess the goodness-of-fit indices of the exploratory models. The absolute fit indices calculated were the chi-square (χ^2), chi-square/degrees of freedom ratio (χ^2/df), and standardized root mean square residual (SRMR). The parsimonious goodness-of-fit index used was the root mean square error of approximation (RMSEA). Lastly, the comparative fit index (CFI) and Tucker-Lewis index (TLI) were used.

RESULTS

A total of 34 five-year-old children (15 males) participated in the study, all attending kindergarten, 16 of them in private schools. Also, 34 six-year-old children (23 males) participated, eight of whom still attended kindergarten, while the other 26 were in first grade; 17 went to private schools.

The socioeconomic profile of the study population, according to the Brazilian Economic Classification Criteria²², portray the unequal Brazilian society, in which most private-school children (more than 77%) belonged to classes A, B1, and B2, whereas those from public schools (more than 71.5%) belonged to classes C1, C2, and D-E. Moreover, while almost half (43.7%) of the householders in the families of public-school children only had basic education, all the parents/guardians of the 5-year-old children and almost all (85.7%) parents/guardians of the 6-year-old children from private schools had a bachelor's degree.

The other results are presented in four sessions: (a) evidence of reliability of the TNL; (b) evidence of validity with data on the children's performance in both the comprehensive and expressive language tasks; (c)

evidence of convergent validity for the TNL; and (d) analysis of instrument dimensionality.

a) Evidence of Reliability

Initially, the agreement coefficient between the two independent evaluators was calculated, employing the Pearson correlation for the comprehensive ($r=1.00$) and expressive narrative scores ($r=0.86$, $p<0.01$) and the total score ($r=0.95$, $p<0.01$) of the children's answers in the TNL. Values equal to or higher than 0.60 are strong, and those equal to or higher than 0.90 are very strong.

The agreement coefficient between the two evaluators was calculated regarding each of the TNL items with Cohen's kappa. The comprehensive language items obtained 1.00, classified as excellent, due to the objective answer required from the children. The interval of the kappa coefficients for expressive language revealed that the coefficient in 29% of the items was below the acceptable.

The internal consistency of the test was calculated with the Cronbach alpha coefficient. The result for narrative language skill (the whole test) was 0.91; for comprehensive language, 0.83; and for expressive language, 0.87.

b) Validity Evidence: Children's performance in the TNL tasks

Narrative Comprehension

Concerning the validity evidence, the children's performance in the narrative comprehension tasks is presented in Table 1. The ANOVA results revealed a significant difference for age ($F [1.64]=8.73$, $p<0.01$), type of school ($F [1.64]=31.64$, $p<0.01$), and task ($F [2.128]=41.44$, $p<0.01$). Post hoc comparisons (Bonferroni) showed significant differences between the tasks. The interactions were not significant.

The children, regardless of age and type of school, performed better, in decreasing order, in listening comprehension of the story with a logical sequence (Broken Boat), then the single picture (Dragon), and lastly without pictures (McDonald's). This performance was measured with the questionnaire with literal

questions while seeing the pictures, except for task one, in which the McDonald's story was told without any pictures. The only inferential question identified in the questionnaire was in task one: "What do you think they should do?" after they had ordered their snacks and the mother realized she had left her wallet at home.

Table 1. The children's performance in the narrative comprehension tasks (N= 68)

| TASKS | AGE | SCHOOL | | | | Total | |
|-----------------------------------|---------|----------------------|------|---------------------|------|-------|------|
| | | Private ^a | | Public ^b | | M | SD |
| | | M | SD | M | SD | | |
| No Picture (McDonald's) | 5 years | 0.65 | 0.13 | 0.47 | 0.10 | 0.55 | 0.14 |
| | 6 years | 0.71* | 0.16 | 0.59* | 0.21 | 0.65 | 0.19 |
| | Total | 0.68 | 0.15 | 0.53 | 0.17 | 0.60 | 0.18 |
| Logical Sequence (Broken Boat) | 5 years | 0.88 | 0.09 | 0.64 | 0.15 | 0.75 | 0.17 |
| | 6 years | 0.87* | 0.10 | 0.76* | 0.15 | 0.81 | 0.14 |
| | Total | 0.87 | 0.10 | 0.70 | 0.16 | 0.78 | 0.16 |
| Single Picture (Dragon) | 5 years | 0.71 | 0.13 | 0.55 | 0.18 | 0.63 | 0.17 |
| | 6 years | 0.78* | 0.11 | 0.67* | 0.15 | 0.73 | 0.14 |
| | Total | 0.75 | 0.13 | 0.61 | 0.17 | 0.68 | 0.17 |

a) Private-school children: Total=33; 5 years old=16; 6 years old=17.

b) Public-school children: Total=35; 5 years old=18; 6 years old=17.

