Early Childhood Education Journal
https://doi.org/10.1007/s10643-026-02190-z

# An Individualized Multimodal Narrative Intervention for Young Children with Neurodevelopmental Disorders

**Júlia Florit-Pons**[2] · **Pilar Prieto**[1,2] · **Alfonso Igualada**[3,4,5]

Received: 18 March 2025 / Accepted: 19 March 2026


## Abstract

This study examines whether an individualized narrative intervention can boost the narrative (macrostructure and perspective-taking) abilities of children with neurodevelopmental disorders (NDD). An experimental group of 16 children with NDD from Catalonia, Spain, received 9 MultiModal Narrative intervention sessions. Children's narrative skills were measured pre- and post-intervention with a retelling task. Their performance was compared to two control groups not receiving the intervention (17 children with NDD and 17 with typical development). After each session, children in the experimental group underwent two Dynamic Assessment measures (retelling task and graduated prompting comprehension questions about the story), tracking their learning. Results showed an improvement in narrative macrostructure but not perspective-taking. The Dynamic Assessment revealed faster learning of macrostructure, individual differences in the prompts needed for successful responses, and that fewer prompts predicted better macrostructure outcomes post-intervention. These findings underscore the relevance of diverse measures in evaluating continuous learning in interventions.

**Keywords** Narrative intervention · Multimodality · Neurodevelopmental disorders · Narrative macrostructure skills · Narrative perspective-taking skills · Dynamic assessment

## Introduction

The ability to construct an oral narrative discourse by retelling or generating a story is one of the key milestones in children's language development. Constructing a narrative discourse involves the organization of linguistic elements into a temporally and causally connected story. This ability is complex, as it requires not only using of the appropriate vocabulary and morphosyntactic structures but also organizing these structures in the correct order while framing them in a relevant socio-communicative context. That is, using language in a way that is adapted to the social situation to effectively communicate with others. Together, these skills allow children to express (a) what happened in the story by introducing the core elements of the story, such as protagonist, problem, attempt and resolution (i.e., narrative macrostructure, also known as story grammar; Favot et al., 2021); and (b) what were the perspectives, emotions and mental states of the characters (i.e., narrative perspective-taking, also known as internal state language or inferential language; Dodd et al., 2011; Siller et al., 2014). These narrative skills have been shown to be both predictors and precursors of later language skills,

✉ Júlia Florit-Pons
julia.florit@upf.edu

Pilar Prieto
pilar.prieto@upf.edu

Alfonso Igualada
aigualada@uoc.edu

1 Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia, Spain

2 Dept. of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Catalonia, Spain

3 NeuroDevelop eHealth Lab, eHealth Center, Universitat Oberta de Catalunya, Barcelona, Catalonia, Spain

4 Dept. of Psychology and Education Sciences, Universitat Oberta de Catalunya (Barcelona), Barcelona, Catalonia, Spain

5 Institut Guttmann, Institut Universitari de Neurorehabilitació, Barcelona, Catalonia, Spain

Published online: 11 April 2026

academic success, and socialization (e.g., Babayiğit et al., 2021; Dickinson & McCabe, 2001). It is during the ages of 4 to 6 that children start producing narrative discourses that are more complex and include the core elements of a story, structured coherently and cohesively. Nevertheless, children with neurodevelopmental disorders (NDD), such as autism or Developmental Language Disorder (DLD), overlap in their difficulties with narrative and socio-communicative skills. Overall, children with NDD (both autism and DLD) have been shown to produce narrative discourses that are significantly below those of age-matched typically developing peers (for reviews, see Baixauli et al., 2016; García-Pérez et al., 2008; Norbury & Bishop, 2003; Winters et al., 2022). They often have difficulties in identifying the different core elements of a story and their narrative discourses include minimal or no reference to characters' emotions and perspectives.

