

ORIGINAL ARTICLE

Open Access



Translation, cultural adaptation, and validation of an Arabic version of the test of narrative language—second edition

Sara Magdy Ibrahim^{1*} , Ossama Ahmed Sobhy², Riham Mohamed ElMaghraby¹ and Nesrine Hazem Hamouda¹

Abstract

Background The significance of narrative skills is evident due to their role in the development of language and their connection to significant social and academic skills. This study aimed to translate, adapt, and validate the Test of Narrative Language-Second Edition (TNL-2) for its use as a tool for the assessment of narrative language in Arabic-speaking Egyptian children. In a cross-sectional study design, the Arabic-translated version of the TNL-2 was administered to 200 typically developing Arabic-speaking Egyptian children ranging in age from 4 years to 15 years and 11 months for validation. The participants were categorized according to their age into ten groups and their scores were analyzed. Face validity was assessed by asking five expert phoniatricians to review the Arabic version of the TNL-2 and complete a questionnaire that assessed the test's effectiveness in measuring different narrative skills.

Results A statistically significant difference was found when comparing the TNL-2 scores among the age groups under study. In addition, there was a significant correlation between standardized Arabic language test scores and the total comprehension and total production subtests' raw scores of the TNL-2. The test-retest reliability and inter-rater agreement demonstrated a high level of reliability and inter-rater agreement. Experts have reached a consensus that the Arabic version of the TNL-2 is capable of evaluating the primary microstructural and macrostructural components of Arabic narratives. Furthermore, it can provide insights into the overall narrative skills of Egyptian Arabic-speaking children.

Conclusion The Arabic-translated version of the TNL-2 demonstrated validity and reliability as an instrument for assessing narrative language comprehension and production skills in Arabic-speaking Egyptian children.

Keywords Narrative language development, Narrative assessment, Arabic storytelling

Background

A narrative refers to the ability to produce a fictional or factual account of meaningful, chronologically sequenced occurrences and experiences [1]. Narrative skills play a crucial role in language development and are closely connected to critical academic skills, namely, reading, comprehension, and writing [2–5]. Narratives are crucial for developing proficient social skills, as evidenced by the observation that children with delayed language development typically exhibit less proficient social communication skills [6].

*Correspondence:

Sara Magdy Ibrahim
s_ibrahim16@alexmed.edu.eg

¹ Phoniatrics Unit, Department of Otorhinolaryngology, Faculty of Medicine, Alexandria University, Champollion street, Alazarita, Alexandria 21131, Egypt

² Audiovestibular Unit, Department of Otorhinolaryngology, Faculty of Medicine, Alexandria University, Champollion street, Alazarita, Alexandria 21131, Egypt



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Studies have shown that narrative competence increases with age alongside language, cognitive, and social skills [5, 7, 8]. Development of narrative abilities commences during the preschool period, typically around the age of 2, progresses throughout the school-age years, and continues to develop through adolescence and even adulthood [5, 7].

Language development is reflected in narrative skills, as the narrator must utilize age-appropriate linguistic abilities to communicate the primary narrative events, including the central theme. Furthermore, linguistic abilities are utilized to express the primary characters' affective states that motivate them to carry out specific actions [9].

In the early stages of development, narratives establish connections between the language used in various contexts, such as the language spoken at home and the language used in schools and literary contexts [10]. Several studies have demonstrated the role of narrative language in predicting academic skills, precisely reading comprehension, writing, and mathematics [2–4]. Likewise, children with poor reading and comprehension abilities have delayed narrative comprehension and production skills compared to their typically developing peers [11].

Delayed narrative skills are evident in various communication disorders, including developmental language disorder, hearing impairment, intellectual disabilities, and autism spectrum disorder [12, 13].

The narrative structural organization consists of macrostructure and microstructure. The macrostructural level reflects story episodes based on Stein and Glenn's story grammar model 1979 [14]. The model analyzes a narrative into a number of temporally sequenced episodes. Episodes encompass a specific setting (person, place, and time), a beginning (a problem that motivates actions), an internal response of the characters (emotions), actions (attempts to solve the problem), and an ending (resolution). These elements comprise the primary components of the narrative, and traditional storytelling requires the incorporation of these pivotal episodes [14]. On the other hand, microstructural level refers to the utilization of productive and complex linguistic elements, including compound sentences, temporal and causal subordinate clauses, adverbial phrases, and adjectives [15]. By integrating all linguistic components and employing intricate microstructural elements, the story's grammatical elements (macrostructure) are clarified, meaning is communicated, and additional details are incorporated [16].

Assessment of narrative skills

Narrative abilities are assessed in terms of macrostructure and microstructure [17]. According to the existing

literature, various tasks have been used to assess narrative language. Research has determined that both storytelling and retelling tasks are adequate for assessing narrative skills as both methods have advantages that rely on the cognitive and linguistic abilities necessary for storytelling [18]. Story-retelling tasks focus on story structure, vocabulary acquisition, articulation, retrieval, and comprehension [18]. The retelling requires the narrator to possess a deep understanding of the original story to retell it accurately [15]. As a result, narratives serve as a means of coordinating sequencing, intricate language, pragmatic competence, and conceptual thinking [19]. Difficulties in retelling tasks arise from children's reluctance to generate the target vocabulary and difficulty scoring [20].

In contrast, storytelling tasks entail newly generated stories encompassing personal or fictional narratives [1]. Personal narratives are accounts of past experiences using characters and temporally coordinated events that might include problems and attempts to solve them [21]. A script is a relatively uncommon form of personal narrative in which an individual is expected to recount regularly recurring events based on multiple personal experiences, such as describing the typical way in which a person spends their holidays [22]. This discourse is characterized by its elaborative nature, as it centers around the narrator's personal experiences and recollections related to a specific subject rather than focusing on a specific occurrence [22].

Fictional narratives entail recounting a story constructed from fabricated events, which are not factual and are derived from the storyteller's imagination. Fictional narratives encompass untold stories, including those prompted by readily available stimuli, such as pictures [23].

Various assessment tools are available for English-speaking children. Examples include the Narrative Assessment Protocol by Bowles et al. [24], Assessment of Story Comprehension by Spencer and Goldstein [25], The Monitoring Indicators of Scholarly Language S. L. Gillam et al. [7], and the Test of Narrative Language by Gillam and Pearson 2004 [26].

The first version of the TNL [26] was designed to evaluate the narrative abilities of children between the ages of 5 and 11 years 11 months in terms of their understanding and creation of narratives, utilizing both actual and fictional stories. The initial iteration of the narrative language test (TNL) consists of two subtests: narrative comprehension and oral narration subtest. Each subtest includes three tasks presented in three different formats: no picture, sequenced pictures, and a single picture. Narrative comprehension is evaluated by presenting a story for each of the different formats to which the child is

required to listen and then answer comprehension questions that contain literal questions that require the child to recall information presented in the story regarding the main story elements, such as characters' names, setting, and main problem, in addition to inferential questions that evaluate the ability of the child to make inferences beyond what was explicitly mentioned in the story. The Child's answers are scored according to the examiner's manual provided [26]. The assessment of narrative production involves three tasks: retelling the first story with no picture, producing a story with a sequence of pictures, and with a single scene picture. Story production is scored for both story content, macrostructure, and microstructure [26].

The authors published the second version of the TNL, TNL 2, in 2017 [27]. This version will be further discussed in the methods section. The primary distinction in the format lies in the inclusion of a picture in the TNL 2 version, whereas the initial version of the McDonald's story lacks a picture. Furthermore, the age range being evaluated is extended to encompass individuals between the ages of 4 and 15 years and 11 months.

TNL has been used in research to assess narrative skills in children with delayed language development, to assess the effects of narrative intervention, to correlate with other measures of language evaluation and working memory, and to predict academic performance in relation to narrative skills level [28–31].

Narratives can be different among cultures. Spanish-speaking children often emphasize the primary characters' internal responses [32]. Japanese children's narratives lean towards producing brief, concise stories as they combine multiple experiences without much elaboration. In contrast, North American children often tell detailed narratives about a single event [33]. Despite cultural variations, certain story elements consistently stand out, including the introduction of main characters, setting, timeframe, and the existence of a problem requiring a solution [19].

Narrative research in Arabic

Narrative language development of children was previously assessed in several languages, including Arabic [34]. About 422 million people speak Arabic, including non-native speakers or those who speak Arabic as a second language. Most native Arabic speakers are present in Egypt, with a population of over 100 million. Several Arabic dialects exist, with the Egyptian, Maghrebi, and Gulf being the most commonly spoken forms [35].

