

Developing a Narrative Test and Determining its Psychometric Properties for Persian-Speaking Children Aged 4 to 5 Years

Fahime Ardeshiri, Reyhane Mohamadi, Elnaz Chitgar, Arezoo Saffarian*

Department of Speech Therapy, Rehabilitation Research Center, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran.

***Corresponding Author:** Arezoo Saffarian, Ph.D

Department of Speech Therapy, Rehabilitation Research Center, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran.

Email: asaffarian.slp@gmail.com

Tel: 02122222059

ORCID ID:

Fahime Ardeshiri: 0000-0001-8920-3637

Reyhane Mohamadi: 0000-0003-2823-5197

Elnaz Chitgar: 0000-0002-3220-2358

Arezoo Saffarian: 0000-0003-1749-2446

Article Info:

Received: 29 Apr 2025

Accepted: 22 Jul 2025

Citation: Ardeshiri E, Mohamadi R, Chitgar E, Saffarian A, Mohamed YE. Developing a Narrative Test and Determining its Psychometric Properties for Persian-Speaking Children Aged 4 to 5 Years. *Journal of Modern Rehabilitation*. 2026;20(1):?-?

Running title: Narrative test and determining its psychometric properties

Abstract

Objectives: Between ages 2-7, children develop the ability to narrate stories with improved detail and organization, including main characters, events, responses, efforts, and settings. Narrative assessment, a tool for evaluating language development, allows speech pathologists to examine syntax, vocabulary, and complexity. This study aimed to develop a tool for evaluating storytelling skills in children aged 4-5.

Methods: This study was conducted in two stages. In the first stage, the story was created, images were designed, and face validity was assessed by speech therapists. In the second stage, a descriptive-analytical study was carried out in kindergartens in Tehran in 2023. Children were asked to narrate stories based on the provided images, and their narratives were scored using story grammar analysis. Psychometric evaluations included construct validity (assessed via paired t-test), inter-rater reliability (measured using Cohen's Kappa), test-retest reliability (analyzed with Pearson correlation), and internal consistency (evaluated using Cronbach's alpha).

Results: The speech therapists evaluated and confirmed the face validity of the story and images, implementing necessary adjustments.. Construct validity was significant ($P = 0.000$), indicating alignment with typical development. Inter-evaluator reliability ($k=0.712$) and test-retest reliability

($r=0.591$) were moderate to good. Internal consistency varied, with $\alpha > 0.5$ in one story and $\alpha < 0.5$ in others.

Discussion: The findings indicate that the test demonstrates validity and moderate reliability, suggesting its suitability for practical applications.

Keywords: Narration -Preschool Age - Psychometric properties-Persian

Highlights

- Storytelling is one of the language skills that develops in children from 2 to 7 years of age.
- Storytelling is one of the predictors of the acquisition of literacy skills.
- Assessing storytelling skills can help diagnose and treat learning disorder early.
- The narrative test is valid and moderately reliable and these properties making it suitable for practical use.

Plain language summary

This research was conducted to develop a tool for assessing storytelling skills in 4-5-year-old Persian-speaking children. After conducting psychometrics involve face validity, construct validity, and inter-rater and test-retest reliability, the efficiency of the tool was shown.

Introduction

Storytelling is a vital skill in children's daily communication from an early age, allowing them to describe real or imaginary events with causal and sequential relationships [1-3]. This skill contributes to language development and predicts future skills in reading, vocabulary, and phonological awareness [4,5]. Storytelling is structured on two levels: microstructure (e.g., word diversity, sentence complexity) and macrostructure (e.g., plot, characters, and events) [6].

Children's narrative development evolves through distinct stages. At ages 2-3, children tell "heap stories," consisting of simple, unconnected descriptions. By 3-4 years, they progress to "sequential stories," where events relate to a main theme but lack plot or causal links. By 4-4.5 years, "primitive stories" emerge, containing a beginning, an action, and an ending. In ages 4.5-5, "chain stories" add basic causal and temporal connections, with four story elements: initiating event, motivation, action, and consequence. True narrative structure typically appears around ages 5-7, with clear plots and character motives, although story complexity continues to grow [7].

