Oral narrative comprehension and individual differences in 7 and 11-year-old children: vocabulary, working memory and Theory of Mind

Abstract: The purpose of this study was to analyze the role of vocabulary, working memory, Theory of Mind, and sustained attention in the comprehension of oral narratives and the generation of inferences in children aged 7 to 11. Reading comprehension skills were used as the criterion to establish the study’s cohort groups. A total of 126 Spanish-speaking Colombian children participated in the study. They were assessed using tasks that measured comprehension of oral narratives, as well as standardized tests of vocabulary, working memory, Theory of Mind, and sustained attention. The results indicated a significant difference in vocabulary between the group of poor comprehenders and the comparison group with better comprehension performance. The poor comprehenders also performed worse than the comparison group in the narrative comprehension test with visual support, as a result of significant differences in responding to emotional inference questions. The correlational analysis showed that the vocabulary variable was associated with all measures of oral narrative comprehension. Additionally, working memory correlated with almost all measures of oral narrative comprehension, except for answering inference questions when the text did not have visual support. Finally, Theory of Mind demonstrated a significant correlation with answering emotional inference questions and the measures without visual support.

Keywords: narrative comprehension; vocabulary; working memory; Theory of Mind; sustained attention

Resumen: El propósito de esta investigación consistió en estudiar el rol del vocabulario, la memoria de trabajo, la teoría de la mente y la atención sostenida en la comprensión de narraciones orales y la generación de inferencias en niños de entre 7 y 11 años. Para ello, se utilizaron las habilidades de comprensión lectora como criterio para establecer los grupos de cohorte del estudio. La muestra, conformada por 126 niños colombianos hispanohablantes, fue evaluada en tareas de comprensión de narraciones orales, así como mediante pruebas estandarizadas en vocabulario, memoria de trabajo, teoría de la mente y atención sostenida. Los resultados mostraron diferencias significativas entre el grupo de baja comprensión lectora respecto al grupo de comparación con mejores desempeños en comprensión en la medida de vocabulario y en la medida de comprensión de narraciones con apoyos visuales, como producto de una diferencia significativa en la respuesta a preguntas de inferencia emocional. El análisis correlacional mostró vínculos significativos entre el vocabulario y todas las medidas de comprensión evaluadas. Por su parte, la memoria de trabajo se correlacionó con todas las medidas de desempeño en comprensión, excepto con la respuesta a preguntas inferenciales. Finalmente, la teoría de la mente mostró una correlación significativa con la respuesta a preguntas de inferencias emocionales y con las medidas de comprensión de narraciones sin apoyos visuales.

Palabras clave: comprensión de narraciones; vocabulario; memoria de trabajo; teoría de la mente; atención sostenida
Comprehension is the result of processing text or discourse and is relevant to personal and social development (Gray, 2017). In order for this to be fully understood, the reader or listener must strategically fill the information gaps that are not explicit in texts or the discourse. This involves adding previous experiences and activating the information about the world as stored in the semantic memory, thus establishing connections with the information from the text, be it either written or oral (Kintsch, 1998; Kintsch & van Dijk, 1978).

Comprehension involves the construction of a coherent, significant text or discourse representation. This process develops a level of representation known as "situation model" (Kintsch, 1998), which constitutes a cognitive representation of the elements, individual properties, relations and facts present in the text, be it either written or oral. Additionally, such representation involves other sources of knowledge that are linked to previous experiences and knowledge about the world. For example, in narrative texts, comprehension may require information about the emotion as a motive, the nature of social interactions or moral judgments. Such additional elements play a key role in narrative text comprehension, and it is important for its influence on a more complete, precise comprehension to be taken into account.

In this regard, constructing the situation model in narrations could be less complex as compared to other types of text, since the previous knowledge that is required in narrative texts comes from the reader's everyday life (Clinton et al., 2020). An inference means the activation of information that is not literally expressed in the text, but that is needed to fill in details for improving the coherence of the sentence or utterance (Currie & Cain, 2015). Graesser et al. (1994) provide an inference classification encompassing the inferences both made during the reading (online) and thereafter (offline). The authors suggest that this classification is placed in a continuum and that it may vary based on the individual abilities, the reader's goals, the text type, the experimental tasks and the methodology followed. Within this classification, the inferences related to the characters' emotional response and the causal consequence inferences stand out. The former refer to the emotions that are experienced by a character as a result of an event or an action, while the latter focus on a causal chain involving physical events and new agent or character plans.

By inferring contents of emotional type from the information that is presented in the text, the reader is able to comprehend the psychological causes of the characters. Thus, emotional inferences enrich the reading experience by allowing the reader to understand the psychological and emotional dimension of the characters. Some authors propose that the reader uses the same mechanisms to understand emotions in their real life as those to comprehend narrative texts (Kim, 2020; Mumper & Gerrig, 2021).