Captions: M = mean; SD = standard deviation.

ANOVA, adjusted with Bonferroni correction. * $p < 0.01$

Most (79.4%) of the 5- and 6-year-old children in this study referred to two or more appropriate actions on the part of the characters. Nevertheless, the children understand the stories told with visual support better. Comparatively, pictures shown in a logical sequence were even more helpful than a single picture. Illustrating the story step-by-step along with the narrative to which the child is listening was effectively used for the better functioning of the whole mental architecture involved in the oral narrative comprehension task. Moreover, regarding the pictureless McDonald's story, all children said they knew McDonald's – except for four children, two from private and two from public schools, who said they did not know any fast-food restaurant.

Narrative Production

The scores of the narrative production tasks are presented in Table 2. The ANOVA results revealed a significant difference for age ($F [1.64]=18.22, p < 0.01$), type of school ($F [1.64]=14.17, p < 0.01$), and task ($F [2.128]=38.65, p < 0.01$). The post hoc test (Bonferroni) revealed a significant difference between all tasks. There was a significant interaction only between age and type of school ($F [1.64] = 4.27, p = 0.04$). The post hoc test (Student's t-test with Bonferroni correction with $p\text{-value} < 0.01$) revealed no significant difference between the scores of the 5- and 6-year-old private-school children ($t[32] = 1.65, p = 0.06$) or between the scores of the 6-year-old public- and private-school children ($t[32] = 1.26, p = 0.22$).

Table 2. The children's performance in the narrative production tasks (N= 68)

| TASKS | AGE | SCHOOL | | | | Total | |
|---------------------------------------|---------|----------------------|------|---------------------|------|-------|------|
| | | Private ^a | | Public ^b | | M | SD |
| | | M | SD | M | SD | | |
| No Picture (McDonald's) | 5 years | 0.39* | 0.10 | 0.23* | 0.10 | 0.31 | 0.13 |
| | 6 years | 0.37 | 0.13 | 0.38 | 0.09 | 0.38 | 0.11 |
| | Total | 0.38 | 0.11 | 0.30 | 0.12 | 0.34 | 0.12 |
| Logical Sequence (Late to school) | 5 years | 0.45* | 0.10 | 0.37* | 0.10 | 0.41 | 0.10 |
| | 6 years | 0.50 | 0.12 | 0.47 | 0.07 | 0.48 | 0.10 |
| | Total | 0.48 | 0.11 | 0.42 | 0.10 | 0.45 | 0.11 |
| Single Picture (Extraterrestrials) | 5 years | 0.50* | 0.16 | 0.37* | 0.11 | 0.43 | 0.15 |
| | 6 years | 0.61 | 0.15 | 0.53 | 0.16 | 0.57 | 0.16 |
| | Total | 0.56 | 0.16 | 0.45 | 0.15 | 0.50 | 0.17 |

a) Private-school children: Total=33; 5 years old=16; 6 years old=17.

b) Public-school children: Total=35; 5 years old=18; 6 years old=17.

Captions: M = mean; SD = standard deviation.

ANOVA, adjusted with Bonferroni correction. * $p < 0.01$

Narrative Comprehension and Production

The scores of the narrative comprehension and production indices are presented in Table 3. It is important to highlight that the narrative language skill

index (NLSI) of all children in this study was either average or superior. This standard is compatible with same-age North American children.

Table 3. Narrative comprehension and production indices

| TASKS | AGE | SCHOOL | | | | Total | |
|-------------------------|---------|----------------------|------|---------------------|------|-------|------|
| | | Private ^a | | Public ^b | | M | SD |
| | | M | SD | M | SD | | |
| Narrative Comprehension | 5 years | 0.73 | 0.07 | 0.55 | 0.12 | 0.63 | 0.13 |
| | 6 years | 0.78 | 0.10 | 0.66 | 0.13 | 0.72 | 0.13 |
| | Total | 0.76* | 0.09 | 0.60* | 0.14 | 0.68* | 0.14 |
| Narrative Production | 5 years | 0.45 | 0.10 | 0.33 | 0.06 | 0.39 | 0.10 |
| | 6 years | 0.51 | 0.10 | 0.46 | 0.08 | 0.49 | 0.09 |
| | Total | 0.48* | 0.10 | 0.40* | 0.10 | 0.44* | 0.11 |

a) Private-school children: Total=33; 5 years old=16; 6 years old=17.

b) Public-school children: Total=35; 5 years old=18; 6 years old=17.

Captions: M = mean; SD = standard deviation.