Research indicates that, alongside language development, cultural and environmental factors also shape storytelling abilities. For example, Asian children, who often rely more on nonverbal cues due to cultural norms, may tell simpler stories compared to their European counterparts. Similarly, children with less language exposure tend to have weaker narrative skills [8-13].

Storytelling assessments provide a comprehensive view of children's language skills by testing various aspects like grammar, vocabulary, and pragmatics, which other methods may overlook [1,16]. These assessments are crucial for identifying language delays and understanding the link between narrative development and broader language skills [17-19].

Standardized storytelling assessments typically use three methods: story production, story retelling, and personal narrative. Examples include Mayer's *Frog, Where Are You?* (1969), the Renfrew Bus Story Test, and the Profiling Oral Narrative Ability (PONA) test, all of which evaluate narrative elements such as microstructure and macrostructure in young children [20-22]. The Multilingual Assessment Instrument for Narratives (MAIN) evaluates bilingual children's storytelling skills across languages [23]. In the Persian, Soleimani et al. (2016) developed a protocol for children aged 5-6, demonstrating strong content validity and significant test-retest

reliability [24]. Similarly, Jafari et al. (2012) created a test for rapid assessment of 6-7-year-old Persian-speaking children's narrative abilities, with high validity and reliability [25]. However, a validated storytelling test specifically for 4-5-year-old Persian-speaking children does not yet exist. This assessment tool would provide speech-language pathologists with an effective instrument for evaluating narrative skills in young children, enabling targeted intervention strategies and potentially mitigating future academic challenges.. This research aims to develop a storytelling evaluation test tailored to Persian-speaking children aged 4-5 years, examining its psychometric properties to ensure it supports accurate evaluation and intervention strategies for early language development.

Method

Participants: This study included typically developing Persian-speaking children in two age groups (4-4.5 years and 4.5-5 years) without language disorders. Participants were selected across four stages:

- **Preliminary Stage:** Stories validated by experts were tested on 10 Persian-speaking children (ages 4-5) from two Tehran preschools, who were asked to retell the stories.
- **Second Phase:** Thirty children from seven preschools in Tehran were selected using random sampling to retell stories with standard images.
- **Third Phase:** A sample of 100 children from various Tehran districts participated in the final test.
- **Fourth Stage:** Among these, 20 children were retested to assess test-retest reliability.

Children were included based on fluency in Persian, scoring above the threshold on the Age and Stage Screening Questionnaire (ASQ), and exclusion of any neurological, developmental, or psychiatric disorders. Exclusions were based on unacceptable ASQ scores, behavioral issues, or lack of cooperation.

Tools:

- **Content Validity Form:** Developed based on Lawshe's method and reviewed by speech-language pathology experts to confirm the relevance of each test item.
- **Personal Information Questionnaire:** A researcher-created 42-item questionnaire on the child's background, family, medical history, and communication.
- **ASQ:** A parent-report screening form for ages 4-60 months, assessing development in communication, motor skills, social, and problem-solving areas, with high reliability ($\alpha = 0.79$).
- **Storytelling Test:** Eight validated stories, tailored to each age group, were assessed for content, face, and construct validity.
- **Recording Devices:** Participants responses were recorded by a Samsung galaxy S21 smartphone and ASUS laptop for subsequent analysis.

Methodology Overview:

This study develops a storytelling test to evaluate narrative skills in young children, structured into four main stages.

Stage 1: Story Selection and Preparation

The researcher, after consultation with children's storytelling experts and referencing literature on narrative development, selected animal characters for the story protagonists to align with the cognitive level of 4-5-year-olds. Based on Paul's guidelines for narrative structure, initial story

drafts were developed focusing on micro and macro structures, including plot elements like setting, sequence, and causality. The draft stories were reviewed by a speech therapy expert for adjustments. Ten draft stories, alongside two questionnaires, were sent to a panel of experts for feedback on necessity and relevance. Following their evaluation, two stories were eliminated, and minor modifications were applied to others. The remaining eight stories were then tested on a small group of children (ten 4-5-year-olds) to verify comprehension and suitability. Upon successful testing, each story was illustrated with three colorful images, aiming for engaging visuals that matched each story's progression.