Text comprehension involves various aspects and requires developing multiple skills. Research in psycholinguistics has established a relation between text comprehension and different cognitive factors, such as vocabulary, working memory and sustained attention (Barreyro et al., 2020; de Bree & Zee, 2020; Lervåg et al., 2018). The connection between the Theory of Mind (ToM) and text comprehension performance (Dore et al., 2018; Kim et al., 2021; Tong et al., 2020) has also been explored. These findings underscore the importance of taking a wide range of cognitive skills and mental processes in the study of text comprehension.
Emerging research has suggested direct and indirect effects from various cognitive and language skills that would be involved in reading comprehension (Kim, 2016, 2017, 2020). In fact, Kim (2020) proposes that, through hierarchical relations, different skills play roles that are more essential than others. In this hierarchical structure, some skills, such as word reading, listening comprehension or reading fluidity, would impact directly on the comprehension, while others, such as ToM, inference generation, monitoring, among others, would have an indirect effect through interactions with other skills.

Vocabulary, or mental lexicon, is defined as information that is stored in the semantic memory, which allows for a phonetic (or graphemic) chain, the morphological and syntactical components thereof, as well as the meaning or semantic load thereof, to be recognized (Raiter & Jaichenco, 2002). Evidence suggests that the mental lexicon extent is correlated to a better comprehension performance (Bråten et al., 2022; Daugaard et al., 2017; Figueroa Sepúlveda & Gallego Ortega, 2021; Shahar-Yames & Prior, 2018; Sterpin et al., 2021). Additionally, vocabulary critically affects the comprehension of oral narrations (Babayiğit & Shapiro, 2020; Barreyro et al., 2023; Cain & Oakhill, 2011; de Bree & Zee, 2020; Kim, 2015, 2016; Suggate et al., 2018; Wolf et al., 2019).

A study that was carried out in Spanish-speaking population of Chilean origin (Figueroa Sepúlveda & Gallego Ortega, 2021) found that vocabulary is the biggest predictor of comprehension in students. Similarly, Strasser and del Rio (2014) found that both the extent and the depth of vocabulary had a mediating effect on the comprehension of stories which were read by an examiner and feature illustrations, while working memory had a direct effect on the measures of comprehension. These findings highlight that a larger lexicon provides a solid foundation to comprehend and process the textual content, both oral and written, more efficiently.

Similarly, numerous studies have proven the relation between text comprehension and working memory (Currie & Cain, 2015; Hannon, 2012; Language and Reading Research Consortium et al., 2018; Lerвåг et al., 2018, Linares & Pelegrina, 2023; Mar et al., 2021). Working memory refers to a temporary information storage and handling system supporting complex cognitive processes (Baddeley, 2021). In a systematic review on research covering narrative skills and executive functions, a positive, moderate relation between working memory and narrative skills was found (Moreira et al., 2023). Likewise, a longitudinal study carried out by Jiang & Farquharson (2018) researched the influence of working memory on the variation of the oral discourse comprehension. The results indicate that the measures of working memory were related to the performance in the oral discourse comprehension. On the other hand, Barreyro et al. (2020) found a significant correlation between measures of working memory and of sustained attention and the performance on inferential questions on a sample of 5- to 6-year-old pre-reader children. These findings underscore the importance of working memory in text comprehension. The ability to efficiently retain and handle information in working memory makes it easier for ideas and the construction of a coherent text representation to be integrated. Therefore, taking the role that working memory has in the design of educational interventions intended to improve reading comprehension into account is key.

Sustained attention, understood as the ability to maintain the attentional focus on a task for extended periods of time (Sarter et al., 2001), plays a crucial role in text comprehension. From a theoretical point of view, an association between sustained attention and comprehension is expected to exist, considering their link to working memory (Baddeley, 2021). Despite its importance, scientific literature on the study of sustained attention in relation to text comprehension is limited. However, recent studies have started to explore this cognitive process and the implications thereof for comprehension performance (Burgos Ortiz & Polanco Barreto, 2019; Hitch et al., 2020; Joseph et al., 2021; Micaи et al., 2020). In this context, a recent study by Barreyro et al. (2020) found a significant correlation between sustained attention and performance on inferential questions. These findings support the idea that the ability to maintain attention in a sustained manner may have a positive impact in reading comprehension.