A special property of the Arabic language is diglossia, which refers to the use of two forms of the language by its speakers for different social situations: Colloquial Arabic (Spoken Arabic) and Modern Standard Arabic

(MSA) [36]. Spoken Arabic is frequently utilized in daily interactions, while MSA (Modern Standard Arabic) is the formal variant utilized in educational contexts, writing, and formal events [37]. The spoken form is the first to be developed as it is used in everyday context and the variety in which the narrative skills are first developed. The standard form is the formal variety used in literal and academic situations used in formal contexts. It is usually first encountered by children in the academic context, in reading and writing, or earlier in the preschool years through their exposure to the media [37, 38]. The two forms are similar regarding various aspects [39]. However, the distance between the Spoken and the Standard forms of Arabic affects the phonological and lexical domains most [40].

The assessment of narrative language skills in Egyptian Arabic has not received much attention. To the best of our knowledge, only one study by Safwat et al. [41] targeted the assessment of narrative skills in preschool children. The objective of the study was to create a battery for evaluating narrative language. This involved having children retell a story using a series of pictures without words and then analyzing both the microstructure and macrostructure of their narrative production. The study comprised a cohort of 60 children, ranging in age from 2 to 6 years, who were native speakers of Egyptian Arabic. The child's performance was evaluated based on the organization of the story, which included elements such as the introduction of the setting and topic, the chronological order of events, the use of references, and the coherence of the narrative. The study findings indicated that the initial component of narrative organization to emerge at the age of 2 years was the utilization of basic verbs to depict action and setting. The study examined various components of language structure, including adjectives and the utilization of prepositions, as well as sentence structure, such as the utilization of simple and compound sentences. They observed a rise in the intricacy of sentences generated as individuals grew older, including the utilization of verb tense and various noun forms [41]. Narrative productivity was assessed by calculating the total number of words, mean length of utterance, and type-token ratio, which refers to the number of different words in relation to the total number of words produced by the child. The mean length of utterance increased across the different age groups [41]. The study was limited in scope as it exclusively focused on preschool children, despite the fact that literature indicates that narrative skills continue to progress throughout adolescence and even into adulthood [5, 7]. Additionally, story comprehension was not assessed [41].

Another study by Khodeir et al. [42] reported on developing and standardizing a test for pragmatic skills

in Egyptian Arabic. The test assessed various pragmatic aspects in children aged 4 years through 10 years, including narrative skills. Narrative skills were evaluated by assessing the child's comprehension of the main story elements and the ability to answer questions about four stories. Narrative production was assessed by eliciting storytelling and retelling from pictures. The study reported a positive correlation between the children's scores and age. Nevertheless, the test did not focus on thoroughly evaluating narrative language proficiency, and the comprehension questions used were mainly literal.

Other studies investigating Arabic narrative production include a study by Ravid et al. [43] which examined the narrative skills of 97 monolingual Palestinian Arabic-speaking children using a story-retelling task. The study concluded that the length of narratives increased with age across seven different age groups. Interestingly, they reported using both the standard Arabic and the Spoken Arabic forms even in younger preschool children, with an increase in the complexity of the lexical and morphosyntactic structure as the grade level increased.

A recent study by Kavar et al. [44] investigated the narrative skills of 30 monolingual Palestinian Arabic-speaking preschool children by comparing the story comprehension and retelling abilities in both its Spoken and Modern Standard forms, focusing on microstructure and macrostructure. Unlike other studies, their findings indicated superior narrative comprehension in the MSA form (except for the theory of mind questions), while better production was observed in the Spoken Arabic form.

Asli-Badarneh et al. [45] conducted a study to assess the narrative abilities of 75 Arabic-speaking Canadian immigrants aged 7 to 12 years, using an Arabic translation of the Test of Narrative Language (TNL). The study aimed to investigate the impact of diglossia on their language skills. The study specifically examined the relationship between microstructure and macrostructure, focusing on the impact of the Standard Arabic lexicon on microstructural elements and its ability to predict macrostructure.

After reviewing the previous studies that focused on the effect of diglossia, it was concluded that better narratives are produced by children in their Spoken form rather than in MSA regarding the length of the story and morphosyntax, and narrative comprehension is easier in Spoken Arabic [43–45]. Research shows that when oral language skills are examined, the Spoken variety is the one in which speakers are more proficient [46]. Therefore, our translation of the TNL-2 of the stories and the comprehension questions directed to the child were in the context of the Spoken form of Egyptian Arabic. The instructions to the clinician and the scoring sheet were translated into MSA.

Based on this brief review of narrative research in Arabic, it is evident that there is a need for further studies focusing on Egyptian Arabic. Considering the significance of narrative language skills in language development and social interactions, a comprehensive assessment tool for narrative language is required. Due to the dearth of research targeting the assessment of narrative language skills of Arabic-speaking Egyptian children, this study focused on the translation, adaptation, and validation of the Test of Narrative Language-Second Edition (TNL-2) [27] for its use in assessing narrative language in Arabic-speaking Egyptian children. Our research question was: Do narrative language skills vary across different age groups, and is the translated version of the TNL-2 a valid and reliable tool to assess the overall narrative skills of an Egyptian Arabic-speaking child?

Methods

The study proceeded in the following steps:

Translation and adaptation of the TNL-2[27].

The translation and cultural adaptation process was carried out in accordance with the principles of good practice, following the subsequent steps [47]:

- 1- Preparation: Prior to translating the TNL-2, permission was obtained from the publisher to translate and culturally adapt the test, as well as its administration to the designated number of participants.¹
- 2- Forward translation and reconciliation: Forward translation was then carried out from the original language (English) to the target language (Arabic) by two independent bilingual certified translators; both were native speakers of the target language. The two forward-translated versions were then reconciled to resolve any discrepancies between the translated versions through an independent native speaker of the Arabic language who had not been involved in any of the forward translations. The translation was mainly intended to capture the concept rather than being a literal translation. Furthermore, the stories and the questions directed toward the child were translated into the spoken form of Egyptian Arabic, the instructions directed toward the clinician, and the scoring sheet were translated to MSA. A single forward translated version was created.

¹ TNL-2 is a commercially available tool. The permission from the publisher grants the authors permission to use the Arabic translated version of the TNL-2 and publish the results of the study. However, the permission prohibits sharing the original or translated version of the test; therefore, the test cannot be provided in the manuscript for review.

Table 1 Adaptations made to the tasks of the TNL-2 in the Arabic version

Task	Original English version of TNL-2	Adapted version of TNL-2	Comments
McDonald's story Tasks 1 and 2	McDonald's story	McDonald's restaurant story	Even though McDonald's is widely available across Egypt, some of the young children were not familiar with it in the pilot study. In the scoring section for narrative production in task 2 both "restaurant" or "McDonald's" were counted as correct.
	Lisa and Raymond Chicken nuggets	Sara and Ahmed Chicken	Common Egyptian first names. There is no Egyptian Arabic translation for Chicken nuggets. Chicken was used instead.
	Large vanilla milk shake	Large juice	For those who were not familiar with McDonald's, we found it more appropriate to use "juice" instead of "milkshake" which is a more common order at restaurants.
Shipwreck Task 3	Shipwreck	Broken ship	The literal translation of Shipwreck is in Modern Standard Arabic, we found the translation to the Spoken Arabic form, an equivalent to "broken ship" easier to understand.
	Maria On the weekend (Saturday/Sunday)	Mariam On the weekend (Friday/Saturday)	A common name in Egypt. The days of the weekend were changed in accordance with the days of the weekend in Egypt (Friday and Saturday)
	Monday	Sunday	The answer to the question "when did Maria take her ship to school" was changed to Sunday which is the first school day of the week in Egypt
	A	A or full mark	The grading system is different among the Egyptian schools. Grading using numerical marks is more common in Egyptian national schools. During scoring of the comprehension questions, both A or full mark such as 10/10 were counted as correct.
Late for school Task 4			No changes were done to adapt the story, it was found culturally appropriate.
Treasure Task 5	Erika and Michael	Rana and Karim	Common Egyptian first names.
	Michael snuck up behind a large rock	Karim hid behind a large rock	The translation of the expression "snuck up" is not commonly used in Spoken Arabic.
	Rolled their eyes	They found it strange and laughed	The literal translation of the expression "rolled their eyes" is not commonly used to show annoyance in Egyptian Arabic. We found it more appropriate that they expressed their disbelief by finding it strange and laughable.
Aliens Task 6			No changes were done to the original version.

3- Backward translation and review: A certified translator made a backward translation of the agreed upon forward translated version to translate the Arabic version back to the test's original language (English). The research team reviewed the backward translated version by two expert phoniatricians in the field of speech-language pathology to reach a final version and confirm the cultural appropriateness of the translated version. Adaptations made to the tasks of the TNL-2 in the Egyptian Arabic translated version are shown in Table 1.