Stage 2: Content and Face Validity Assessment

To establish content validity, a panel of experts (five speech-language pathologists professors, two experienced therapists, and a linguist) evaluated the age-appropriateness of each narrative for the target population.. This process involved rating each story element as essential, useful, or non-essential (Table 1). To establish face validity, the same panel reviewed whether the illustrations supported the storyline in an age-appropriate way and ensured clarity. Questions posed included the appropriateness of the images, sequence alignment with the story, and clarity for aiding children in narrative retelling (Table 2,3). Experts were also surveyed about the priority and necessity of using stories in the test, the results of which are shown in Table 4. Following panel revisions, the researcher further assessed the final draft with a small group of 10 children divided into two age categories (4-4.5 and 4.5-5 years), using children's responses to finalize the stories.

Table 1: Content Validity Ratio

Sentence number	Stories									
	Mr. Mouse's shopping	The farmer's farm	Playing with the ball	Helping	Birthday gift	Neighbor	Kindergarten	Chicks	New friend	Visiting grandparents
Sentence 1	Unnecessary	Unnecessary	Unnecessary	Unnecessary	Unnecessary	Unnecessary	Unnecessary	Unnecessary	Unnecessary	Unnecessary
Sentence 2	1	0.75	0.75	1	1	1	0.75	0.75	1	0.75
Sentence 3	1	1	0.75	1	1	1	1	1	1	1
Sentence 4	1	1	0.75	1	-	1	1	1	1	1

Sentence°	1	-	-	-	-	1	-	-	1	1
------------------	---	---	---	---	---	---	---	---	---	---

Table 2: Questions related to story elements level 1

Content validity questions about story elements	Story 1:(Mr. mouse's shopping)	Story 2: (the farmer's farm)	Story 3: play with the ball	Story4: (helping)	Story 5: (birthday gift)
Does the story have cohesion?	Sentences should be connected with appropriate conjunctions or related words	Moderate		Yes	Moderate
Does the stories have suitable introduction, continuation and conclusion?	Introduction: In a big forest, there live the little mouse	Yes	Introduction: a shy bear lived in the forest with its mother	Yes	Yes
Does the story have a main event?	The event is incomplete: what should it buy	Hunger and the need for food should be addressed	Yes, but it needs revision	Yes	The challenge is not well expressed
Does the story have an action following the main event?	Yes	Yes	Yes	Yes	Yes
Are there at least basic story elements (consequence, characters, events, etc.) present?	No	No, the events should be complete	moderate	moderate	moderate

Is the vocabulary diversity in the story appropriate?	Yes	Yes	Yes	Kind deer	Low
Is the syntactic complexity in the story appropriate?	Yes	Yes	Yes	Yes	Low
Are the character in the story appropriately chosen?	Yes	Yes	No, a character should be added	Yes	Yes
Is the sequence of the story appropriate?	Yes	Yes	Yes	Yes	Yes

Table 3: Questions related to story elements level 2

Content validity questions about story elements	Story 6: neighbor	Story 7: (kindergarten)	Story 8: chicks	Story9: (new friend)	Story 10: (visiting grandparents)
Does the story have cohesion?	Need revision	Need revision	You can rearrange the sentences in the last section and include a conclusion	Moderate	Yes
Does the stories have suitable introduction, continuation and conclusion?	Yes	Yes	Yes	The casual relationships are not well expressed	No
Does the story have a main event?	Yes	Yes	Yes	The challenge is	The challenge is not good; it

				Not very good	Doesn't need a solution, etc.
Does the story have an action following the main event?	Yes	Yes	Yes	No	Yes
Are there at least basic story elements (consequence, characters, events, etc.) present?	Yes	Yes	Yes	Yes	Yes
Is the vocabulary diversity in the story appropriate?	Yes	Yes, it could be improved	Yes	Yes	Yes
Is the syntactic complexity in the story appropriate?	Yes	Yes	Yes	Yes	Yes
Are the character in the story appropriately chosen?	Yes	Yes	Yes	Yes	I think it would be better if there were no snakes, as animals are used in all the stories; it could also be about humans.
Is the sequence of the story appropriate?	Yes	Yes	It would be better if the sentences were rearranged	No	No