ToM refers to the ability of putting oneself in someone else’s place, of imagining what they think and feel, as well as understanding and foreseeing their behavior (Baron-Cohen, 2008, p. 90). It is estimated that around the age of 4 and a half years, children go through the false belief task (Wimmer & Perner, 1983), showing that they are able to correctly predict the “wrong” action by the character that has been deceived in the story. Thus, there is predominance in ToM research on preschoolers. However, some studies propose that both ToM and the narrative processing abilities continue to increase their
complexity in subsequent stages (Devine & Hughes 2013, 2016; Doree et al., 2018). Dore et al. (2018) propose that ToM could be the missing piece that would allow to explain text comprehension, since it would be difficult for a child who has a hard time attributing emotions and intentions to others to have the ability to comprehend narrative texts and generate emotional inferences.

Recent studies show the significant role of ToM in the comprehension of narrations (Boerma et al., 2017; Kim, 2015, 2016, 2017, 2020). For example, Kim et al. (2021) evaluated 132 fourth-graders with an average age of 10.39 years on cognitive measures, such as vocabulary, ToM, working memory, comprehension monitoring, grammatical knowledge and attentional control. The comprehension measures included tasks with narrative, informative texts, and a task that evaluated the recall of mental states in the presented texts. The findings reported ToM as the only predictor of the grade in which children included mental states in their recall of narrative and informative texts. Additionally, the measure of ToM was positively correlated to the comprehension measures. The results are consistent with other studies (Atkinson et al., 2017; Boerma et al., 2017; Florit et al., 2020; Happé & Frith, 2020; Kim, 2020; Kim et al, 2021; Manfredi et al., 2020), and highlight the importance of ToM in the comprehension of narrative texts and the influence thereof on cognitive processing during reading.

The purpose of this research is to analyze the difficulties specific to the comprehension of narrations in children with low performance in comprehension tasks as compared to a group that gets better results. It will be researched how variables, such as vocabulary, working memory, ToM and sustained attention, affect the comprehension of narrations. This analysis will allow for the factors affecting the comprehension of narrations in children to be better understood, which may be crucial for developing intervention strategies that are effective.

Method

Participants
The sample consisted of 126 Colombian students between 7 and 11 years old, 68 males (76.4 %) and 21 females (23.5 %). The participants were divided in two groups based on their performance according to the reading comprehension subtest from the Magallanes Scale Test for Reading and Writing (Escala Magallanes de Lectura y Escritura) (Toro et al., 2002). The distribution of groups was as follows:

- **Group with low performance in text comprehension.** Consisting of schooled, typically developed 43 boys and 19 girls (average age: 9.37 years), with reading speed and non-verbal IQ within normal range, but with difficulties in text comprehension. No medical record of delayed development or language. (Group BC, n = 62). Performance equal or below the 25th percentile in the reading comprehension subtest from the Magallanes Scale Test for Reading and Writing (Toro et al., 2002).

- **Comparison group with better comprehension performance.** Consisting of schooled, typically developed 40 boys and 24 girls (average age: 9.12 years); no medical record of delayed development or language; in turn, with reading speed, text comprehension, non-verbal IQ within normal range. (Group C, n = 64). Performance above the 25th percentile in the reading comprehension subtest from the Magallanes Scale Test for Reading and Writing (Toro et al., 2002).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characterization of the Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Group BC</td>
<td>62</td>
</tr>
<tr>
<td>Comparison Group</td>
<td>64</td>
</tr>
</tbody>
</table>

*Note. SD: Standard Deviation*

Instruments

**Working Memory Task**
The Number Recall subtest from the Kaufman Assessment Battery for Children (K-ABC) (Kaufman & Kaufman, 2014) was used. The purpose of the test is to measure working memory’s ability to store verbal information. In this test, the subject must repeat orally a sequence of digits in the presented order. The sequences progressively increase in complexity, and every correct sequence
awards one point. The total score is calculated by subtracting the failed items from the item of higher number.

**Vocabulary Task**

The vocabulary subtest from the Kaufman Brief Intelligence Test (K-BIT) (Kaufman & Kaufman, 2000) was administered. The test is a measure of the verbal ability that requires oral responses and inquires on the quantity of words that the examinee knows; it consists of 82 items, which evaluate vocabulary. It presents two parts: Part A, which evaluates expressive vocabulary, requires for the person to provide the name of an object represented graphically; and Part B, which requires for the subject to give an answer from two clues: a descriptive verbal expression and a word missing some letters. For example, the description “a dark color” is presented, and an incomplete word: “B–C–” is provided for the correct answer “black”. Scores are dichotomous in nature: the correct answer is awarded 1 point, and the incorrect answer is awarded 0; there are no partial scores. The Spanish adaptation presents suitable levels of validity and reliability.