Face validity

Face validity was assessed by presenting the Arabic-translated version of the TNL-2 to five experienced phoniatricians with at least 10 years of experience in child speech-language pathology and asking them to answer a questionnaire about the ability of the TNL-2 to assess different narrative comprehension and production skills, by giving a score from 1 to 5, denoting a poor to excellent ability of the TNL-2 to assess the skills in question.

Pilot study

A pilot study was done on 20 participants to ensure the test's clarity and cultural appropriateness, the children's

ability to understand the stories, and the examiner's ability to administer the test. Additionally, reliability and inter-rater agreement were investigated on the 20 participants (data for the pilot study are included as [Supplementary material](#)).

Application of the Arabic-translated, culturally-adapted version of the TNL-2

Our research was conducted in a cross-sectional design on 200 typically developing children aged 4.0 through 15.11 years. The study participants were divided into ten groups according to their age. Each group consisted of 20 participants.

Inclusion criteria: Typically developing children aged 4 years through 15 years, 11 months.

Exclusion criteria:

1. Children with delayed language development, currently or by history.
2. Children with intellectual disability.
3. Children with hearing or visual impairment.
4. Children with neurodevelopmental disorders, such as ADHD.

The participants were selected from the relatives of the patients attending the outpatient clinic of the phoniatics unit to assess the validity and reliability of the translated Arabic version of the TNL-2. The study was conducted from April 2022 to March 2023.

All participants were assessed by the following protocol of evaluation:

- a. Elementary diagnostic procedures:
 - History taking which included personal data and history of delayed language development or other developmental disorders.
- b. Clinical diagnostic procedures:
 - Psychometric evaluation by Stanford Binet Scale 4th edition to assess intelligence quotient and mental age in order to exclude intellectual disability. An Arabic validated version of the Stanford Binet scale was used in the study, and the scores obtained by the children were given according to Arabic norm-referenced measures [48].
 - Standardized Arabic language test, as a screening tool for the children's linguistic skills and for exclusion of delayed language development [49]. The Arabic language test is formed of five subtests: semantics, expressive language, receptive language, pragmatics, and prosody. The child's score for each subtest and the total score were compared to the means for the child's age group to ensure

all participants' linguistic skills were adequate for their age. The test was completed in about 20 min. The test was administered in the Spoken Arabic form.

- Finally, the Arabic-translated version of the TNL-2 was applied.

Description of the TNL-2

The TNL-2 is a measure of comprehension and production of connected speech used to tell stories. It assesses children's ability from 4 years 0 months through 15 years and 11 months to tell and comprehend three types of stories: scripts, personal narratives, and fictional narratives. The TNL-2 consists of six tasks organized into two subtests (comprehension and production). The comprehension and production tasks are presented alternatively. Comprehension tasks include Task 1, Task 3, and Task 5, while production tasks comprise Task 2, Task 4, and Task 6.

Comprehension subtest

The comprehension subtest comprises task 1, "McDonald's story," task 3, "shipwreck story," and task 5, "Treasure story." The stories are narrated to the child, and they are required to listen carefully and answer comprehension questions that assess the ability to recall essential story elements and events (such as the name of the characters, time, and the main problem). The questions also assess the ability of the child to make inferences and non-literal interpretations about the story.

Production subtest

The production subtest is also composed of three tasks: task 2 is retelling the "McDonald's story." The second production task is task 4, "late for school," in which the child is required to generate a story based on a sequence of five pictures. The third and final production task is task 6, "Aliens," in which the child is required to generate a fictional narrative based on a picture.

Test timing and scoring

After listening to each story, the child is asked questions in the comprehension subtest. The child receives 1 point for every correct answer. The first task has a maximum score of 20. Task 3 has a maximum score of 14. Lastly, task 5 has a maximum score of 13. The maximum total raw score for the comprehension subtest is 47.

The production subtest is evaluated by listening to a voice recording of the child's speech at least three times. Every production is evaluated based on its narrative content and grammatical elements. The child is assigned one point for every story element that is mentioned. The evaluated story grammar elements include the utilization

of temporal relations, causal relations, accurate grammar, dialogue inclusion, sequencing, and complete episodes.

Task 2 has a maximum score of 31, task 4 has a maximum score of 25, while task 6 has a maximum score of 30. The maximum total production raw score is 86.

The results of the three comprehension tasks are combined to form a total comprehension raw score. Similarly, the results of the three production tasks form a production subtest raw score. Raw scores are converted to scaled scores and percentile ranks according to the child's age. Age equivalents can also be obtained. The comprehension and production scaled scores are combined to form a total scaled score from which a composite (narrative language ability index) can be obtained. Descriptive terms are used to describe the scaled scores and narrative language ability index ranging from very poor to very superior.

A digital voice recorder was used to record the entire test during its application. The recordings were replayed later to fill the scoring sheet as appropriate. Administration of the test required about 15–20 min. However, the scoring time varied from one participant to another, as the recordings of the production subtests were required to be replayed at least three times to be scored appropriately.

Reliability testing

The 20 participants of the pilot study were reassessed after 2 weeks to obtain test-retest reliability. Additionally, two independent expert-phoniatricians were asked to listen to the recordings of the 20 participants in the pilot study and score the participants separately to test for inter-rater agreement.

Statistical methodology

The data were analyzed with the Statistical Package for the Social Science (SPSS) for version 25 (SPSS Inc, Chicago, IL). Kolmogorov–Smirnov's test of normality revealed significance in the distribution of most variables, so non-parametric statistics were adopted. Data were described using minimum, maximum, mean, standard deviation, standard error of the mean, 95% CI of the mean, median, 95% CI of the median, 25th–75th percentile, and inter-quartile range. Categorical variables were described using frequency and percentage. Comparisons were carried out between more than two independent, not-normally distributed subgroups using the Kruskal–Wallis test. Pearson's correlation was used. Intra-class correlation (ICC) was also used to assess agreement. Cicchetti guidelines were utilized for the evaluation of the ICC coefficient value. Interpretation for ICC: Cicchetti (1994) gives the following often quoted guidelines for interpretation for kappa or ICC inter-rater agreement

measures: Less than 0.40 (poor), between 0.40 and 0.59 (fair), between 0.60 and 0.74 (good), and between 0.75 and 1.00 (excellent). During sample size calculation, beta error accepted up to 20% with a power of study of 80%. An alpha level was set to 5% with a significance level of 95%. Statistical significance was tested at p -value $< .05$.

Results

Distribution of the participants according to sex, Stanford Binet scale, and Arabic language test results

Table 2 shows the distribution of the study participants in 10 age groups and their sex distribution. The results indicated that the children demonstrated at least average intelligence and overall general IQ on the Stanford Binet subtests. The participants' results on the standardized Arabic language test were within the range considered "adequate for their age" when compared to the normative data.

Face validity

The summary of the responses from the five expert phoniatricians is displayed in Table 3. All experts unanimously concurred that the TNL-2 possesses exceptional proficiency in comprehensively understanding a child's narrative abilities at a specific age. It also evaluates the child's capacity to generate narratives' fundamental microstructural and macrostructural components.

Application of the TNL-2 results

Median scores for the comprehension subtests and the total raw score for comprehension are reported in Table 4. Median scores for the production subtests and production total raw score are reported in Table 5.

The Kruskal–Wallis test revealed a statistically significant difference in all assessed subtests of the TNL-2 across different age groups. This difference was reflected in a statistically significant increase in raw scores for both the comprehension and production subtests across the age groups.

A statistically significant difference was found between the assessed age groups regarding the raw scores of the comprehension subtest. A statistically significant difference was found in the McDonald's story raw scores ($p < 0.001$), the shipwreck story raw scores ($p < 0.001$), and the Treasure story raw scores ($p < 0.001$). The total comprehension raw score obtained by combining the raw scores of the three previously mentioned stories showed a statistically significant difference between the age groups ($p < 0.001$), as depicted in Table 4.