Table 4: Priority of each story based on Content Validity Ratio

	Stories	
		Priority or necessity of using stories in storytelling test

Level 1	Story 1 (Mr. Mouse's shopping)	1
	Story 2 (the farmer's farm)	1
	Story 3 (playing with the ball)	1
	Story4 (helping)	1
	Story 5 (birthday gift)	0.50
Level 2	Story6 (neighbor)	1
	Story 7 (kindergarten)	0.50
	Story 8 (chicks)	1
	Story 9(new friend)	1
	Story 10 (visiting grandparents)	0.25

Stage 3. Test Implementation and Data Collection

After content and face validity adjustments, the storytelling test was administered to 100 preschoolers, selected via random sampling from various kindergartens. Parental consent was secured for each participant. The storytelling sessions were administered individually in quiet kindergarten classrooms. Each child listened to a narrative and was subsequently asked to retell it. This procedure, which included audio recording of the narrations, was designed to comprehensively assess children's storytelling abilities by capturing nuanced aspects of their narrative production. Children's retellings were scored on 10 key story elements. A subsample of 25 children was asked to narrate all eight stories, with the maximum possible score of 40 points. The stories were then transcribed and analyzed according to story grammar, including elements such as the initiating event, main characters, response, plan, and consequences. The presence of each structural element was scored with either a "1" or "0" based on the child's narrative output.

Stage 4. Validity and Reliability Testing

The final test aimed to determine both construct validity and reliability. Construct validity assessed whether the test effectively differentiated between the two age groups. Given that storytelling skills improve with age, children in the older group (4.5-5 years) were expected to demonstrate greater narrative complexity. This was tested by comparing average scores between the age groups. Scores were assigned based on the presence of narrative elements, with age-related growth confirming construct validity.

To measure reliability, both test-retest reliability and inter-rater reliability were employed. Test-retest reliability was calculated by re-administering the test to 20 randomly selected children after 20 days, with Pearson's correlation applied to evaluate score consistency. Inter-rater reliability was verified by having a second examiner independently score 10 children from each age group. Kappa statistics assessed scoring agreement between examiners, while Cronbach's alpha coefficient determined internal consistency across test items.

Data Analysis

Data analysis was performed using SPSS 16, with descriptive and inferential statistical techniques applied. Content validity ratios and indexes were calculated per Lawshe's model, while t-tests measured construct validity by analyzing total test scores in correlation with age. Pearson's correlation coefficient evaluated the consistency of the test over time, and Cronbach's alpha was used to confirm internal consistency.

Result

The study involved two groups of 50 typically developing children. The first group (30 girls, 20 boys) averaged 50.48 months old with a 2.032 standard deviation, while the second group (26 girls, 24 boys) averaged 57.32 months with a 1.7485 standard deviation. Content validity was assessed by expert review of story elements. Experts identified the introductory phrase, "Once upon a time, there was no one but the merciful God," as unnecessary in all stories, along with select phrases in stories 2, 3, 7, and 8. All other statements were rated necessary by all experts, affirming their inclusion per the Content Validity Ratio (CVR) formula.

In the face validity phase, experts suggested changes for better story alignment. For "Mr. Mouse in the Clouds of Thought," they advised adding a store image in the first illustration and showing cheese in the third, with a depiction of lack of money for clarity. In "Rabbit Farm," they recommended a full fence around the farm so the rabbit couldn't access the lettuce. For construct validity, analysis showed that children aged 4-4.5 scored an average of 31.40 on first-level stories and 16.80 on second-level stories. The 4.5-5 age group scored higher, with 38.84 and 34.04, indicating age-appropriate story levels (Table 5).

Table 5: Construct validity results using independent t-test

	Levels	Mean	Count	P-value
Level 1	Group of children 4 to 4.5 years	31.40	25	0.000
	Group of children 4.5 to 5 years	38.84		
Level 2	Group of children 4 to 4.5 years	16.80	25	0.000
	Group of children 4.5 to 5 years	34.04		

The research aimed to assess the internal consistency of story structures (initial event, middle response, plan, attempt, consequences, reaction, main characters, hero, context). Each element scored one if present, zero if absent. Children's storytelling data was analyzed using SPSS, with Cronbach's alpha reported for each story. Internal consistency of elements is shown in Table 6. The relatively low Cronbach's alpha observed in this tool can be explained by developmental differences in children's narrative abilities. Research shows that children under age 7 typically do not consistently produce or recognize all narrative components such as setting, characters and temporal structure. Since narrative skills emerge gradually during early childhood, the variability

in children's responses reflects normal developmental patterns rather than a flaw in the instrument itself [33].