**Sustained Attention Task**

Likewise, the Symbol Search subtest from the Wechsler Intelligence Scale for Children (WISC-V) (Wechsler, 2014) was used. The child observes one or two target symbols and a search group; the task is to decide if any of the target symbols are present within the search group in a time-bound manner (120 seconds). The test begins with example and training items intended to ensure that the instruction is clear. The direct score is the sum of correct answers minus the number of incorrect answers plus the bonus points for time when appropriate. This test allows to assessing the visual discrimination of abstract stimuli, cognitive flexibility, perceptive speed and precision, among others (Cayssials, 1998).

**Theory of Mind Task**

This basic ability to relate to others and attributing other states, beliefs and emotions was evaluated as follows: The participant began with the classic false belief task of “Sally and Ana” (Baron-Cohen, 2008), which assesses the first-order mental representation about the others’ belief. It was necessary for the child to pass the false belief task in order to continue to the next, social awareness task. In order to evaluate this inter-subjective ability in a more complex way, the Metida de Pata (“Foot in mouth”) test (Baron-Cohen et al., 1999) was applied. It is intended for children between 7 and 11 years old, and presents a good internal consistency (Cronbach’s alpha = .82). The Argentinean version carried out by Grañana et al. (1999), and the adaptation of this version suitable to the Colombian context by Gutierrez (2012), were used. Detecting a “foot in mouth” requires from the subject to discriminate between the state of the speaker’s and their partner’s knowledge. Moreover, an evaluation of the emotional impact on the partner is required. This test is based on presenting 10 voice-recorded stories involving two or three characters and, at least two sentences, wherein a simple, everyday language is used. In order to detect the foot in mouth, the child must answer correctly four inquiries in each story. Therein, the subject is verified as identifying in the story if someone said something that should not be said, what was said that should not be said, or some approximation thereto, if the comprehension question is asked, and lastly, if the “foot in mouth” is acknowledged as a consequence of a false belief. One point per story is awarded if the child manages to answer correctly all four questions.

**Reading Comprehension Task**

The purpose of the reading comprehension subtest from the Magallanes Scale Test for Reading and Writing (Toro et al., 2002) is to assess the reader's ability to comprehend the message of the text, including the inference of the characters’ emotional states, their physical location, the sequence of events and the cause-effect relationships. The narrative texts and the questions are characterized by covering everyday subjects, employing a varied vocabulary, usual phrases and, occasionally, metaphors. During the application, a text according to the participant’s school level is selected, followed by 10 questions evaluating both inferential and literal recall aspects by multiple selection with only one answer. There is no set time limit, but unnecessary delays are recommended to be avoided, and the texts remain available during the question phase. The discriminating validity of the test, as evaluated by using the Cohen’s kappa index (.677), supports its capacity to distinguish between poor-comprehension groups and those with better performance. In the present study, this instrument was used as inclusion criterion for the poor-comprehension group (Group BC) and the comparison group with better
comprehension performance. Specific levels, as determined by percentile scores obtained by using the software TIPI-SOFT EMLE-TALE 2000 (different for each age group) were used to categorize the participants. Those with performances equal to or below the 25th percentile grouped in the test Level C (level characterized by the inability to extract meaning from the text, understand the characters' situation, the events that are taking place, the sequence thereof, including the cause-effect relationships) were placed in the Group BC, while the remaining participants with scores above the 25th percentile were placed in the comparison group (with test Levels A and B).

**Comprehension of Oral Narrations Task**

The text comprehension subtest from the Prolec-R instrument (Cuetos et al., 2007) evaluated the speaker’s ability to extract the text meaning and merging it with their previous knowledge. In the original version of the test, each subject was assigned three out of the four proposed texts. These texts were presented orally, followed by two inferential questions and two literal information questions for each text. The correct answers were awarded 1, and the incorrect ones, 0. These questions were verbally answered by the participants.

**Comprehension of Narrations with Illustrations Task (ad hoc)**

Lastly, a task device was designed to evaluate the comprehension of oral narrations. To this effect, the tales *La sorpresa de Nandi* (Browne, 2011) and *Hipo no nada* (Bernasconi, 2007) were adapted. The task consisted in presenting the story orally and featuring illustrations that followed the narrations. Then, an evaluation was carried out by means of three types of questions: explanatory-type inferentials, emotional-type inferentials and literal information recall. The test as a whole consisted of four questions of emotional inferences, four questions of explanatory inferences and four questions of the literal content of the text. For example, for the tale *Hipo no nada* a question referred to emotional inference was: ¿How did Hipo feel for not being able to play with the rest of the animals? The valid answer related to the idea that the hippopotamus was unsuccessful when trying to play with other animals. As an explanatory inference question, “Why did Hipo feel embarrassed in the moles’ house?” was asked. The correct answer was that the hippopotamus was too big and could not move freely in the moles’ house. And as a literal information question, “Where were the birds?” was asked. The right answer was linked to the memory of the birds being on top of the tree. On the other hand, only the first statement issued by the participant after the examiner asked the question was accounted for as a valid answer, thus awarding one point per right answer. Such measure is the summation of the answer to emotional inference questions, the answer to explanatory inferences questions and text literal recall. The device showed a good internal consistency as it is an ad-hoc instrument (Cronbach’s alpha = .732).