A statistically significant difference was found between the assessed age groups regarding the raw scores of the production subtest. A statistically significant difference was found in the McDonald's story raw scores ($p < 0.001$),

Table 2 Sex distribution, Stanford Binet, and Arabic language test results in the studied age groups

Sex	n (%)	Age groups (years)												Test of significance p value						
		(4 - <5) (n=20)	(5 - <6) (n=20)	(6 - <7) (n=20)	(7 - <8) (n=20)	(8 - <9) (n=20)	(9 - <10) (n=20)	(10 - <11) (n=20)	(11 - <12) (n=20)	(12 - <14) (n=20)	(14 - <16) (n=20)									
		M	F	M	F	M	F	M	F	M	F	M	F	M	F					
Stanford Binet Verbal IQ	- Min-Max - Median - 95% CI for median	14 (70%) 92.00-118.00 103.50 98.00-112.00	6 (30%) 90.00-115.00 98.00 96.00-104.00	4 (20%) 90.00-118.00 98.00 96.00-104.00	16 (80%) 90.00-115.00 98.00 96.00-104.00	6 (30%) 92.00-116.00 102.00 95.00-110.00	14 (70%) 90.00-116.00 98.00 94.00-106.00	8 (40%) 90.00-120.00 98.00 94.00-100.00	14 (70%) 90.00-110.00 98.00 98.00-104.00	6 (30%) 90.00-110.00 98.00 98.00-104.00	14 (70%) 90.00-110.00 98.00 98.00-104.00	6 (30%) 90.00-110.00 98.00 98.00-104.00	13 (65%) 90.00-104.00 96.00 90.00-100.00	7 (35%) 90.00-104.00 96.00 90.00-100.00	12 (60%) 90.00-116.00 94.00 90.00-98.00	8 (40%) 90.00-116.00 94.00 90.00-98.00	10 (50%) 90.00-103.00 90.00 -	11 (55%) 90.00-104.00 94.00 92.00-97.00	9 (45%) 90.00-104.00 94.00 92.00-97.00	$H_{(KW)}(df=9)=43.642$ $p<0.001^*$
Stanford Binet Abstract IQ	- Min-Max - Median - 95% CI for median	90.00-110.00 98.00 96.00-102.00	90.00-118.00 102.00 98.00-110.00	90.00-120.00 100.00 94.00-114.00	90.00-125.00 98.00 96.00-108.00	90.00-130.00 96.00 94.00-110.00	90.00-128.00 94.00 92.00-104.00	90.00-118.00 96.00 92.00-104.00	90.00-120.00 100.00 94.00-108.00	90.00-110.00 98.00 96.00-100.00	90.00-110.00 98.00 96.00-100.00	90.00-110.00 98.00 96.00-100.00	90.00-110.00 98.00 96.00-100.00	90.00-108.00 92.00 92.00-98.00	90.00-110.00 98.00 96.00-100.00	90.00-110.00 98.00 96.00-100.00	90.00-102.00 95.50 92.00-98.00	90.00-101.00 93.00 92.00-96.00	90.00-101.00 93.00 92.00-96.00	$H_{(KW)}(df=9)=21.952$ $p=0.009^*$
Stanford Binet General IQ	- Min-Max - Median - 95% CI for median	90.00-114.00 101.00 99.00-107.00	90.00-117.00 102.00 98.00-103.00	92.00-117.00 99.00 97.00-111.00	90.00-123.00 99.00 98.00-104.00	90.00-120.00 97.50 93.00-101.00	90.00-117.00 98.00 94.00-103.00	90.00-111.00 97.00 93.00-101.00	90.00-110.00 99.50 93.00-103.00	90.00-110.00 99.50 93.00-103.00	90.00-110.00 99.50 93.00-103.00	90.00-110.00 99.50 93.00-103.00	90.00-110.00 99.50 93.00-103.00	90.00-102.00 95.50 92.00-98.00	90.00-102.00 95.50 92.00-98.00	90.00-102.00 95.50 92.00-98.00	90.00-102.00 95.50 92.00-98.00	90.00-101.00 93.00 92.00-96.00	90.00-101.00 93.00 92.00-96.00	$H_{(KW)}(df=9)=31.434$ $p<0.001^*$
Arabic Language Test (Total raw score)	- Min-Max - Median 95% CI for median	283.00-332.00 317.00 293.00-323.00	289.00-342.00 322.50 319.00-327.00	304.00-345.00 331.00 322.00-339.00	330.00-345.00 340.50 338.00-345.00	339.00-345.00 345.00 345.00-345.00	342.00-345.00 345.00 345.00-345.00	345.00-345.00 345.00 345.00-345.00	342.00-345.00 345.00 345.00-345.00	342.00-345.00 345.00 345.00-345.00	342.00-345.00 345.00 345.00-345.00	336.00-345.00 345.00 345.00-345.00	345.00-345.00 345.00 345.00-345.00	345.00-345.00 345.00 -	345.00-345.00 345.00 -	341.00-345.00 345.00 345.00-345.00	345.00-345.00 345.00 345.00-345.00	339.00-345.00 345.00 345.00-345.00	339.00-345.00 345.00 345.00-345.00	$H_{(KW)}(df=9)=148.790$ $p<0.001^*$

n number of patients, Min-Max minimum-maximum, Std. Dev. standard deviation, CI confidence interval, KS Kolmogorov-Smirnov, H:KW Kruskal-Wallis H, NS statistically not significant (p ≥ 0.05)

* Statistically significant (p < 0.05)

Table 3 Summary of the rating of the five expert phoniatricians to the TNL-2 face validity questionnaire

Please give each item a score regarding the ability of the TNL-2 to achieve the following items:	1 Poor 2 Uncertain 3 Fair 4 Good 5 Excellent				
	1 n(%)	2 n(%)	3 n(%)	4 n(%)	5 n(%)
1. Assess the auditory comprehension of a story sequence				2(40%)	3(60%)
2. Assess the ability of the child to answer literal comprehension questions in relation to a story				2(40%)	3(60%)
3. Assess the ability of the child to make non literal interpretation of events/ make inferences in relation to a story				2(40%)	3(60%)
4. Assess the ability of the child to understand the main problem/ dilemma in a story				2(40%)	3(60%)
5. Assess the ability of the child to retell a story				2(40%)	3(60%)
6. Assess the ability of the child to produce a complete narrative episode with its minimum requirements (beginning, middle, and ending)				2(40%)	3(60%)
7. Assess the ability of the child to produce temporal and causal relations in a story sequence				1(20%)	4(80%)
8. Assess the ability of the child to understand and express different emotion and internal responses in relation to a story event		1(20%)	2(40%)		2(40%)
9. Assess the ability of the child to produce the main microstructure (morphosyntax) and macrostructure (story grammar) elements of narratives					5(100%)
10. Gives an idea about the overall narrative abilities of the child at a given age					5(100%)

the late-for-school story raw scores ($p < 0.001$), and the Aliens story raw scores ($p < 0.001$). The total production raw score, obtained by combining the raw scores of the three previously mentioned stories, showed a statistically significant difference between the age groups with a p -value ($p < 0.001$; Table 5).

Correlation between the Arabic language test scores and the TNL-2 scores

Figure 1 shows a strong positive correlation between the TNL-2 comprehension total raw score and the Arabic language test total raw score in 200 measurement points ($p < .0001$).

Figure 2 shows a strong positive correlation between the TNL-2 production total raw score and the Arabic language test total raw score in 200 measurement points ($p < .0001$).

Test reliability results

Test-retest reliability: (data provided as Supplementary material)

A very high positive correlation was found between the total comprehension raw score of the TNL-2 at the test and the retest times of assessment ($r = 0.977$, $p < .001^*$). An excellent degree of reliability was found between the total comprehension raw score of the Test of Narrative Language measurements. The single measure ICC was .970 with a 95% confidence interval from .908 to .989 ($F(19,19) = 80.543$, $p < .001$).

A strong positive correlation was found between the total production raw score of the Test of Narrative

Language at the initial assessment and the subsequent retest times ($r = 0.981$, $p < .001^*$). An excellent degree of reliability was found between the total production raw score of the Test of Narrative Language measurements. The single measure ICC was .970 with a 95% confidence interval from .839 to .991 ($F(19,19) = 106.179$, $p < .001$).

Inter-rater agreement (data provided as Supplementary material)

There was an excellent degree of inter-rater agreement between the total comprehension raw score of the TNL-2 measurements. The single measure ICC was .994 with a 95% confidence interval from .986 to .998 ($F(19,19) = 332.191$, $p < .001$).

In addition, there was an excellent degree of inter-rater agreement between the total production raw score of the TNL-2 measurements. The single ICC was .993 with a 95% confidence interval from .991 to .999 ($F(19,19) = 270.191$, $p < .001$).

Discussion

Narratives are regarded as a significant measure of language development and a means of structuring language comprehension, abstract reasoning, and sequencing of events [19]. Moreover, narratives are linked to social, literacy, and academic skills development [2–6].