To foster oral narrative skills in children with NDD, many researchers and practitioners have been devoted to designing and implementing narrative-based interventions. As reported by a recent systematic literature review examining 24 narrative intervention studies for children with language disorder aged between 3 and 18 (Favot et al., 2021), most interventions have been developed to target children's narrative macrostructure skills and/or linguistic productivity or complexity within the discourse (also known as narrative microstructure). Although studies have reported the beneficial effects of narrative instruction on both macrostructure and microstructure, the effect sizes seem to be larger for macrostructure, suggesting that it is easier for children to improve in terms of identifying and reproducing the story events. Interestingly, recent studies have shown that oral narrative instruction also transfers to literacy skills, such as listening, reading, or writing in both children with NDD (see e.g., Whalon et al., 2019) and those with typical development (TD) (Spencer & Petersen, 2018). However, most narrative interventions have focused on narrative macrostructure and have generally avoided training and assessing narrative perspective-taking skills (see Dodd et al., 2011; Hessling & Schuele, 2020 for exceptions). These perspective-taking skills are not only relevant for understanding how the characters feel and think throughout the story but are also closely tied to the development of inferential processes, as children need to go beyond literal meanings to interpret the characters' emotions and intentions.

To implement effective interventions that adapt to children's needs, speech-language professionals typically emphasize their need to incorporate tools to track children's learning ability and potential for learning throughout the intervention. One of these tools is Dynamic Assessment (DA), which is an interactive evaluation method that differs from traditional "static" testing by actively supporting the child during assessment. Instead of only measuring independent performance at a single point in time, DA incorporates prompts, feedback, or teaching episodes to evaluate how well a child can learn when given support (Miller et al., 2001). DA can help practitioners understand an individual's learning needs and identify and predict learning potential (Bamford et al., 2022). In fact, DA has been found to be effective in predicting children's language difficulties and diagnosing language disorders (for reviews, see Bamford et al., 2022; Orellana et al., 2019). Using DA during the language intervention process can have diverse purposes, such as monitoring changes throughout the intervention, planning or adjusting the intervention delivery, and predicting changes after the intervention (Camilleri & Botting, 2013; for a review, see Bamford et al., 2022). Within the context of narrative abilities, DA has been implemented using different procedures, such as test-train-retest or prompts together with testing (e.g., Fiestas & Peña, 2018; Peña et al., 2006, 2014). For instance, Peña et al. (2006), implemented two individualized narrative intervention sessions incorporating DA in children aged 6 to 7 with language impairment and typical development. They showed that the intervention had small to moderate effect sizes on children's learning after the sessions. In line with this, Fiestas and Peña (2018) reported that typically-developing 6- to 9-year-old children's narrative macrostructural abilities improved after a brief 2-session learning experience in both English and Spanish. Nevertheless, to our knowledge, no previous narrative-based intervention studies have examined the effect of DA prompting on predicting the change in children with NDD after intervention.

Finally, it should be considered that most narrative interventions have been designed for English-speaking or English-Spanish bilingual children and have not considered other languages (for some exceptions, see Delgado-Cruz et al., 2022, 2024). Assessing the effectiveness of narrative interventions in other languages and contexts might be of interest, as each context has its own differential characteristics. For instance, in the context of Catalonia, where the current study was conducted, preschool and early school education is mostly focused on early literacy skills, rather than oral discourse, meaning that professionals lack validated materials in the Catalan language. Therefore, assessing other educational and clinical contexts has important implications. It provides research evidence on the generalizability of narrative interventions as well as practical implications for professionals, who will be provided with validated materials in their language that are adapted to their context.

This study aims to assess whether an intensive implementation of a multi-tiered narrative intervention (see section

*The MultiModal Narrative intervention* for more details) can boost both narrative macrostructure and perspective-taking skills of children with NDD. To do so, two sets of complementary measures were used. The first consisted of static measures comparing pre- and post-intervention oral narrative outcomes. The second involved DA session-by-session measures to assess continuous learning and learning potential. Crucially, we tested whether individual performance on DA narrative measures can predict macrostructure gains after the intervention. First, we hypothesize that children's oral narrative skills might improve significantly after receiving an intervention that systematically trains narrative macrostructure and narrative perspective-taking. Second, we expect to observe children's continuous narrative learning process through the DA, which will serve to predict their narrative skills post-intervention.

## Methods

To address our aim, we followed a non-randomized control trial design that included three groups (an experimental group with NDD children, a control group with NDD children, and a control group with TD children). Having these three groups allowed for comparing the effectiveness of the intervention at two different levels. First, the comparison with the control NDD group allowed us to differentiate the intervention effects from the general developmental progress and control for task repetition effects. Second, the control TD group served as a baseline to compare NDD children's narrative abilities to those of children with typical development, and thus this helped to assess whether after the intervention, NDD children's performance is similar to that of children with TD.