ANOVA, adjusted with Bonferroni correction. * $p < 0.01$

Post hoc test (Student's t-test, adjusted with Bonferroni correction. * $p < 0.01$)

The ANOVA results revealed a significant difference for age ($F [1.64] = 17.04, p < 0.01$), type of school ($F [1.64] = 29.64, p < 0.01$), and narrative dimension ($F [1.64] = 353.79, p < 0.01$). There was a significant interaction only between the type of school and narrative dimension ($F [1.64] = 7.48, p < 0.01$). The post hoc test (Student's t-test with Bonferroni correction with p -value < 0.01) revealed a significant difference

between expressive and comprehensive narrative, for both public ($t [33] = 10.19, p < 0.01$) and private schools ($t [33] = 18.24, p < 0.01$), with higher scores for comprehensive narrative. There was likewise a significant difference between public and private schools when the comprehensive ($t [66] = 5.34, p < 0.01$) and expressive narrative dimensions ($t [66] = 3.24, p < 0.01$) were considered.

Oral Narrative (Total Score)

The children's performance in the test total score showed a significant correlation between the TNL and age ($r=0.41$, $p<0.01$). The ANOVA was used to compare the children's performance in oral narrative, with the mean (M) and standard deviation (SD) and using age and type of school as factors. From the private schools, the 16 five-year-old children obtained $M=0.54$ and $SD=0.07$, while the 17 six-year-old children obtained $M=0.59$ and $SD=0.09$. The total of the 33 private-school children was $M=0.57$ and $SD=0.09$. From the public school, the scores of the 35 children were $M=0.46$ and $SD=0.10$; for the 18 five-year-old children, they were $M=0.40$ and $SD=0.07$, and for the 17 six-year-old children, $M=0.53$ and $SD=0.08$. For all 68 children, the scores were $M=0.52$ and $SD=0.11$; those 5 years old obtained $M=0.47$ and $SD=0.10$, while those 6 years old obtained $M=0.56$ and $SD=0.09$.

In short, the ANOVA results revealed a significant difference for age ($F [1.64]= 19.428$, $p<0.01$) and type of school ($F [1.64]=31.87$, $p<0.01$). The 6-year-old children performed better than the 5-year-old ones. The children who studied in private schools performed better than their public-school peers. There was no significant interaction between age and type of school ($F [1.64]=3.79$, $p=0.06$).

c) Evidence of convergent validity for the Test of Narrative Language

The relationship between the TNL and LDA was examined with the Pearson correlation test. A significant

correlation was found between the total scores of the TNL and LDA ($r=0.39$, $p=0.01$) and between the expressive language scores of the TNL and LDA ($r=0.34$, $p<0.01$). However, no statistically significant correlation was found between the receptive language scores of the TNL and LDA ($r=0.10$, $p=0.41$).

d) Analysis of the TNL Dimensionality

Exploratory Factor Analysis (EFA)

The extraction method with the main component analysis for the TNL was used for the first EFA, based on the Eigenvalue (3.202), with 53.36% explained variance. This analysis, conducted without factor specification, presented the one-factor solution, which included all the TNL subtests ($KMO = 0.81$; Bartlett's test of sphericity $\chi [15] = 130.02$, $p<0.01$). The factor loadings for each subitem were as follows: a) comprehension tasks: McDonald's=0.84; Broken Boat=0.74; Dragon=0.73; b) production tasks: Extraterrestrials=0.74; McDonald's=0.72; Late to school=0.60).

The second EFA, conducted with the main component analysis method, oblique rotation (Promax) and two-factor fixed number, presented a solution described in Table 4 (fixed factor analysis), which includes all TNL subtests ($KMO = 0.83$; Bartlett's test of sphericity $\chi [15] = 130.02$, $p<0.01$). Component 1, with an eigenvalue of 3.202, has a percentage of explained variance of 53.36; component 2, with an eigenvalue of 0.904, has a percentage of explained variance of 15.06.

Table 4. Factorial analysis with fixed factors

| | TNL | |
|----------------------------|-----------------|--------|
| | Factor Loadings | |
| | Components | |
| Subtests | 1 | 2 |
| Eigenvalue | 3.202 | 0.904 |
| MACCOMP | 0.841 | 0.542* |
| BOATCOMP | 0.819 | 0.329 |
| DRAGONCOMP | 0.700 | 0.525* |
| ETPROD | 0.579* | 0.800 |
| MACPROD | 0.797* | 0.318* |
| LATEPROD | 0.334 | 0.906 |
| % Total Explained Variance | | 68.42 |

*Extraction method: Main component analysis, with oblique rotation (Promax); KMO=0.83; Bartlett's test of sphericity; $p < 0.01$.

Captions: TNL = Test of Narrative Language; MACCOMP = "McDonald's" comprehension task; BOATCOMP = "Broken Boat" comprehension task; DRAGONCOMP = "Dragon" comprehension task; ETPROD = "Extraterrestrials" production task; MACPROD = "McDonald's" production task; LATEPROD = "Late to school" production task.