Table 6: Results of internal consistency using Cronbach's alpha coefficient

	Story 1	Story 2	Story 3	Story 4	Story 5	Story 6	Story 7	Story 8
Cronbach's alpha	0.346	0.329	0.222	0.327	0.326	0.283	0.534	0.158

Test-Retest Reliability:

The Pearson correlation for test reliability showed an average correlation ($r=0.425$) in the 4-4.5-year age group across two sessions with a 20-day gap, while the 4.5-5-year group had a strong correlation ($r=0.719$). The overall correlation between sessions was 0.591, indicating average consistency.

Inter-rater Reliability:

The Kappa coefficient for the 4-4.5-year group was 0.639, indicating good reliability, while the 4.5-5-year group had a Kappa of 0.744, also good. An overall Kappa of 0.712 confirms good inter-rater reliability. The paired t-test ($p=0.000$) showed a significant difference between the groups.

Descriptive Statistics:

In this study examined the average percentage of narrative elements present in children's storytelling, categorized by age. Data were collected from two groups of 50 children (each retelling age-appropriate stories) and two groups of 25 children (each retelling 8 stories). The percentage of children expressing each narrative element in each story was calculated, and the mean was derived based on the number of stories (4 or 8).

The results showed that all children in group one included the narrative element of "attempt" in their stories. Following that, the elements with the highest frequencies were, respectively: plan, main character, initial response, hero, consequence, and initiating event. In this group, the narrative element with the lowest frequency was "reaction".

In group two, all children included the narrative element of "hero." This was followed by high frequencies for initiating event, initial response, main character, plan, attempt, and consequence. The least frequently expressed element was "reaction" although it was significantly more frequent than in group one. Furthermore, when the groups were asked to retell stories from both levels, the results showed stronger performance across all narrative elements – except for the "setting" element, where the difference was minimal.

Table7: The percentage of each narrative element expressed by the children

	Level 1 stories (n=50)	level 2 stories (n=50)	Age group 4 to 4.5 years (both stories)	Age group 4.5 to 5 years (both stories)
Initial event	% 90/50	% 99	% 90	% 100
Internal response	% 96	% 98	% 99/50	% 100

plan	%٩٨	%٩٧	%٤٩	%١٠٠
Attempt	%١٠٠	%٩٥	%٤٩	%٩٨
Consequence	%٩٣	%٩١/٥	%٥٠/٥	%٨٢
Reaction	%٤	%٥١/٥	%٢/٥	%٨٦
Main characters	%٩٧	%٩٨	%٩١/٢٥	%١٠٠
Hero	%٩٥	%١٠٠	%٧٩	%٩٨
Context of time	%١٩	%٥٩	%١٢/٥	%٦٠
Context of place	%٧٣	%٥٣/٥	%٦٠/٢٥	%٥٢

Discussion

We designed and validated an assessment tool to evaluate storytelling ability through picture retelling in children aged 4-5 years, divided into two age groups (4.0-4.5 and 4.5-5.0 years). The instrument comprises eight stories, each containing three sequential images. The face validity of the images and story text was established through consensus among speech-language pathologists specializing in this field and age group. After the investigations, the construct validity was confirmed ($P=0.000$), that is, the stories are aligned with the natural development of children. Inter-rater reliability ($k=0.712$) and test-retest reliability ($r=0.591$) were both good, indicating reliable performance. The Internal consistency of one of the stories was acceptable with $\alpha < 0.5$, while for the remaining stories $\alpha > 0.5$. Since the theoretical basis of the current research is the development of children's storytelling skills with age [11], to determine the validity of the construct, it is necessary to compare the two groups involved in the research. The test of each story was scored based on 10 narrative elements, and 25 children from each group had to tell all 8 stories. The maximum score that children could get in each level was 40.