**Procedure**

The children were evaluated in two individual sessions of about 60 minutes with a 10-minute recess. In the first session, K-BIT subtests, matrices and vocabulary (Kaufman & Kaufman, 2000), the text comprehension subtest of the Magallanes Scale Test for Reading and Writing (Toro et al., 2002), and social awareness tests were administered. In the second session, the Symbol Search subtest from the WISC-V (Wechsler, 2014), the Number Recall subtest from the K-ABC (Kaufman & Kaufman, 2014), the text comprehension subtest from the Prolec-R (Cuetos et al., 2007), and the comprehension of oral narrations task were administered.

Informed consent, which allowed to putting on record the voluntary participation, was asked from parents and/or guardians. This cleared the use of data in a confidential manner and for research purposes only. The tests were applied after the parents or guardians gave their written authorization.

**Results**

Data was analyzed by using the SPSS statistical package, thus fixing the level of statistical significance in .05. Preliminary descriptive tests, which are detailed in Table 2, were conducted. Then, the study of the differences between the group of children with poor comprehension performance (Group BC) and the comparison group with better comprehension performance, both in comprehension measures and in cognitive measures, was undertaken. Finally, the potential correlations between the types of variables above were analyzed.
Table 2
Descriptive statistics of cognitive measures and narrative comprehension measures

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
<th>As.</th>
<th>Cu.</th>
<th>KS</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-ABC Working Memory Test</td>
<td>54.12</td>
<td>17.05</td>
<td>20</td>
<td>95</td>
<td>.470</td>
<td>-.345</td>
<td>.143</td>
<td>.00</td>
</tr>
<tr>
<td>Theory of Mind Test and “Foot in Mouth” Test</td>
<td>41.37</td>
<td>24.09</td>
<td>0</td>
<td>90</td>
<td>.100</td>
<td>-.860</td>
<td>.106</td>
<td>.001</td>
</tr>
<tr>
<td>WISC-V Sustained Attention Test</td>
<td>44.32</td>
<td>12</td>
<td>15</td>
<td>75</td>
<td>.020</td>
<td>-.221</td>
<td>.102</td>
<td>.003</td>
</tr>
<tr>
<td>Prolec-R</td>
<td>86.77</td>
<td>14.11</td>
<td>25</td>
<td>100</td>
<td>-1.35</td>
<td>2.60</td>
<td>.183</td>
<td>.00</td>
</tr>
<tr>
<td>Prolec-R Inferences</td>
<td>85.18</td>
<td>16.62</td>
<td>33.33</td>
<td>100</td>
<td>-.913</td>
<td>.164</td>
<td>.274</td>
<td>.00</td>
</tr>
<tr>
<td>Prolec-R Literal Information</td>
<td>88.62</td>
<td>15.58</td>
<td>16.67</td>
<td>100</td>
<td>-1.92</td>
<td>4.84</td>
<td>.291</td>
<td>.00</td>
</tr>
<tr>
<td>Narration Task (Ad Hoc)</td>
<td>90.01</td>
<td>10.27</td>
<td>50</td>
<td>100</td>
<td>-1.27</td>
<td>1.93</td>
<td>.231</td>
<td>.00</td>
</tr>
<tr>
<td>Answers to Emotional Inference Questions</td>
<td>82.14</td>
<td>19.96</td>
<td>25</td>
<td>100</td>
<td>-.850</td>
<td>-.039</td>
<td>.291</td>
<td>.00</td>
</tr>
<tr>
<td>Answers to Explanatory Inference Questions</td>
<td>93.45</td>
<td>11.47</td>
<td>50</td>
<td>100</td>
<td>-1.34</td>
<td>0.46</td>
<td>.462</td>
<td>.00</td>
</tr>
<tr>
<td>Answers to Literal Information Questions</td>
<td>93.84</td>
<td>12.11</td>
<td>50</td>
<td>100</td>
<td>-1.81</td>
<td>2.52</td>
<td>.472</td>
<td>.00</td>
</tr>
</tbody>
</table>

Notes. M: mean; SD: Standard deviation; Min.: Minimum; Max.: Maximum; As.: Asymmetry; Cu.: Kurtosis; KS: Kolmogorov-Smirnov test; p: Level of statistical significance.