Confirmatory Factor Analysis (CFA)

The goodness-of-fit indices for all exploratory models (single-factor and two-factor models) for the

TNL are shown in Table 5. Even though the two-factor model also had good fit indices, this did not occur with the RMSEA.

Table 5. Goodness-of-fit indices of the Test of Narrative Language

| Models | Goodness-of-fit Indices | | | | | |
|---------------|-------------------------|-------------|--------|-------|-------|-------|
| | χ^2 (df) | χ^2/df | RMSEA | SRMR | CFI | TLI |
| Single-factor | 11.926 (9) | 1.325 | 0.079 | 0.057 | 0.969 | 0.948 |
| Two-factor | 11.575 (8) | 1.446 | 0.089* | 0.055 | 0.965 | 0.934 |

Captions: χ^2 = chi-square; χ^2/df = chi-square/degrees of freedom ratio; RMSEA = Root mean square error of approximation; SRMR = Standardized root mean square residual; CFI = Comparative fit index: goodness-of-fit; TLI = Comparative fit index: Tucker-Lewis Index.

$\chi^2/df < 2$ or 3; *RMSEA ≤ 0.07 ; SRMR ≤ 0.08 ; CFI and TLI > 0.90 or 0.95.

DISCUSSION

The objective of this study was to obtain validity evidence for the Brazilian Portuguese version of the TNL for clinical use. The validity evidence with data on the children's performance are discussed below, considering the comprehension tasks alone, then the production tasks alone, and the interaction between the type of school and narrative dimension. Afterward, the evidence of convergent validity for the TNL was approached, and lastly, the instrument dimensionality was discussed.

Narrative Comprehension Tasks

The better performance of the 6-year-old children in the TNL comprehension tasks indicates the expected

evolution with age, corroborating the literature²³. The larger receptive vocabulary in this age may make it easier to understand words and the relationship between them to establish the meaning of the text they hear. Moreover, 6-year-old children answer questions more precisely than those 4 years and 6 months old – including the more complex questions on the solution of the problem in the story and its consequence²⁴. There seems to be a continuous evolution from 4 years old that does not depend on linguistic aspects alone, such as the relationship and integration between statements and world knowledge.

This study corroborates better result findings among private-school children in their oral narrative discourse tasks²³. This datum indicates that public-school children seem to have fewer opportunities to access

the diversity of cultural goods, which may restrict their construction of world knowledge. Oral language development problems have been related to socioeconomic influences in less favored populations, characterizing delayed or poor vocabulary and syntax^{25,26}. On the other hand, the quality of the oral narrative of both private- and public-school children has improved after tutoring by the evaluator²⁷. In other words, by means of tutoring, all children showed potential skills in the zone of proximal development²⁸.

The children's social class and their parents'/guardians' educational attainment should not be viewed as the only factors that influence the quality of the children's narrative comprehension. The time they spend in preschool and the home environment, including the means used to stimulate language and the quality of the parent-child interaction, should be considered as well^{29,30}.

Concerning the context of the comprehension tasks, the children's worst performance was in the McDonald's story, told without pictures. Good image resources are believed to integrate the text that is being heard; they are essential to 3-to-4-year-old children³¹. Seemingly, the absence of pictures made it more difficult for the 5- and 6-year-old children in this study to understand the McDonald's story.

Better performance in story comprehension was achieved with exposure to the logical sequence, in the Broken Boat story. The images worked as mnemonic support, as they explicitly represented the answer to certain questions.

Next, the best mean was obtained in the narrative comprehension with a single picture, which seems to have caused the same effect of the pictures in the logical sequence, though less intensely. The Dragon story posed two problems. The main one is presented in the scene in which the dragon spits fireballs towards Michele while she is taking gold from the chest. When asked what the problems in this story were, most (60.3%) of the 5- and 6-year-old children referred to only one of them, and 83% of these answered the problem shown explicitly in the picture. The influence of the picture on the children's answers was evident, although the task of registering, integrating, and processing all information in the story was more difficult in this task than in the one with a logical sequence.

Using visual images while they hear the story helps them understand oral narratives because information processing is facilitated³²⁻³⁴. The listening comprehension of stories is directly influenced by the

operational memory skills, knowledge of grammar, inference, and theory of mind, as shown in this study³⁵.

Reflecting on the abovementioned skills³⁵, it is noticed that presenting the visual images along with the auditory information of the storytelling helps them function better. Thus, the visual images prevented the overload of the operational memory, helped understand the personal (she), spatial (there), temporal (when she got there...), and causal references (... because she fell on the mud), and confirmed the inferences about certain solutions to the problem in the story and the theory of mind concerning the assumptions of the characters' mental states based on explicit facial expressions.