An analysis of the children's performance scores revealed distinct patterns across age groups and story levels. In the 4.0-4.5-year age group, Level 1 narratives yielded a maximum score of 36 (achieved by one participant) and a minimum score of 27. For Level 2 stories within this age group, scores ranged from 12 to 24, with two children attaining the maximum score. The older cohort (4.5-5.0 years) demonstrated higher performance on Level 2 stories, with scores ranging from 31 to 37.. In level 1 stories, the lowest score was 31 and the highest score was 40, with the majority of children scoring between 38 and 40. Furthermore, a greater number of children in the older age group (4.5-5.0 years) scored above 30 compared to their younger counterparts (4.0-4.5 years).. As previous studies have confirmed, children under 5 years of age acquire various language elements, but the full and accurate use of these language elements to fully express narrative elements is not yet fully developed in children under 5 years of age [31]. In other words, despite the complete structure of micro-structure in 4-year-old children, there is a significant difference between the storytelling skills and the use of macro-structure in 4-year-old and 5-year-old children [16]. One-year-old children as studies on the Storytelling Test of 6- to 7-year-old children show, children can create more complex stories as they get older, and they can also use more story elements to create stories, including description. Emotional and mental states and causal relationships, which are not seen in this age group[2]. In addition to the current study, another test construction study examining syntactic comprehension in Persian-speaking children aged 4 to 10 years reports significant differences in both the types and complexity of syntactic structures across age groups. The findings demonstrate that syntactic structures become progressively more complex and diverse with increasing age [32]. Overall, the results of the reviewed studies are consistent with this study and show that with increasing age, complexity is seen in story elements and structures. The statistical analysis of the independent t-test in this study also showed a significant difference

between the two groups, similar to previous studies. This result means that this research has a theoretical basis that story-telling skills increase with age, and as the age of 5 approaches, the content of the story and the overall structure of the story become more complete[11]. According to the development process of story writing from the point of view of Appleby, who states that children in the age range of 4 to 4.5 years have stories with a main character and three-story grammars including the beginning of the event, attempt or action, and consequences. In the results section, children in this age group were able to tell stories in such a way that these elements had high scores[7].

In order to determine Inter-rater reliability, 10 children from each age group were randomly selected and their verbal output was evaluated by another examiner. Of the verbal output of 10 children in group 1, who were 4 to 5.4 years old, the verbal output of 5 children, that is, half of the children that were reanalyzed, had similar results in all 4 stories. Among the other 5 children's verbal output, two children differed in the scores of one story, the total score of one was one score higher and the total score of the other was one score lower. Also, one person had a difference in the scores of 2 stories, that the score of one story was higher and the score of the other story was lower, so the total score was not different from the first order. In the first group, this statistic was 0.6, which indicates good performance, that is, inter-rater reliability is in the good range.

On the other hand, In the second group where the children were 4.5 to 5 years old, among the 10 children whose verbal output was evaluated by the second examiner, the verbal output of 8 children, i.e. more than half of the children, had similar results to the first analysis. Two other children also had a difference of one point in a story compared to the previous time. As expected from the scores, the kappa statistic in the second group will be higher than the first group. The Kappa statistic in this group was 0.7, which means that the inter-rater reliability is in a good range.

If we take a look at the Kappa statistic that was used to determine the Inter-rater reliability of the storytelling test by evaluating 20 children, we see a number of 0.7. Good performance is also considered in the analysis of this number. Various factors may influence this result. The first factor is that since the raters of the two groups were exchanged in the second assessment, they may have applied the mentality of the first group to the second group. This suggests the potential for rater bias, wherein an evaluator who first assessed the older age group might subsequently expect higher performance from the younger group. Such expectations could lead to stricter scoring criteria being applied to the younger children, resulting in artificially lower scores compared to initial assessments..

The results of these statistical analyzes also show that this tool is consistent with the theories related to the natural development of children and follows the natural development of children's storytelling skills. According to the appropriate result of construct validity, we find that this instrument can assess what the project manager intended and distinguish between the skills of older children (4.5 to 5 years) and younger children (4 to 4.5 years).