As shown in Table 2, the scores (expressed as a percentage) in all tests showed distributions that were significantly distanced from the normal, asymptotic distribution. For this reason, non-parametric tests were employed for subsequent analyses in the study of said scores.

In a second time, the poor-comprehension-performance group was compared to the comparison group by using the Mann-Whitney U test with the purpose of examining the inter-group differences in the vocabulary, working memory, ToM and sustained attention measures, and the oral narration comprehension measures (answer to inferential-type and text literal recall questions). Table 3 shows both groups’ performances.

Table 3
Performances of the poor-comprehension-performance group (Group BC) and the comparison group in cognitive and comprehension measures

<table>
<thead>
<tr>
<th></th>
<th>Group BC</th>
<th></th>
<th>Comparison Group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>K-BIT Vocabulary Test</td>
<td>60.47</td>
<td>11.12</td>
<td>63.97</td>
<td>9.89</td>
</tr>
<tr>
<td>K-ABC Working Memory Test</td>
<td>51.29</td>
<td>15.09</td>
<td>56.87</td>
<td>18.46</td>
</tr>
<tr>
<td>Theory of Mind Test and “Foot in Mouth” Test</td>
<td>41.22</td>
<td>24.48</td>
<td>41.51</td>
<td>23.90</td>
</tr>
<tr>
<td>WISC-V Sustained Attention Test</td>
<td>43.46</td>
<td>11.68</td>
<td>45.15</td>
<td>12.34</td>
</tr>
<tr>
<td>Prolec-R</td>
<td>84.13</td>
<td>16.23</td>
<td>89.32</td>
<td>11.25</td>
</tr>
<tr>
<td>Prolec-R Inferences</td>
<td>82.79</td>
<td>18.59</td>
<td>87.50</td>
<td>14.24</td>
</tr>
<tr>
<td>Prolec-R Literal Information</td>
<td>85.48</td>
<td>18.22</td>
<td>91.66</td>
<td>11.87</td>
</tr>
<tr>
<td>Narration Task (Ad Hoc)</td>
<td>87.90</td>
<td>11.35</td>
<td>92.05</td>
<td>8.71</td>
</tr>
<tr>
<td>Answers to Emotional Inference Questions</td>
<td>78.22</td>
<td>20.98</td>
<td>85.93</td>
<td>18.29</td>
</tr>
<tr>
<td>Answer to Explanatory Inference Questions</td>
<td>91.93</td>
<td>12.62</td>
<td>94.92</td>
<td>10.13</td>
</tr>
<tr>
<td>Answer to Literal Information Questions</td>
<td>92.74</td>
<td>12.30</td>
<td>94.92</td>
<td>11.93</td>
</tr>
</tbody>
</table>

Notes. M: mean; SD: Standard Deviation.
The results in cognitive measures showed that the mean performance in the poor-reading-comprehension group was significantly lower than that of the comparison group in the vocabulary measure \(U = 1505, z = -2.33, p < .01\). However, no significant differences were found in the groups in the following measures: Working memory \(U = 1663, z = -1.57, p > .05\), ToM \(U = 1974, z = -0.04, p > .05\) and sustained attention \(U = 1830, z = -.755, p > .05\).

Performances in the oral narration comprehension measures revealed significant inter-group differences of performance in the ad hoc comprehension of narrations task \(U = 1559, z = -2.16, p < .05\), as well as in the answer to emotional inference questions \(U = 1573, z = -2.18, p < .05\). The remaining measures did not show differences between the two groups: Prolec-R \(U = 1650, z = -1.68, p > .05\), Prolec-R inferences \(U = 1746, z = -1.24, p > .05\), Prolec-R literal information \(U = 1639, z = -1.86, p > .05\), answer to explanatory inference questions \(U = 1772, z = -1.36, p > .05\) and answer to literal recall questions \(U = 1792, z = -1.29, p > .05\).

For the purpose of analyzing the relations between vocabulary, working memory, ToM, sustained attention and performance in the oral narration comprehension measures, a study on the correlations was carried out by employing the Spearman’s rank statistic (Rho) in the general group.

Table 4 shows the correlation values between the evaluated measures. The vocabulary measure was significantly correlated to all Prolec-R comprehension measures \((Rho = .354, p < .01\), Prolec-R inferences \((Rho = .311, p < .01\), Prolec-R literal information \((Rho = .207, p < .01\), comprehension task \((Rho = .411, p < .01\), answer to emotional inference questions \((Rho = .221, p < .013\), answer to explanatory inference questions \((Rho = .353, p < .01\) and answer to literal information \((Rho = .393, p < .01\).