The TNL questionnaires had literal questions to assess the retrieval of essential information in the story. There must be a balance between the quantity and quality of the questions related to macrostructural elements. Furthermore, none of the questions investigated the comprehension of microstructural elements, such as the comprehension of anaphoric elements. Comprehension tasks verify the mental representations the listener developed, through which the child apprehends the story they are being told. The issue, then, is: What is the best way to investigate them? The questionnaire is valid, but retelling the story would be interesting as well, as it allows assessing how the child organizes the relevant ideas to give them meaning.

Narrative Production Tasks

Reflecting on the absence of significant difference in narrative production between 5- and 6-year-old private-school children, it may be thought that 5-year-old children narrate more, while 6-year-old children narrate better. The 5-year-old children are beginning to use more connectives³⁶, and their vocabulary is being enriched; however, they still need to better organize the elements they produce in the narrative³⁷. Comparatively, the 6-year-old children, having acquired the gains inherent to their cognitive-linguistic development, begin mastering basic macrostructure in their narrative productions³⁸.

The absence of significant difference between the scores obtained by the 6-year-old public- and private-school children stood out. Seemingly, the process of learning to read and write is a factor that marks such a difference, providing an evolution in the expressive language, especially for the 5-year-old public-school children, favoring their ability to produce better narratives by 6 years old.

The context of the tasks with visual images furnished more elaborate narratives than the ones without them, corroborating other studies^{31,32}. It is also interesting to highlight the best result in narrative production using a single picture, in contrast with the logical sequence of images, as that one provides smaller support to structure the story.

From 6 to 7 years old, children enrich their narrative, imagining elements or events that are not in the pictures³¹. The single picture may have given greater freedom to the children, as they resorted to their life experiences and mental representations instigated by the picture to narrate their story.

The simultaneously shown five pictures in a logical sequence not only are greater in number but also revealed the plot step-by-step. Hence, the children probably did not use cohesive elements, or used few of them, when they articulated each event, restarting their speech in every panel. It can be stated that different narrative production contexts influence how the story is produced.

Retelling the pictureless McDonald's story required the children's working memory skills to organize and linguistically express the retrieved more important information. As there were no pictures to help them retrieve the information, the children in this study obtained the worst performance. The TNL 2³⁹, published in 2017 by the same authors of TNL, includes a picture to perform this same task.

A recent study sheds light on this discussion, as it assessed the types of storytelling and retelling tasks, both with pictures, in 4-to-5-year-old children⁴⁰. Greater lexical diversity was observed in retelling the story, and greater sentence complexity, in telling the story. When applying the TNL, it was not possible to compare the same pictureless McDonald's story retelling task with another pictured story retelling task or pictureless story production.

Furthermore, in this TNL narrative production task, a single score index (lexical diversity) was considered. According to the authors of the TNL¹¹, the child must say the word or an acceptable synonym, not necessarily in the same sentence or order of the original story. In the other production tasks, with the logical sequence (Late to school) and a single picture (Extraterrestrials), other indices were also assessed, such as a) grammaticality (sentence complexity with cohesive temporal and causal relationships, verb conjugation, and presence of errors), b) lexical diversity (vocabulary and references), c) contextualization cues (setting and characters),

and d) other elements in the story (development, conclusion, meaning, and creativity). Thus, it was not possible to better compare the performance in the pictured and pictureless tasks because the children's narrative productions were not analyzed based on the same indices. As in a previously mentioned study⁴⁰, it is suggested that different types of tasks be assessed with the same elicitation (storytelling and retelling, both with pictures). Hence, it could not be concluded that the children in the present study performed worse in the McDonald's story retelling task, due to the type of task, the absence of pictures, or the type of score used.

It is important to mention that there are other tasks for comparison, such as the report, included in the TNL 2, published in 2017³⁹ and developed by the same authors of the TNL. The main difference between the TNL and TNL 2 is that this last one includes a report task and the picture for the McDonald's story retelling task. If the child's performance is poor in a given narrative assessment, based only on one type of task and from the same context, it should not be concluded that they have difficulties narrating. Creating or retelling a story is different from reporting a life experience. The report narrative reveals the representation developed by the person regarding their own experiences in the environment where they live.

The TNL proposed an in-depth analysis of the macro- and microstructure in the production tasks, except for figuring the characters' mental states. The protocol made objective scoring easier in the tasks, as it directed the evaluator's attention to certain grouped items in the macrostructural components in the story. Nevertheless, based on the reliability evidence for narrative production, the manual could provide further examples of acceptable answers and instruct inexperienced evaluators in further detail on how to score the test.

Narrative Comprehension and Production Tasks

Regarding the interaction between schools and narrative dimensions, it is believed that the scores obtained by the public- and private-school children are closer in production than in comprehension because the analysis of narrative production has similar micro- and macrostructural criteria. In all tasks, the analysis was based on keywords that should be present in the macrostructure of the story created by the child. Two tasks (Late to school and Extraterrestrials) also assessed the presence of microstructural elements distributed between the items named as events or

elements in the story, vocabulary, grammar, and the meaning of the story.