To our knowledge, no currently available Egyptian Arabic tool for assessing narratives addresses the full range of narrative abilities and the broad age range assessed by the TNL-2. The primary objective of this study was to create an Egyptian Arabic version of the TNL-2, which

Table 4 Test of Narrative Language comprehension subtests and total comprehension raw score in the studied age group

	Age groups (years)										Test of significance p value	
	(4-<5) (n = 20)	(5-<6) (n = 20)	(6-<7) (n = 20)	(7-<8) (n = 20)	(8-<9) (n = 20)	(9-<10) (n = 20)	(10-<11) (n = 20)	(11-<12) (n = 20)	(12-<14) (n = 20)	(14-<16) (n = 20)		
McDonald's raw score	- Min-Max	3.00-14.00	5.00-13.00	5.00-19.00	8.00-19.00	8.00-19.00	9.00-18.00	6.00-18.00	11.00-20.00	11.00-18.00	11.00-17.00	$H_{(KW)}(df=9) = 105.513$
	- Median	6.00	9.00	10.00	11.50	12.00	13.50	13.50	15.50	15.00	15.00	$p < .001^*$
	- 95% CI for median	5.00-7.00	8.00-10.00	9.00-12.00	10.00-14.00	12.00-17.00	12.00-15.00	12.00-15.00	12.00-15.00	14.00-16.00	14.00-17.00	15.00-16.00
Shipwreck raw score	- Min-Max	1.00-8.00	1.00-9.00	3.00-10.00	5.00-12.00	5.00-14.00	7.00-13.00	5.00-14.00	9.00-14.00	7.00-13.00	7.00-13.00	$H_{(KW)}(df=9) = 116.833$
	- Median	4.50	5.00	7.00	7.50	9.00	10.00	10.00	11.00	11.00	11.00	$p < .001^*$
	- 95% CI for median	3.00-5.00	5.00-7.00	7.00-10.00	7.00-9.00	8.00-10.00	10.00-12.00	9.00-12.00	11.00-12.00	11.00-12.00	11.00-12.00	
Treasure raw score	- Min-Max	0.00-9.00	3.00-8.00	4.00-10.00	3.00-12.00	4.00-18.00	4.00-12.00	3.00-13.00	7.00-13.00	8.00-12.00	8.00-12.00	$H_{(KW)}(df=9) = 106.001$
	- Median	3.50	5.00	6.50	7.50	7.50	8.00	9.00	9.00	10.00	10.00	$p < .001^*$
	- 95% CI for median	3.00-6.00	5.00-6.00	5.00-7.00	6.00-8.00	6.00-9.00	8.00-9.00	9.00-10.00	9.00-11.00	10.00-11.00	10.00-11.00	
Comprehension total raw score	- Min-Max	4.00-31.00	13.00-28.00	13.00-39.00	18.00-43.00	22.00-43.00	24.00-39.00	14.00-43.00	29.00-43.00	30.00-41.00	29.00-39.00	$H_{(KW)}(df=9) = 128.697$
	- Median	13.00	18.50	22.00	27.50	28.00	31.00	32.50	36.00	35.00	36.00	$p < .001^*$
	- 95% CI for median	12.00-18.00	16.00-22.00	21.00-27.00	24.00-32.00	24.00-31.00	30.00-33.00	30.00-35.00	34.00-38.00	34.00-38.00	36.00-37.00	

n number of patients, Min-Max minimum-maximum, Std. Dev. standard deviation, CI confidence interval, KS Kolmogorov-Smirnov, H:KW Kruskal-Wallis H, NS statistically not significant (p ≥ 0.05)

* Statistically significant (p < 0.05)

Table 5 Test of Narrative Language production subtests and total production raw score in the studied age groups

	Age groups (years)										Test of significance p value	
	(4-<5) (n = 20)	(5-<6) (n = 20)	(6-<7) (n = 20)	(7-<8) (n = 20)	(8-<9) (n = 20)	(9-<10) (n = 20)	(10-<11) (n = 20)	(11-<12) (n = 20)	(12-<14) (n = 20)	(14-<16) (n = 20)		
McDonald's raw score	- Min-Max	4.00-13.00	7.00-18.00	8.00-25.00	8.00-23.00	8.00-24.00	7.00-24.00	12.00-24.00	16.00-25.00	15.00-27.00	16.00-27.00	$H_{(KW)}(df = 9) = 114.458$
	- Median	8.50	12.50	13.00	15.50	17.00	19.00	18.00	21.00	20.50	19.50	$p < .001^*$
	- 95% CI for median	6.00-10.00	11.00-15.00	11.00-16.00	14.00-19.00	16.00-20.00	18.00-21.00	17.00-20.00	20.00-23.00	20.00-23.00	18.00-22.00	19.00-21.00
Late for school raw score	- Min-Max	12.00-9.00	3.00-11.00	6.00-18.00	6.00-18.00	6.00-19.00	5.00-21.00	9.00-22.00	11.00-22.00	10.00-24.00	12.00-20.00	$H_{(KW)}(df = 9) = 114.458$
	- Median	6.00	6.00	9.50	13.00	13.00	14.50	17.50	18.00	17.00	16.50	$p < .001^*$
	- 95% CI for median	5.00-8.00	5.00-9.00	8.00-12.00	10.00-15.00	12.00-18.00	13.00-17.00	14.00-18.00	17.00-20.00	16.00-19.00	16.00-17.00	16.00-17.00
Aliens raw score	- Min-Max	1.00-9.00	3.00-12.00	4.00-18.00	5.00-21.00	3.00-19.00	5.00-21.00	7.00-22.00	7.00-25.00	8.00-23.00	3.00-22.00	$H_{(KW)}(df = 9) = 76.485$
	- Median	5.50	7.50	9.00	10.00	11.00	13.00	13.50	15.50	15.50	13.00	$p < .001^*$
	- 95% CI for median	5.00-6.00	6.00-9.00	7.00-13.00	8.00-14.00	9.00-14.00	13.00-18.00	11.00-16.00	12.00-18.00	12.00-18.00	11.00-16.00	11.00-16.00
Production total raw score	- Min-Max	7.00-28.00	11.00-36.00	24.00-58.00	20.00-60.00	18.00-61.00	17.00-65.00	34.00-58.00	36.00-64.00	40.00-66.00	37.00-61.00	$H_{(KW)}(df = 9) = 121.694$
	- Median	21.00	28.50	34.00	39.00	42.00	46.50	47.00	56.00	50.50	49.50	$p < .001^*$
	- 95% CI for median	16.00-23.00	22.00-33.00	27.00-37.00	36.00-45.00	34.00-50.00	43.00-51.00	43.00-53.00	51.00-59.00	46.00-57.00	46.00-55.00	46.00-55.00

n number of patients, Min-Max minimum-maximum, Std. Dev. standard deviation, CI confidence interval, KS Kolmogorov-Smirnov, H:KW Kruskal-Wallis H, NS statistically not significant (p ≥ 0.05)

* Statistically significant (p < 0.05)

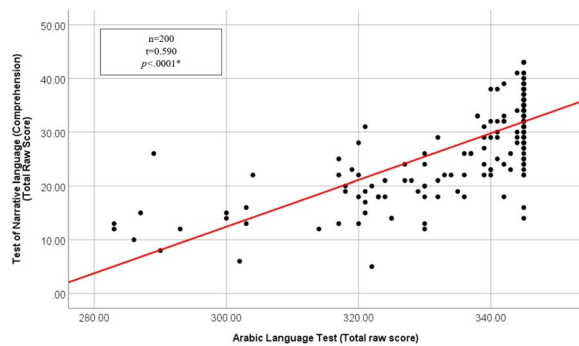


Fig. 1 Correlation between the TNL-2 total comprehension raw score and the Arabic language test total raw score. Scatter plot with best-fit (regression) showing strongly positive correlation between Test of Narrative Language comprehension total raw score and Arabic Language Test total raw score in 200 points of measurements