Performance in the K-ABC working memory subtest was significantly associated to the measures: Prolec-R \((Rho = .354, p < .01\), Prolec-R inferences \((Rho = .322, p < .01\), Prolec-R literal information \((Rho = .277, p < .01\), comprehension task \((Rho = .397, p < .01\), answer to emotional inference questions \((Rho = .298, p < .01\) and answer to literal information \((Rho = .425, p < .01\). The ToM measures and next tests showed significant correlations: Prolec-R \((Rho = .308, p < .01\), Prolec-R inferences \((Rho = .244, p < .006\), comprehension task \((Rho = .241, p < .006\), answer to emotional inference questions \((Rho = .261, p < .003\). The sustained attention measure failed to reveal correlations that were significant with the evaluated measures on comprehension.

Table 4
Spearman’s Correlation between Cognitive Measures and Oral Narration Comprehension Measures

<table>
<thead>
<tr>
<th></th>
<th>K-BIT Vocabulary</th>
<th>K-ABC Working Memory</th>
<th>ToM and “Foot in Mouth” Test</th>
<th>WISC-V Sustained Attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prolec-R</td>
<td>.399**</td>
<td>.354**</td>
<td>.308**</td>
<td>.084</td>
</tr>
<tr>
<td>Prolec-R Inferences</td>
<td>.302**</td>
<td>.311**</td>
<td>.244**</td>
<td>.107</td>
</tr>
<tr>
<td>Prolec-R Literal Information</td>
<td>.407**</td>
<td>.277**</td>
<td>.314**</td>
<td>.047</td>
</tr>
<tr>
<td>Narration Task (ad hoc)</td>
<td>.411**</td>
<td>.396**</td>
<td>.241**</td>
<td>.045</td>
</tr>
<tr>
<td>Answers to Emotional Inference Questions</td>
<td>.221*</td>
<td>.298**</td>
<td>.261**</td>
<td>-.049</td>
</tr>
<tr>
<td>Answer to Explanatory Inference Questions</td>
<td>.353***</td>
<td>.082</td>
<td>.157</td>
<td>.106</td>
</tr>
<tr>
<td>Answer to Literal Information Questions</td>
<td>.393**</td>
<td>.425**</td>
<td>.067</td>
<td>.043</td>
</tr>
</tbody>
</table>

*p < .05 **p < .01

Discussion

The purpose of this research was to study the oral narration comprehension and its relation to working memory’s ability, vocabulary, ToM and sustained attention in children between 7 and 11 years old. Two groups were compared: the first one, consisting of typically developed children with poor comprehension performance, and the second, a comparison group consisting of their peers with better reading comprehension performances.

The results of this study indicate that vocabulary is a good predictor of comprehension in readers between 7 and 11 years old, since significant inter-group differences in this cognitive variable are identified. Similarly, these findings showed significant correlations between vocabulary and all
comprehension measures that were evaluated. Consequently, the results suggest that an extensive mental lexicon, as well as inference generation, increases the chance of comprehending a story that is orally narrated. These findings are consistent with the work of Strasser and Del Rio (2014), in which both the extent and the depth of vocabulary have a mediating effect on the comprehension of stories which are read by an examiner and that feature illustrations. They are also linked in the study of Barreyro et al. (2023), who found a correlation between vocabulary and some comprehension measures, such as inference generation, in preschoolers. Additionally, the authors found an effect in vocabulary and inference generation to foresee the comprehension of narrations that are presented orally. Likewise, the results of the present research are consistent with the findings of other studies (Barreyro et al., 2020; Gottheil et al., 2019; Sterpin et al., 2021).

On the other hand, the key role that working memory plays in oral narration comprehension, as well as in the ability to make inferences and remember literal details from the text, is remarkable. This conclusion is supported by the correlations that were found between working memory and the comprehension measures as evaluated in narrations with illustrations (ad hoc task), as well as in narrations without illustrations (Prolec-R). These results are consistent with the findings of Strasser and Del Rio (2014), who identified a direct effect from working memory on the comprehension measures. Also, they are supported by a recent systematic review that included 38 studies and showed the connection between working memory and the narrative comprehension (Moreira et al., 2023). However, it must be pointed out that, while the working memory variable shows correlation with most comprehension measures, no significant differences between groups were observed on this cognitive measure. Therefore, it is possible that working memory is involved in oral narration comprehension, but its impact cannot be completely captured by the correlation analysis. Lastly, a significant relation between working memory and the text literal recall measures was identified. This could suggest that a greater capacity to store information in a temporary manner is associated to better oral narration comprehension, since it makes it easier for literal information to be recalled.