The analysis of narrative comprehension not only did not investigate the understanding of the microstructural elements but also varied in the quantity and quality of the macrostructural questions, leading to further variability in the children's answers.

Lastly, the validity evidence demonstrated in the children's performance in the total score obtained in the test indicates that the TNL distinguishes the narrative performance between the ages researched.

The evidence of convergent validity for the TNL suggests a weak correlation between its receptive language scores and those of the LDA, due to different and mutually unrelated comprehensive language aspects considered in each test. While the LDA has objective questions that require linguistic skills, particularly on the lexical-semantic level, the TNL assesses the comprehensive dimension of language applied to the stories created for the test itself.

Analysis of the TNL Dimensionality

The best way to explain the findings on the TNL dimensionality would be to consider the EFA results with a two-factor fixed number. Considering factor 1 as the comprehension, it was observed that the three tests proposed to assess comprehension (McDonald's, Broken Boat, and Dragon) had a greater loading in this factor. However, the McDonald's and Dragon tasks had moderate factor loadings in the factor named production.

There were questions in these tasks (McDonald's and Dragon) addressing the characters' attempts to deal with the situation before them – contrary to the Broken Boat task, whose questionnaire did not have such questions. This difference is probably related to the importance of this macrostructure element (attempts) to the development of the plot in the narrative production.

Considering factor 2 as the production, two of the tests proposed to assess production (Extraterrestrials and Late to school) had a greater loading in this factor. However, the McDonald's task, which assesses production, had a higher factor loading in the comprehension factor.

The McDonald's task assessed narrative production through story retelling. However, this production was not assessed regarding microstructure and organization of ideas; rather, the precision of the words used to retrieve the original information in the story was

assessed. Retelling, also known as story reproduction, has been considered a narrative comprehension task, in which the fidelity to the information in the original text is verified²⁴. In the TNL, retelling the McDonald's story verified and scored this very aspect. Therefore, the way this retelling is analyzed in the TNL corroborated what is being considered as an oral narrative comprehension task.

The Extraterrestrials production task had a high factor loading in the production factor. Nevertheless, it also had a moderate factor loading in comprehension. This last result may refer to the degree of comprehension of the picture, expressed in the narrative generated by the children. Before narrating, the children must integrate the various elements in the picture, verify whether they are familiar visual stimuli stored in their memory, observe the characters' facial expressions, access their world knowledge to construct the perceptive unification of the picture, and develop their narrative. Hence, the narrative production task depends on understanding the picture in all its complexity. Perhaps, if the correction system in the task had considered the obligatory keywords to the narrative, it might have mitigated the effect of the comprehension related to the presentation of a single picture.

The narrative comprehension assessment should explore further the mental representations the children make throughout the story and use more inferential items in the questionnaire. On the other hand, the narrative production assessment was expected to require, besides all the elements encompassed in the protocol, the presence of keywords related to a single picture, as it did in the logical sequence task.

The close relationship between comprehensive and expressive language poses some difficulties in distinguishing them in oral narrative tasks. Nevertheless, either oral narrative comprehension or production can be predominantly assessed, despite the interference of one dimension with the other. The specific functioning on each linguistic level in both dimensions can also be assessed. The oral narrative is a unique task, requiring cognitive-linguistic functioning in the whole dimension of oral language. Therefore, it is necessary to reflect on the assessment method (task and context), the precise selection of the indicators, and the analysis system that best express the dimension being assessed. Thus, flaws are seen in narrative production, while narrative comprehension remains intact.

In the confirmatory factor analysis, the absolute and comparative fit indices are expected to have $\chi^2/$

df values lower than 2 or 3; the SRMR, lower than or equal to 0.08 to indicate a good fit; and the RMSEA, lower than or equal to 0.07 to indicate a good fit. The CFI and TLI indices must be higher than 0.90 or 0.95 or close to these values. In this study, both the single- and two-factor models met the proposed values. Hence, by the principle of parsimony, the best model for the TNL would be the one with a single-factor structure. Thus, for the 5-to-6-year-old Brazilian children, the TNL would be an instrument to assess the overall narrative skill.

The world literature on children's oral narrative is vast from the children's performance standpoint, obtained with various tests and tasks. On the other hand, it is scarce when the object of study is the very test or task. In this sense, the present study has contributed to the analysis of the children's oral narrative assessment means, with the TNL as its object and, as its objective, to furnish evidence of the psychometric qualities of reliability and validity of the TNL version for clinical use. Nonetheless, the restricted number of participants in the study and age range, from 5 years to 6 years and 11 months old, were limitations that can be solved in further research on the topic.