The good Inter-rater reliability and test-retest reliability, which had moderate results, suggest that the tool can be valid and provide consistent results for its users, i.e., researchers conducting research designs or clinicians in clinical settings, in the future.

It can be stated that children remember the stories after 20 days and with the details they already had in mind and also hearing the story for the second time, they can have better and more complete outputs. As a result, they may have different results compared to the first evaluation, which leads to a decrease in the correlation coefficient between the results. In the present study, only one story had an acceptable internal consistency. This can also have various reasons. For example, the number of children participating in the sampling of each story was limited. We know that the larger

the sample size, the greater the internal consistency. On the other hand, a group of children who were included in the sample were in the kindergarten environment and were separated from their position and entered a separate room. This could create stress or not focus on the story based on the character and moral pattern of the children. As a result, children did not pay attention to the details of the stories. For example, children often had difficulty recounting attempt, consequences, and contexts of time and place that require more attention and care. Another factor that accounted for the difference in children's outcomes was the availability of a rich environment for the development of children's storytelling. Children who were more familiar with books and reading, paid more attention to the details of the stories, and as a result, they recounted those details more.

The most important factor affecting children's grades and, consequently, the internal consistency of stories, is the natural development of children's storytelling. As mentioned, 4-4.5-year-old children describe only 3 elements of a story, including the initial event, attempt, and its consequences. 4.5 to 5-year-old children also express 4 elements of the story, including the initial event, plot, attempt, and its consequences. This is why the correlation of elements in these stories was often unacceptable[28].

Conclusion

According to the measures taken to determine the face and content validity of the storytelling test in this research and with the help of experts, the images and texts of the stories were adapted. The result of this research was the preparation of 4 stories in the age group of 4 to 4.5 years and 4 stories in the age group of 4.5 to 5 years, whose texts and images were suitable for children in this age range. In addition, the results of the statistical analysis also tell us that this tool follows the principle that story telling skills increase with age, and as children approach the age of 5, the content of the story and the overall structure of the story. Consistent with established developmental trajectories in children's narrative abilities, this instrument demonstrates appropriate construct validity. The findings indicate that the assessment tool successfully measures its intended constructs and effectively discriminates between the storytelling skills of older (4.5-5.0 years) and younger (4.0-4.5 years) age groups. In addition, good interrater reliability and test-retest reliability show that the test is a suitable tool and its use is practical.

Ethical consideration

Compliance with ethical guidelines

This study was approved by the Ethics Committee of Iran University of Medical Sciences (Code: IR.IUMS.REC.1401.472).

Funding

The study was funded by the Iran University of Medical Sciences (Grant No.: 1402-1-6-25155).

Authors' contributions

Conceptualization and study design: Reyhane Mohamadi, Fahime Ardeshtiri; Data acquisition: Fahime Ardeshtiri; Data analysis and interpretation: Arezoo Saffarian, Fahime Ardeshtiri; Statistical analysis: Fahime Ardeshtiri; Initial draft preparation: Elnaz Chitgar; Review, editing and final approval: Reyhane Mohamadi, Arezoo Saffarian, Fahime Ardeshtiri.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgments

The authors thank the officials of schools in regions 1, 2, 3, 5, 6, 7, 9, 10, 12.