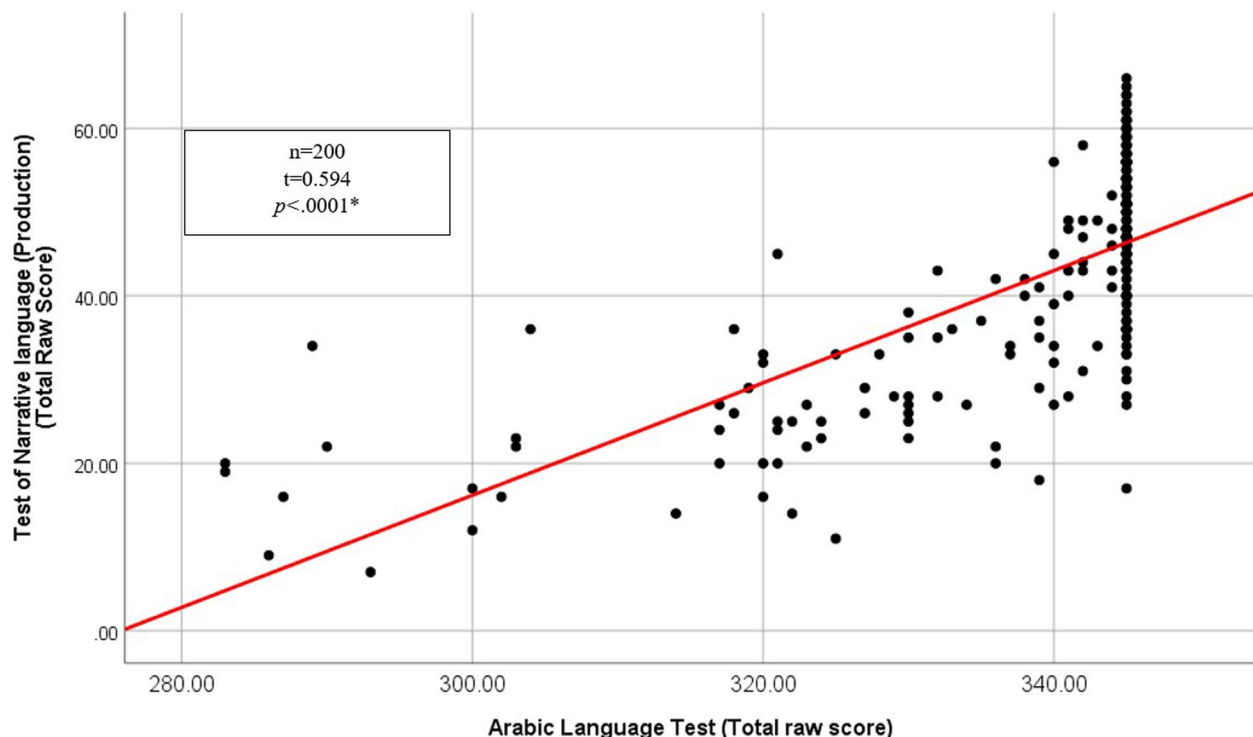


Fig. 2 Correlation between the TNL-2 total Production raw score and the Arabic language test total raw score. Scatter plot with best-fit (regression) showing strongly positive correlation between Test of Narrative Language production total raw score and Arabic Language Test total raw score in 200 points of measurements

can be utilized as an assessment tool for evaluating the progress of Arabic language skills, particularly in the area of narrative development. This study considered the lack of literature that examines narrative skills in Egyptian Arabic-speaking children.

Several studies have used the TNL to assess the narrative skills of children with delayed language development,

to evaluate the effectiveness of narrative interventions, and to correlate the performance of children in narrative tasks to different academic skills such as reading [28–31].

The TNL [26] has been translated, culturally adapted, and validated to other languages, such as Portuguese, to assess children's narrative skills. They concluded that the TNL can differentiate between different age groups regarding their narrative skills [50]. Additionally, an Arabic version was used to assess Arabic microstructure in Arabic-speaking children in Canada with respect to diglossia. The study concluded a significant relationship between microstructure and story grammar elements, with evidence of the role of the standard Arabic lexicon in predicting macrostructural elements [45].

The TNL-2 was specifically chosen in the current study for translation and adaptation into Egyptian Arabic for several reasons. First, the test assesses the main narra-

tive dimensions: macrostructure and microstructure. Both fundamental aspects are represented in the TNL-2 as the narrative production tasks are scored based on the story content and complexity. Children's narrative productions are scored based on semantic and morpho-syntactic elements, including conjunctions, temporal relations, correct grammar, story grammar elements, and

the production of complete narratives [27]. These aspects are reported in the literature as narrative language's most critical linguistic representatives [51].

Furthermore, TNL-2 assesses both comprehension and production [27]. Comprehension tasks include literal and inferential questions that tap into the children's cognitive and pragmatic abilities [52]. Additionally, we aimed to validate the test in order to use it to assess narrative skills for those with normal and disordered language skills later on. Some language disorders, such as specific language disorders, are known to have a discrepancy between receptive and expressive language skills [26]. Incorporating both comprehension and production tasks in the TNL-2 would be advantageous for capturing these distinctions and facilitating the diagnosis and monitoring of intervention programs.

Furthermore, the TNL-2 assesses narrative comprehension and production in a wide age range (4.0 through 15.11) with normative data for children aged 4 years through 15 years and 11 months [27]. The assessment of narrative skills encompasses various formats, including story retelling, picture sequencing, and the interpretation of a single picture as a script, personal narrative, or fictional narrative. This approach enables the examination of children's narrative abilities through a diverse range of tasks [27]. The test does not evaluate the spontaneous production of narratives, known as open narrative assessment methods [53].

In contrast to the structured methods, such as story retelling or using pictures to elicit narratives, in open methods, the child is required to produce a spontaneous account of a familiar situation, which requires memory skills, linguistic competence, and cognitive maturation. In that case, the examiner has no control over the narrative's subject, thereby posing challenges in standardizing assessment instruments and drawing comparisons among subjects [53, 54]. The TNL-2 serves as an assessment tool that fills the gap where the evaluation of narratives is concerned, and other previously used tools assessing narratives in Egyptian Arabic were lacking. One study by Safwat et al. targeted the assessment of narrative skills in preschool Arabic-speaking Egyptian children. The study limitations included assessment of preschool children only, and assessment of narrative comprehension was not included [41].

Another study by Khodeir et al. (2017) assessed various pragmatic aspects in children from the age of 4 years through 10 years, including narrative skills. Even though the assessment of a broader age range was included in addition to evaluating comprehension skills, the questions were mainly literal, targeting the main story elements without evaluating the ability to make inferences [42].

In the present study, comprehension of stories was assessed by listening to three stories, followed by comprehension questions the child was required to answer. The results showed a statistically significant increase in the comprehension scores of the three stories across the age groups. This finding can be attributed to the ability of children to develop story comprehension skills with age, in parallel with receptive language and cognitive skills [55]. Our finding is also supported by the strong correlation obtained by the Arabic language test scores, which constitute receptive and expressive components and the total raw comprehension score of the TNL-2. Comprehension questions for each story were a mixture of literal and inferential questions. The ability to make inferences continues to develop with pragmatic language development as the child matures, and the ability of children to make implications starts as early as 4 or 5 years old and continues to develop with age [56]. Earlier studies have reported that making inferences is an ability of late acquisition observed by the age of 8 years [57]. The development of these inferential skills enables children to answer more comprehension questions correctly. Therefore, increased comprehension scores in the current study were observed with increasing age.

A statistically significant increase in the scores of the production subtests was found across the age groups. Narrative production was assessed by retelling a story script while looking at the appropriate picture, producing a personal narrative based on a sequence of five pictures of familiar events, and producing a fictional narrative while looking at a picture. The scores for the production subtests were given based on the number of the correct elements produced in addition to the use of specific microstructural elements specified in the scoring sheet. The lower scores obtained in the younger age groups are explained by the younger children producing fewer story elements [58], with less use of temporal and causal relations [59], as observed in the sample included in the current study. Additionally, the use of correct morphology, complete episodes, and dialogue was noticed in the narratives of older children. The fact that these elements were scored in the test rendered the scores of children in the younger age groups less than the older ones. Subsequently, significant differences were found among the groups. This finding agrees with Safwat et al. (2013), who reported that the use of references increased with age. The study also reported an increase in the complexity of the sentences produced with age, such as the use of verb tense and different noun forms [41].

These findings were also supported in the current study by a strong positive correlation between the total scores of the Arabic language test and the total production raw scores of the TNL-2, showing that language

development and narrative language skills continue to develop with age. However, it should be taken into consideration that the ALT used in the current study is a screening tool for children from 2 to 8 years, which is the age range of language development save for the more complex linguistic pragmatic skills that continue to develop throughout adolescence [8]. This explains the ceiling effect found in our study in the results of the ALT around the age of 8 years, as most of the children obtained a total score around and above this age.

The study results agree with research on the development of narrative skills that show that children continue to develop their narrative skills during maturation and produce narratives containing the main macrostructural elements in the form of initial events, problems, consequences, emotions, attempts at solving the problem, and resolution, all the while incorporating the use of language, temporal relations, and causal relations to produce a coherent narrative [60, 61].

It is also noticed in the present study results that the medians of the scores of the children in the retelling task were higher than those of the personal narratives. The scores for fictional narratives were found to be the lowest, especially in the younger age groups. The later development of fictional narratives explains this finding, as cognitive development continues [62]. Additionally, story-retelling tasks are easier for children than personal narratives, especially for the younger age groups. Our findings agree with literature demonstrating that story retelling is easier than story generation, and personal events are recounted more readily than telling fictional stories [19].

The Arabic version of the TNL-2 was proved to be reliable in the current study by assessing the test-retest reliability of the test items, which showed a very high positive correlation between test-retest results of the total comprehension raw scores, total production raw scores, and the narrative language index. Additionally, inter-rater agreement was measured and showed an excellent degree of agreement between raters regarding the same previously mentioned items. Face validity was verified through the evaluation of the translated version of the TNL-2 by five expert phoniatricians. They were asked to review the test and complete a questionnaire assessing its effectiveness in measuring different narrative skills. Experts unanimously concur that the Arabic iteration of the TNL-2 possesses the capacity to evaluate the primary microstructural and macrostructural components of narratives, providing insight into the overall narrative abilities of Egyptian children who speak Arabic. Furthermore, the test's internal validity was confirmed by the strong positive correlation between the test and retest scores.