The evaluator must verify, with preestablished criteria, in which form and context the child showed better and worse narrative comprehension and production. Thus, a narrative test must help investigate the actual level of development and the zone of proximal development²⁶. This investigation enables the analysis of the facilitating and complicating aspects in the narratives, which construct the differential diagnosis of language disorders. Hence, it directs the therapeutic or educational planning, as it helps understand specific needs.

More important than assessing the oral narrative is paying attention to the narrator, perceiving them as a biopsychosocial subject, and respecting their well-being. Only after identifying the need for improvement in the children's oral narrative skills, it is possible to provide a more effective intervention – which is why a standardized oral narrative test for the Brazilian population is relevant.

CONCLUSION

The results of this study suggest that the TNL, adapted to Brazilian Portuguese, is valid to assess the oral narrative of the target audience for which it is meant. Considering the interdependence between comprehension and production, the possibility of predominance in the assessment of one or another

dimension of language was discussed. The contexts of assessment were analyzed based on the extent to which they favored the children's performance and how they did so in the six tasks of the test. It was noticed that the absence of images caused greater difficulty to both understand and produce narratives.

The potential of the results must be considered for the future construction of a standardized measure to assess narratives. The TNL approach to assess narrative comprehension and production with three tasks in each dimension confirms the importance of this quantity and diversification, so that the evaluator can better understand the child's competence. The forms, contexts, and criteria that defined the oral narrative assessment must match to favor the comparison of the child's performance in all tasks. Thus, it is believed that the oral narrative assessment can provide consistent data for the differential diagnosis of oral language disorders, and hence, plan the care provided in the speech-language-hearing clinic and analyze the oral language development program in educational settings.

ACKNOWLEDGMENTS

This article is an integral part of the author's doctoral thesis in the Postgraduate Program in Psychology of the *Universidade Federal do Rio de Janeiro* (UFRJ). The author extends her gratitude to Prof. Jane Correa, for her supervision, and to the schools and their administrators, for authorizing and enabling the children's invaluable participation. My deepest gratitude for their time and cooperation!

REFERENCES

1. Perroni MC. *Desenvolvimento do discurso narrativo*. São Paulo: Martins Fontes; 1992.
2. Silva ACF, Ferreira AA, Queiroga BAM. Development of oral narrative and level of mother's education. *Rev. CEFAC*. 2014;16(1):174-86.
3. Labov W, Waletzky J. Narrative analysis: oral versions of personal experience. In: Helm J, editor. *Essays on the verbal and visual arts*. Seattle: University of Washington Press; 1967. p.12-44.
4. Rungrojsuwan S. The development of narrative macrostructure of Thai children. *rEFlections*. 2019;26(2):1-23.
5. Norbury CF, Bishop DVM. Narrative skills of children with communication impairments. *Int J Lang Commun Disord*. 2003;38(3):287-313.