References

1. Nakhshab, M., M. Arabpour, and Y.J.J.o.R.i.R.S. Kazemi, *Narrative assessment methods in children: a systematic review*. 2015. **10**(5): p. 704-719.
2. Beytollahi, S., Z. Soleymani, and S.J.I.J.o.M.S. Jalaie, *The development of a new test for consecutive assessment of narrative skills in Iranian school-age children*. 2020. **45**(6): p. 425.
3. Isbell, R., et al., *The effects of storytelling and story reading on the oral language complexity and story comprehension of young children*. 2004. **32**: p. 157-163.
4. Wellman, R.L., *Narrative Ability of Children with Speech and Language Deficits and its Potential to Predict Later Literacy Skills*. 2009, Case Western Reserve University.
5. Safwat, R.F., et al., *Assessment of narrative skills in preschool children*. 2013. **29**: p. 130-135.
6. Justice, L.M., et al., *A scalable tool for assessing children's language abilities within a narrative context: The NAP (Narrative Assessment Protocol)*. 2010. **25**(2): p. 218-234.
7. Applebee, A.N., *The child's concept of story: Ages two to seventeen*. 1978.
8. Silva, M., K. Strasser, and K.J.E.C.R.Q. Cain, *Early narrative skills in Chilean preschool: Questions scaffold the production of coherent narratives*. 2014. **29**(2): p. 205-213.
9. Reese, E., et al., *Maternal elaborative reminiscing increases low-income children's narrative skills relative to dialogic reading*. 2010. **21**(3): p. 318-342.
10. Hutchison, D., *Gender differences in the macrostructural narrative language of second-grade African American children*. 2012, Thesis, Ohio State University. Retrieved from <http://kb.osu.edu/dspace>
11. Melzi, G., A. Schick, and E.J.A.e.p. Bostwick, *Latino children's narrative competencies over the preschool years*. 2013. **27**(115): p. 1-14.
12. Nielsen, D.C., L.D. Friesen, and J.J.J.o.E. Fink, *The effectiveness of a model of language-focused classroom instruction on the vocabulary and narrative development of kindergarten children*. 2012. **192**(2-3): p. 63-77.
13. Vernon-Feagans, L., et al., *Early language and literacy skills in low-income African American and Hispanic children*. 2001. **1**: p. 192-210.
14. Bowles, R.P., et al., *Development of the Narrative Assessment Protocol-2: A tool for examining young children's narrative skill*. 2020. **51**(2): p. 390-404.
15. Reilly, J., et al., *"Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome*. 2004. **88**(2): p. 229-247.
16. Hipfner-Boucher, K., et al., *Relationships between preschoolers' oral language and phonological awareness*. 2014. **34**(2): p. 178-197.
17. Norbury, C.F., D.V.J.I.j.o.l. Bishop, and c. disorders, *Narrative skills of children with communication impairments*. 2003. **38**(3): p. 287-313.
18. Pankratz, M.E., et al., *The diagnostic and predictive validity of the Renfrew Bus Story*. 2007.
19. Sebina, B., J. Setter, and M.J.I.J.o.B. Daller, *The Setswana speech rhythm of 6-7 year-old Setswana-English bilingual children*. 2021. **25**(3): p. 592-605.
20. Mercer, M., *Frog, where are you*. 1969, New York: Dial Books for Young Readers.
21. Renfrew, C., *At-A-Glance Test Review: The Renfrew Bus Story (RBS)*.

22. Westerveld, M.F. and G.T.J.I.J.o.S.-L.P. Gillon, *Oral narrative context effects on poor readers' spoken language performance: Story retelling, story generation, and personal narratives*. 2010. **12**(2): p. 132-141.
23. Gagarina, N.j.V., et al., *MAIN: Multilingual assessment instrument for narratives*. 2012. **56**: p. 155-155.
24. Soleymani, Z., et al., *Language sample analysis: development of a valid language assessment tool and determining the reliability of outcome measures for Farsi-speaking children*. 2016. **13**(2): p. 275-291.
25. Jafari, S., et al., *Developing a story retelling test for the assessment of language structure in Persian-speaking children*. 2012. **21**(3): p. 51-61.
26. Haynes, S.N., G.T. Smith, and J.D. Hunsley, *Scientific foundations of clinical assessment*. 2011: Routledge.
27. Bornstein, R.F., *Face validity in psychological assessment: Implications for a unified model of validity*. 1996.
28. Paul, R. and C. Norbury, *Language Disorders from Infancy Through Adolescence-E-Book: Language Disorders from Infancy Through Adolescence-E-Book*. 2012: Elsevier Health Sciences.
29. Karami, A.J.T.R.P., *Introduction to test construction and psychological tests*. 2008.
30. Kline, P., *Handbook of psychological testing*. 2013: Routledge.
31. Owens Jr, R.E., *Language development: An introduction*. 2001: Allyn & Bacon.
32. Mohamadi, R., et al., *Development of the Persian syntax comprehension test*. 2019. **124**: p. 22-29.
33. Demir, O. E., Levine, S. C., & Goldin-Meadow, S. (2015). Narrative skill in children with early unilateral brain injury: A possible limit to compensation. *Developmental Science*, 18(1), 35-52.