### Participants

Fifty children ($N=50$) from an initially recruited sample of 66 participated in the study (see Supplementary Material 1 for the CONSORT 2010 flow diagram depicting the process from enrollment until analysis). Sample size was calculated a priori using G*Power Version 3.1.9.6 (Faul et al., 2009) with inputs of 0.05 for alpha, 3 groups, and 2 for the number of measurements (pre- and post-intervention). The necessary sample size to have an acceptable level of power (1-β error probability = 0.8) and to detect small effect sizes of 0.25 was 42 participants.

Participants were TD and NDD children aged 3 to 7 years from Catalonia, Spain (see Table 1 for participants' characteristics and Supplementary Material 2 for the distribution of participants within the NDD groups). All the participants were Spanish-Catalan bilinguals. All children were enrolled in public schools and were therefore taught in a general education setting. TD participants were recruited from a public school in a city in the metropolitan area of Barcelona. As NDD were receiving additional intervention services apart from the general education service at school, they were recruited through contact with speech-language therapists working at public and private speech-language therapy services, or with language-specialized teachers working at public schools. NDD children were required to meet the following inclusion criteria to be eligible for participation in the study: (a) have a previous professional diagnosis or reported risk of autism or DLD (following a transdiagnostic

**Table 1** Participants' characteristics

| Variable | Control TD | Control NDD | Experimental NDD |
|---|---|---|---|
| *N* of participants (*N* of females and males) | 17 (8 F, 9 M) | 17 (5 F, 12 M) | 16 (6 F, 10 M) |
| *N* of participants with a diagnosis (*N* of F and M) | – | 9 (2 F, 7 M) | 9 (1 F, 8 M) |
| *N* of participants with reported risk (*N* of F and M) | – | 8 (3 F, 5 M) | 7 (5 F, 2 M) |
| Age: Mean (SD) | 5.45 (0.24) | 4.75 (0.78) | 5.20 (0.96) |
| Age: Range | 5.08–5.75 | 3.92–7.08 | 3.5–7.33 |
| TSI: Mean (SD) | 102 (0) | 95.06 (9.24) | 102.43 (12.04) |
| TSI: Range | 102 | 87.7–110.8 | 70.5–120.2 |
| CELF: Mean (SD) | 93.06 (12.0) | 82.94 (15.33) | 81.38 (13.67) |
| CELF: Range | 76–115 | 58–106 | 58–105 |
| PleaseApp: Mean (SD) | 29.65 (6.36) | 19.71 (4.90) | 24.00 (6.99) |
| PleaseApp: Range | 19–42 | 10–30 | 13–38 |
| K-BIT: Mean (SD) | 103.06 (13.44) | 101.65 (13.26) | 101.56 (13.70) |
| K-BIT: Range | 83–123 | 74–123 | 82–127 |

The territorial socioeconomic index (TSI) is a value calculated by the Statistical Institute of Catalonia that summarizes the socioeconomic profile of a population in a certain area. The average IST value for Catalonia is 100, with values ranging from 36.5 to 137.1. For this study, we used this value calculated at the municipality level (considering different neighborhoods within big cities) where the school or speech-therapy center that the children were attending was located. For more information see https://www.idescat.cat/pub/?id=ist&lang=en

approach, Astle et al., 2022); (b) have a vocabulary size of at least 50 functional words; (c) be able to use 2-word combinations systematically; and (d) receive weekly intervention sessions with a certified specialist.

In this study, a transdiagnostic approach was followed (Astle et al., 2022) for the following reasons. First, some children did not yet have a differential diagnosis, because they were still young and also because there was a lot of variability in how diagnostic labels were assigned. This is in line with the general reports that young children are often undiagnosed (McGregor, 2020). To solve this issue, some participants were considered to be at risk of being diagnosed with autism or DLD. Second, despite not having an official diagnosis, at the time of testing, all children were undergoing assessments and were in the process of receiving a diagnosis. The fact that children with autism and DLD have been shown to have overlapping difficulties with oral narrative and pragmatic abilities (e.g., Norbury et al., 2014; Norbury & Bishop, 2003) and that there is recent evidence suggesting that oral language interventions have similar aims and effects in these groups of children (see Donolato et al., 2023) also contributed to following a transdiagnostic approach. Finally, we follow CATALISE's (Bishop et al., 2016, 2017) approach to prioritize more the needs children have (i.e., functional diagnosis) rather than a diagnostic label. For this, we ensured that all participating children had a professional report describing discourse-level linguistic difficulties. Their usual speech-language therapists also confirmed that a narrative intervention was appropriate for them. This applied even if children had not previously been exposed to any narrative intervention either at the school level or at the individual speech-therapy level.