6. Pinto G, Tarchi C, Bigozzi L. The relationship between oral and written narratives: a three-year longitudinal study of narrative cohesion, coherence, and structure. *Br J Educ Psychol.* 2015;85(4):551-69.
7. Pinto G, Tarchi C, Bigozzi L. Development in narrative competencies from oral to written stories in five to seven-year-old children. *Early Child Res Q.* 2016;36:1-10.
8. Almeida ANS, Freitag RMK. Narrativas de alfabetizandos e a competência linguística. *Linguagem & Ensino.* 2014;17(2):551-71.
9. Wajskop G. Linguagem oral e brincadeira letrada nas creches. *Educação & Realidade.* 2017;42(4):1355-74.
10. Silva LC, Lamônica DAC, Hage SRV. Instruments to assess children's language and speech translated and adapted into Brazilian Portuguese: an integrative literature review. *Rev. CEFAC.* 2021;23(2):e12520.
11. Gillam RB, Pearson NA. *Test of Narrative Language (TNL).* Austin: Pro-Ed; 2004.
12. Domsch C, Richels C, Saldana M, Coleman C, Wimberly C, Maxwell L. Narrative skill and syntactic complexity in school-age children with and without late language emergence. *Int J Lang Commun Disord.* 2012;47(2):197-207.
13. Pearce WM, Williams C, Steed W. Dialectal grammatical differences in oral narratives of school-aged indigenous Australian children. *Int J Speech Lang Pathol.* 2015;17(4):335-45.
14. Squires KE, Lugo-Neris MJ, Peña ED, Bedore LM, Bohman TM, Gillam RB. Story retelling by bilingual children with language impairments and typically developing controls. *Int J Lang Commun Disord.* 2013;49(1):60-74.
15. Gillam S, Olszewski A, Fargo J, Gillam R. Classroom-based narrative and vocabulary instruction: results of an early-stage, nonrandomized comparison study. *Lang Speech Hear Serv Sch.* 2014;45(3):204-19.
16. Timler GR. Use of the children's Communication Checklist-2 for classification of language impairment risk in young school-age children with attention-deficit/hyperactivity disorder. *Am J Speech Lang Pathol.* 2014;23(1):73-83.
17. Costa GM, Rossi NF, Giacheti CM. Performance of Brazilian Portuguese speakers in The Test of Narrative Language (TNL). *CoDAS.* 2018;30(4):e20170148.
18. Santos G dos. *Adaptação transcultural e evidências de validade do Test of Narrative Language para o português brasileiro [dissertation].* Rio de Janeiro (RJ): Universidade Federal do Rio de Janeiro; 2019.
19. Zaccaron R, D'Ely RCSF, Xhafaj DCP. Estudo piloto: um processo importante de adaptação e refinamento para uma pesquisa quase experimental em aquisição de L2. *Revista do GELNE.* 2018;20(1):30-41.
20. Menezes MLN. *Avaliação do desenvolvimento da linguagem.* Rio de Janeiro (RJ): Editora Autor; 2004.
21. Rossi NF, Lindau TA, Gillam RB, Giacheti CM. Cultural adaptation of the Test of Narrative Language (TNL) into Brazilian Portuguese. *CoDAS.* 2016;28(5):507-16.
22. ABEP: Associação Brasileira de Empresas de Pesquisa [homepage on the internet]. Critério de classificação econômica Brasil – Critério Brasil; 2015. [accessed on 2021 Jul 29]. Available at: <http://www.abep.org/criterio-brasil>
23. Prando ML, Jacobsen GM, Gonçalves HA, Siqueira LS, Moraes AL, Pureza J et al. Discurso narrativo oral infantil. In: Fonseca RP, Prando ML, Zimmermann N, editors. *Avaliação de linguagem e funções executivas em crianças - Tarefas para Avaliação Neuropsicológica.* São Paulo (SP): Memmon; 2016. p. 106-38.
24. Brandão ACP, Spinillo AG. Aspectos gerais e específicos na compreensão de textos. *Psicologia: Reflexão e Crítica.* 1998;11(2):0.
25. Dockrell JE, Marshall CR. Measurement issues: assessing language skills in young children. *Child Adolesc. Psychiatry Ment. Health.* 2015;20(2):116-25.
26. High PC, Klass P. Literacy promotion: an essential component of primary care pediatric practice. *Pediatrics.* 2014;134(2):404-9.
27. Silveira HG, Brocchi BS, Perissinoto J, Puglisi ML. Tutoring effects in the narrative skills of typically developing children. *CoDAS.* 2019;31(2):1-8.
28. Vygotsky LS. *A formação social da mente.* São Paulo (SP): Martins Fontes; 1991.
29. Dias NM, Bueno JOS, Pontes JM, Mecca TP. Linguagem oral e escrita na Educação Infantil: relação com variáveis ambientais. *Psicologia Escolar e Educacional.* 2019;23:1-10.

30. Souza MSL, Cáceres-Assenço AM. Do vocabulary and narrative skills correlate in preschoolers with typical language development? *CoDAS*. 2021;33(6):1-7.
31. Repina TA. Development of imagination. In: Zaporozhets AV, Elkonin DB, editors. *The psychology of preschool children*. Cambridge: MIT Press; 1971. p. 255-77.
32. Spinillo AG. O efeito da representação pictográfica na produção de narrativas. *Psicologia: Teoria e Pesquisa*. 1991;7(3):311-26.
33. Spinillo AG. Era uma vez... e foram felizes para sempre: esquema narrativo e variações experimentais. *Temas em Psicologia*. 1993;1(1):67-77.
34. Joffe VL, Cain K, Marić N. Comprehension problems in children with specific language impairment: does mental imagery training help? *Int J Lang Commun Disord*. 2007;42(6):648-64.
35. Kim Y-SG. Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *J Exp Child Psychol*. 2016;141:101-20.
36. Soares SM, Silva BC. Relações temporais na narrativa infantil em dados de fala de português brasileiro. *Linguagens & Cidadania*. 2017;19(n especial):1-13.
37. Hudson JA, Shapiro LR. From knowing to telling: the development of children's scripts, stories, and personal narratives. In: McCabe A, Peterson C, editors. *Developing narrative structure*. Hillsdale: Lawrence Erlbaum Associates; 1991. p. 89-136.
38. Berman R. On the ability to relate events in narrative. *Discourse Process*. 1988;11(4):469-97.
39. Gillam RB, Pearson NA. *Test of Narrative Language (TNL-2)*. Austin: Pro-Ed; 2017.
40. Kapantzoglou M, Fergadiotis G, Restrepo MA. Language sample analysis and Elicitation technique effects in bilingual children with and without language impairment. *J Speech Lang Hear Res*. 2017;60(10):2852-64.