The Arabic-translated version of the TNL-2 utilized in the current study demonstrated its validity, reliability, and comprehensiveness as an assessment tool for evaluating various narrative skills, encompassing both the understanding and production of narratives. The TNL-2 can be utilized to evaluate the progression of narrative skills in children across various age cohorts, enabling the determination of their present narrative proficiency levels in relation to normative data.

The current study has a number of limitations. The Arabic version of the TNL-2 was not administered to atypically developing children in order to evaluate the test's capacity to differentiate between children with typical language development and those with delayed language development caused by various factors. This could serve as a guiding principle for future research. Caution should be exercised when interpreting the results because the sample used in the study was not normally distributed. This was due to the cross-sectional design of the study, which resulted in the presence of outliers and caused some age groups to deviate from a normal distribution. As a result, a statistical analysis was conducted using a test specifically designed for non-normally distributed populations. Additional data should be collected on a larger scale of cities and schools to obtain more substantial evidence for generalization, applicability, and test standardization.

Conclusion

The Arabic-translated version of the TNL-2 is a valid and reliable tool that can be used to assess the comprehension and production of narrative language skills in Egyptian Arabic-speaking children. Further application of the test on a larger sample of children is recommended. The Arabic version of the TNL-2 is suggested to be used to evaluate narrative skills in children with delayed language development and to assess the results of language intervention on narrative language.

Abbreviations

ADHD	Attention-deficit hyperactivity disorder
ICC	Intraclass correlation
IQ	Intelligence quotient
MSA	Modern standard Arabic
TNL	Test of Narrative Language
TNL-2	Test of Narrative Language-Second Edition

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43163-024-00603-7>.

Supplementary Material 1.

Acknowledgements

Not applicable.

Author's contributions

SMI contributed by designing the study, collecting and interpreting the data, drafting, and revising the manuscript. OAS and RMI contributed by writing and revising the manuscript. NHH contributed by writing the manuscript and data analysis and interpretation.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector.

Availability of data and materials

All data that support the findings of this study are included in this article. Further enquiries can be directed to the corresponding author.

Declarations**Ethical approval and consent to participate**

Permission to use the TNL-2 for the translation into Arabic, validation, and adaptation was obtained from the copyright holder, Pro-Ed, Inc. The Faculty of Medicine, Alexandria University ethics committee, approved the study protocol with IRB NO:00012098. Written informed consent was obtained from all caregivers after explaining the objectives of the study. The study was conducted at the Unit of Phoniatrics, department of Otorhinolaryngology, Alexandria Main University Hospital.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Received: 18 December 2023 Accepted: 8 March 2024

Published online: 05 April 2024

References

- Engel S (2012) *The stories children tell: making sense of the narratives of childhood*. Henry Holt and Company, New York
- Suggate S, Schaughency E, McAnally H, Reese E (2018) From infancy to adolescence: the longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension. *Cognit. Dev.* 47:82–95. <https://doi.org/10.1016/j.cogdev.2018.04.005>
- Kirby MS, Spencer TD, Chen Y-JI (2021) Oral narrative instruction improves kindergarten writing. *Read. Writ. Quar.* 37:574–591. https://doi.org/10.1044/2021_LSHSS-20-00162
- O'Neill DK, Pearce MJ, Pick JL (2004) Preschool children's narratives and performance on the Peabody Individualized Achievement Test–Revised: evidence of a relation between early narrative and later mathematical ability. *First Lang.* 24:149–183. <https://doi.org/10.1177/0142723704043529>
- Lervåg A, Hulme C, Melby-Lervåg M (2018) Unpicking the developmental relationship between oral language skills and reading comprehension: it's simple, but complex. *Child Dev.* 89:1821–1838. <https://doi.org/10.1111/cdev.12861>
- Colozzo P, Gillam RB, Wood M, Schnell RD, Johnston JR (2011) Content and form in the narratives of children with specific language impairment. *J. Speech Lang. Hear. Res.* 54:1609–1627. [https://doi.org/10.1044/1092-4388\(2011\)10-0247](https://doi.org/10.1044/1092-4388(2011)10-0247)
- Gillam S, Gillam R, Fargo J, Olszewski A, Segura H (2016) Monitoring indicators of scholarly language (MISL): a progress-monitoring tool for scoring narratives produced by school-age children with language impairments. *Commun. Dis. Quarterly.* 38:96–106. <https://doi.org/10.3389/feduc.2022.918127>
- Nippold MA (2017) Reading comprehension deficits in adolescents: addressing underlying language abilities. *Lang. Speech Hear. Serv. Sch.* 48:125–131. https://doi.org/10.1044/2016_LSHSS-16-0048
- Berman RA (2003) Genre and modality in developing discourse abilities. In: Moder CL, Martinovic-Zic A (eds) *Discourse across languages and cultures*. John Benjamins Publishing Company, Amsterdam. 329–356
- Nikolopoulos TP, Lloyd H, Starczewski H, Gallaway C (2003) Using SNAP Dragons to monitor narrative abilities in young deaf children following cochlear implantation. *Int. J. Pediatr. Otorhinolaryngol.* 67:535–541. [https://doi.org/10.1016/s0165-5876\(03\)00034-x](https://doi.org/10.1016/s0165-5876(03)00034-x)
- Westerveld MF, Gillon GT (2008) Oral narrative intervention for children with mixed reading disability. *Child Lang. Teach. Ther.* 24:31–54. <https://doi.org/10.1177/0265659007084567>
- Glisson L, Leitão S, Claessen M (2019) Evaluating the efficacy of a small-group oral narrative intervention programme for pre-primary children with narrative difficulties in a mainstream school setting. *Australian J. Learn. Difficult.* 241–20. <https://doi.org/10.1080/19404158.2019.1596138>
- Zamani P, Soleymani Z, Jalaie S, Zarandy MM (2018) The effects of narrative-based language intervention (NBLI) on spoken narrative structures in Persian-speaking cochlear implanted children: A prospective randomized control trial. *Int. J. Pediatr. Otorhinolaryngol.* 112:141–150. <https://doi.org/10.1016/j.ijporl.2018.06.048>
- Stein NL, Glenn CG (1979) An analysis of story comprehension in elementary school children. *New Direct. Dis. Proc.* 2:53–120
- Petersen DB (2011) A systematic review of narrative-based language intervention with children who have language impairment. *Commun. Disord. Quar.* 32:207–220. <https://doi.org/10.1177/1525740109353937>
- Spencer TD, Kajjan M, Petersen DB, Bilyk N (2013) Effects of an individualized narrative intervention on children's storytelling and comprehension skills. *J. Early Intervent.* 35:243–269. <https://doi.org/10.1177/1053815114540002>
- Petersen DB, Gillam SL, Gillam RB (2008) Emerging procedures in narrative assessment: the index of narrative complexity. *Topic. lang. Disord.* 28:115–130. <https://doi.org/10.1097/01.TLD.0000318933.46925.86>
- Merritt DD, Liles BZ (1989) Narrative analysis: clinical applications of story generation and story retelling. *J. Speech Hear. Disord.* 54:438–447. <https://doi.org/10.1044/jshd.5403.438>
- Spencer TD, Petersen DB (2020) Narrative intervention: principles to practice. *Lang. Speech Hear. Serv. Sch.* 51:1081–1096. https://doi.org/10.1044/2020_LSHSS-20-00015
- Hadley EB, Dickinson DK (2020) Measuring young children's word knowledge: a conceptual review. *J. Early Child. Liter.* 20:223–251. <https://doi.org/10.1177/1468798417753713>
- Westerveld MF, Gillon GT (2010) Oral narrative context effects on poor readers' spoken language performance: story retelling, story generation, and personal narratives. *Int. J. Speech Lang. Pathol.* 12:132–141. <https://doi.org/10.3109/17549500903414440>
- Hayward DV, Gillam RB, Lien P (2007) Retelling a script-based story: do children with and without language impairments focus on script and story elements? *Am. J. Speech Lang. Pathol.* 16:235–245. [https://doi.org/10.1044/1058-0360\(2007\)028](https://doi.org/10.1044/1058-0360(2007)028)
- R. B. Gillam, N. A. Pearson (2017) *Test of narrative language*. Pro-ed, Austin.
- Justice LM, Bowles R, Pence K, Gosse C (2010) A scalable tool for assessing children's language abilities within a narrative context: the NAP (Narrative Assessment Protocol). *Early Child. Res. Quarterly.* 25:218–234. <https://doi.org/10.1016/j.jecresq.2009.11.002>
- Spencer TD, Goldstein H (2019) *Assessment of Story Comprehension (ASCTM) Manual*. Brookes Publishing Company, Baltimore, Paul H
- R. B. Gillam, N. A. Pearson (2004). *TNL: Test of narrative language*. Pro-ed, Austin.
- R. B. Gillam, N. A. Pearson (2017). *Test of Narrative Language-second edition*. Pro-ed, Austin.
- Gillam SL, Olszewski A, Fargo J, Gillam RB (2014) Classroom-based narrative and vocabulary instruction: results of an early-stage, nonrandomized comparison study. *Lang. Speech Hear. Serv. Sch.* 45:204–219. https://doi.org/10.1044/2014_LSHSS-13-0008
- Loucks T, Chon H, Han W (2012) Audiovocal integration in adults who stutter. *Int. J. Lang. Commun. Disord.* 47:451–456. <https://doi.org/10.1111/j.1460-6984.2011.00111.x>