Regarding the sustained attention measure, no significant differences or a correlation with the comprehension measures were found. These results differ from those found in other studies (Barreyro et al., 2020; Strasser & Del Rio, 2014), but they can be supported as follows: Firstly, the limited number of studies inquiring about this cognitive process in text comprehension and production (Moreira et al., 2023) is underscored. Secondly, the methodological variation as implemented in order to operationalize the attention measure could be linked to the differences that were found. For example, Strasser and del Rio (2014) found that attention and inhibitory control explained the comprehension performance in preschoolers. However, attentional control was evaluated by using a checklist that was filled by the examiner after administering the narration comprehension measure. Other studies evaluate the total time in problem resolution or by the number of right answers in a preset time period. Thirdly, maturational development could have explanatory value in the results of this study. The comparative analysis of Barreyro et al. (2020) showed a higher ability to maintain sustained attention, as well as better comprehension performances in 6 year-old children as compared to 5 year-old children. These results are consistent with those reported by other studies (Currie & Cain, 2015; Lervåg, 2018; Linares & Pelegrina, 2023). Therefore, the findings of this research could suggest that children older than 7 years are not in a developmental stage wherein sustained attention has a direct incidence on oral narration comprehension.

The ability to understand emotional states, beliefs and others’ emotions (ToM) could be linked to better performances in the narrative process ability. In fact, authors such as Kim (2020) propose a causal relation between ToM and the narrative skills, which are developed at the same time during the different stages of neurodevelopment. In this order of ideas, 7 year-old children would have the ability to understand mixed emotions and nonliteral situations (Devine & Hughes, 2013, 2016; Manfredi et al., 2020; Tong et al., 2019, 2020). And in the narrative process, for this age, a more advanced follow-up of the characters’ mental and emotional states would be expected. For example, the follow-up of the character’s goals. Nyhout (2015) found in 7 year-old children a spontaneous representation about the characters’ goal. By using the processing time paradigm, the children’s response time was noticed to be longer when the sentences were inconsistent with the goal as presented at the beginning of the narration (presentation in oral mode) in relation to the sentences that were consistent.
The results of the present research not only proved a significant correlation between the ToM measure and the performances on the answer to emotional inference questions. Relevant links between the ToM and the two comprehension measures, with or without visual aid are also revealed. Likewise, the Prolec-R inference measure (without visual aid) can also be associated to this cognitive variable. The findings show that the ability to understand emotional states, beliefs and others’ emotions could be linked to better performances in oral narration comprehension and to the generation of emotional inferences from a narrative text for children between 7 and 11 years old.

According to the results of this study, the following recommendations for educational interventions can be suggested: (1) it is essential for docents to include integral strategies when teaching reading comprehension. Such strategies have to focus on the expansion and improvement of vocabulary, as well as on the development of mental abilities. (2) Fostering the generation of emotional inferences can contribute to a better performance in narrative discourse comprehension. For example, it could be focused on the development of narrative skills that are complex and suitable to the age of the students, such as following up the characters’ advanced mental and emotional states. (3) The results of this study suggest the need to diversify the material that are used to cover narrative text comprehension in the classroom. In addition to illustrated books, graphic novels or videos, the use of podcasts, radio programs, audiobooks, voice recordings, among others, could also be considered.

Some limitations of the present study must be taken into account. The sample has been formed following a nonrandom criterion for convenience by taking educational institutions that volunteered to take part in the study. Similarly, the descriptive-correlational nature of the study is underscored as a limitation, since while relations between the variables can be identified, this design fails to allow to establishing causal relations. Also, other variables should be included in order to evaluate attentional processes and the link thereof to narration comprehension with higher precision, for example, monitoring comprehension and inhibition.

Another limitation of the study is linked to the type of measure that is used for working memory, since the study examined the storage capacity and failed to incorporate active manipulation. On the other hand, it should be highlighted that the answer to explanatory questions in the task with visual aids was the only comprehension measure that did not show any correlation to working memory. This could be explained by the fact that, in this specific task, the readers were provided with illustrations to support the narration. Therefore, they did not need to use cognitive resources from their working memory, but relied upon their vocabulary. In this sense, a significant correlation between the vocabulary measure and this inferential comprehension was found. However, in order to prove this hypothesis, an interaction or moderation analysis must be carried out, which is one of the limitations of the present study.

Further research based on this research could explore more accurately the link between ToM and the narrative skills in children older than 7 years of age, considering that there has been little research in this age group (Dore et al., 2018). In particular, it would be interesting to explore the narration comprehension from an intervention program on ToM tasks adapted to the neurodevelopment stage. Finally, researching clinical groups on the incidence of visual aids on narration comprehension would add empirical value to their educational inclusion processes.

Referencias


J. P. C. has contributed in 1, 2, 3, 5, 6, 8, 12, 13; J. P. B. in 1, 3, 6, 10, 14.

**Scientific editor in charge:** Dra. Cecilia Cracco.