30. Catts HW, Compton D, Tomblin JB, Bridges MS (2012) Prevalence and nature of late-emerging poor readers. *J. Educ. Psychol.* 104:1–30. <https://doi.org/10.1037/a0025323>
31. Costa GM, Rossi NF, Giacheti CM (2012) Desempenho de falantes do português brasileiro no "Test of Narrative Language (TNL)." *Codas.* 30:e20170148. <https://doi.org/10.1590/2317-1782/20182017148>
32. Castilla-Earls A, Petersen D, Spencer T, Hammer K (2015) Narrative development in monolingual Spanish-speaking preschool children. *Early Educ. Dev.* 26:1166–1186. <https://doi.org/10.1080/10409289.2015.1027623>
33. Minami M (2002) Culture-specific language styles: the development of oral narrative and literacy. *Multilingual Matters*, Bristol
34. Sah W, Tornig P (2019) Storybook narratives in Mandarin-speaking pre adolescents and without autism spectrum disorder: internal state language and theory of mind abilities. *Taiwan J. Ling.* 17:67–89. [https://doi.org/10.6519/TJL.201907_17\(2\).0003](https://doi.org/10.6519/TJL.201907_17(2).0003)
35. G. Julian (2020) What are the most spoken languages in the world. Fluent in 3 months. <http://tony-silva.com/eslef/miscstudent/downloadpa geararticles/mostspokenlangs-fluentin3months.pdf>. Accessed 31 August 2022
36. Ferguson CA (1959) Diglossia. *Word.* 15:325–340. <https://doi.org/10.1080/00437956.1959.11659702>
37. Albirini A (2016) Modern Arabic sociolinguistics: diglossia, variation, codeswitching, attitudes and identity. Routledge, London
38. Leikin M, Ibrahim R, Eghbaria H (2014) The influence of diglossia in Arabic on narrative ability: evidence from analysis of the linguistic and narrative structure of discourse among pre-school children. *Read. Writ.* 27:733–747. <https://doi.org/10.1007/s11145-013-9462-3>
39. Kavar K, Walters J, Fine J (2019) Narrative production in Arabic-speaking adolescents with and without hearing loss. *J. Deaf Stud. Deaf Educ.* 24:255–269. <https://doi.org/10.1093/deafed/eny048>
40. Saiegh-Haddad E (2014) B. Spolsky (2014) Acquiring literacy in a diglossic context: problems and prospects. In: Saiegh-Haddad E, Joshi RM (eds) *Handbook of Arabic literacy: Insights and perspectives*. Springer, Heidelberg, pp 225–240
41. Safwat RF, H. M. EL-Dessouky, S. S. Shohdi, I. A. Hussien, (2013) Assessment of narrative skills in preschool children. *Egypt. J. Otolaryngol.* 29:130–135. <https://doi.org/10.7123/01.EJO.0000426472.83521.26>
42. Khodeir MS, Hegazi MA, Saleh MM (2017) Development and standardization of a test for pragmatic language skills in Egyptian Arabic: The Egyptian Arabic Pragmatic Language Test (EAPLT). *Folia Phoniatic. Logop.* 69:209–218. <https://doi.org/10.1159/000485656>
43. Ravid D, Naoum D (2014) & S. Nasser (2014) Narrative development in Arabic: story re-telling. In: Saiegh-Haddad E, Joshi RM (eds) *Handbook of Arabic literacy: Insights and perspectives*. Springer, Heidelberg, pp 153–170
44. Kavar K, Saiegh-Haddad E, Armon-Lotem S (2023) Text complexity and variety factors in narrative retelling and narrative comprehension among Arabic-speaking preschool children. *First Lang.* 43:355–379. <https://doi.org/10.1177/01427237221149800>
45. Asli-Badarneh A, Hipfner-Boucher K, Bumgardner XC, AlJanaideh R, Saiegh Haddad E (2023) Narrative microstructure and macrostructure skills in Arabic diglossia: the case of Arab immigrant children in Canada. *Int. J. Biling.* 27:349–373. <https://doi.org/10.1177/13670069221077306>
46. Albirini A (2019) Why Standard Arabic is not a second language for native speakers of Arabic. *Al-Arabiyya.* 52:49–72
47. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P (2005) Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health.* 8:94–104. <https://doi.org/10.1111/j.1524-4733.2005.04054.x>
48. Melika L (1998) The Stanford Binet Intelligence Scale. In: Melika L (ed) *Arabic Examiner's Handbook*. Dar El Maref Publishing, Cairo. 13–32
49. M. Kotby, A. Khairy, M. Barakah (1995) Language testing of Arabic speaking children. *Proceeding of the XVIII World Congress of the International Association of Logopedics and Phoniatic*
50. G. d. Santos, (2022) Validity evidence of the Test of Narrative Language (TNL) adapted to Brazilian Portuguese. *Rev. CEFAC.* 24:e6321. <https://doi.org/10.1590/1982-0216/20222416321>
51. Dockrell JE, Marshall CR (2015) Measurement issues: assessing language skills in young children. *Child Adolesc. Ment. Health.* 20:116–125. <https://doi.org/10.1111/camh.12072>
52. Vandewalle E, Boets B, Boons T, Ghesquière P, Zink I (2012) Oral language and narrative skills in children with specific language impairment with and without literacy delay: a three-year longitudinal study. *Res. Dev. Disabil.* 33:1857–1870. <https://doi.org/10.1016/j.ridd.2012.05.004>
53. Berman RA (1995) Narrative competence and storytelling performance: how children tell stories in different contexts. *J. Narrat. Life History.* 5:285–313. <https://doi.org/10.1075/jnlh.5.4.01nar>
54. Berman L, Slobin D (1994) Relating events in narrative: across linguistic developmental study. *Lawrence Erlbaum Associates Inc*, Hillsdale
55. Kintsch W (1998) *Comprehension: a paradigm for cognition*. Cambridge University Press, Cambridge
56. Berman RA (2008) The psycholinguistics of developing text construction. *J. Child. Lang.* 35:735–771. <https://doi.org/10.1017/S0305000908008787>
57. Wilson E, Katsos N (2022) Pragmatic, linguistic and cognitive factors in young children's development of quantity, relevance and word learning inferences. *J. Child Lang.* 49:1065–1092. <https://doi.org/10.1017/S0305000921000453>
58. Rakhlin NV, Li N, Aljughaiman A, Grigorenko EL (2020) Narrative language markers of Arabic language development and impairment. *J. Speech Lang. Hear. Res.* 63:3472–3487. https://doi.org/10.1044/2020_JSLHR-20-00082
59. Eisenberg Sarita L, Ukrainetz Teresa A, Hsu Jennifer R, Kaderavek Joan N, Justice Laura M, Gillam Ronald B (2008) Noun Phrase Elaboration in Children's Spoken Stories. *Lang. Speech Hear. Serv. Sch.* 39:145–157. [https://doi.org/10.1044/0161-1461\(2008/014\)](https://doi.org/10.1044/0161-1461(2008/014))
60. Muñoz ML, Gillam RB, Peña ED, Gullely-Faehnle A (2003) Measures of language development in fictional narratives of Latino children. *Lang. Speech Hear. Serv. Sch.* 34:332–342. [https://doi.org/10.1044/0161-1461\(2003/027\)](https://doi.org/10.1044/0161-1461(2003/027))
61. Peterson C, McCabe A (2013) *Developmental psycholinguistics: Three ways of looking at a child's narrative*. Springer Science & Business Media, London
62. Roth FP, Spekman NJ (1986) Narrative discourse: spontaneously generated stories of learning-disabled and normally achieving students. *J. Speech Hear. Disord.* 51:8–23. <https://doi.org/10.1044/jshd.5101.08>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.