Before starting the intervention, we evaluated participants' core language abilities (either CELF Preschool-2 or CELF-5 core language score, depending on the child's age; Wiig et al., 2009, 2013), pragmatic abilities (PleaseApp; Andrés-Roqueta et al., 2024), and non-verbal IQ (K-BIT Matrices; Kaufman & Kaufman, 1990). Pairwise *t*-tests were used to compare groups (see Table 2 for all the statistical results, with Cohen's *d* as a measure of effect size). Interestingly, no significant differences in core language, receptive pragmatics and non-verbal IQ were found between the participants with autism or DLD (CELF: $p = .070$; PleaseApp: $p = .129$; K-BIT: $p = .987$).

## Design

In this study, we used a non-randomized controlled trial with a between-subjects experimental design involving three different groups, which included an experimental group that received the MultiModal Narrative intervention and two control groups. Figure 1 shows a schematic representation of the study design and the intervention protocol.

Recruitment and pre-intervention assessments were completed before allocation to each group. Participants were not randomly assigned to the groups, as allocation was based on the willingness and availability of the child's speech-language therapist to implement the intervention sessions, together with families' acceptance of changing their child's usual intervention practice to the trial presented in this paper.

## The MultiModal Narrative Intervention

The intensive tier of the MultiModal Narrative (MMN) intervention was implemented in this study. The MMN is a multi-tiered narrative-based intervention for children with TD or NDD that trains both narrative macrostructure and perspective-taking through a set of educational strategies. These strategies include interactive and multimodal enactment strategies using Catalan, the vehicular language in schools in Catalonia. It was co-created with 93 teachers and speech-language therapists working in the Catalan educational and clinical system, with the aim of being an evidence-based intervention program (see Florit-Pons et al., 2025b). An initial pilot study showed that a larger implementation of the intervention was feasible (see Florit-Pons at al., 2025a).

The intensive tier (i.e., designed for children with NDD) of the MMN involves nine intervention sessions dedicated to training oral narrative skills through story retelling and story generation activities. The intervention centers on three stories about a capybara, such that each story is trained during three consecutive sessions followed by a step-by-step protocol (see Supplementary Material 3 for a schematic description of the procedure). All intervention sessions start with the same four initial activities. First, the interventionist briefly introduces the activity and the story. Next, the interventionist and child watch audiovisual cartoons together on a laptop, followed by a video of a storyteller retelling and

**Table 2** Summary of the t-test results and Cohen's d by Group

| Variable | Experimental NDD vs. Control TD | Experimental NDD vs. Control NDD | Control TD vs. Control NDD |
|---|---|---|---|
| Age | $t(16.83) = -1.00$, $p = .330$, $d = 0.36$ | $t(28.90) = 1.50$, $p = .144$, $d = 0.53$ | $t(19.15) = 3.57$, $p = .002$, $d = 1.23$ |
| TSI | $t(15.00) = 0.143$, $p = .888$, $d = 0.05$ | $t(28.13) = 1.96$, $p = .059$, $d = 0.69$ | $t(16.00) = 3.01$, $p = .007$, $d = 1.06$ |
| CELF | $t(29.90) = -2.60$, $p = .014$, $d = -0.91$ | $t(30.92) = -0.31$, $p = .759$, $d = -0.11$ | $t(30.25) = 2.14$, $p = .040$, $d = 0.74$ |
| PleaseApp | $t(30.27) = -2.42$, $p = .022$, $d = -0.85$ | $t(26.72) = 2.03$, $p = .052$, $d = 0.72$ | $t(30.03) = 5.11$, $p < .001$, $d = 1.75$ |
| K-BIT | $t(30.79) = -0.32$, $p = .754$, $d = -0.11$ | $t(30.72) = -0.02$, $p = .986$, $d = -0.01$ | $t(31.99) = 0.31$, $p = .760$, $d = 0.11$ |

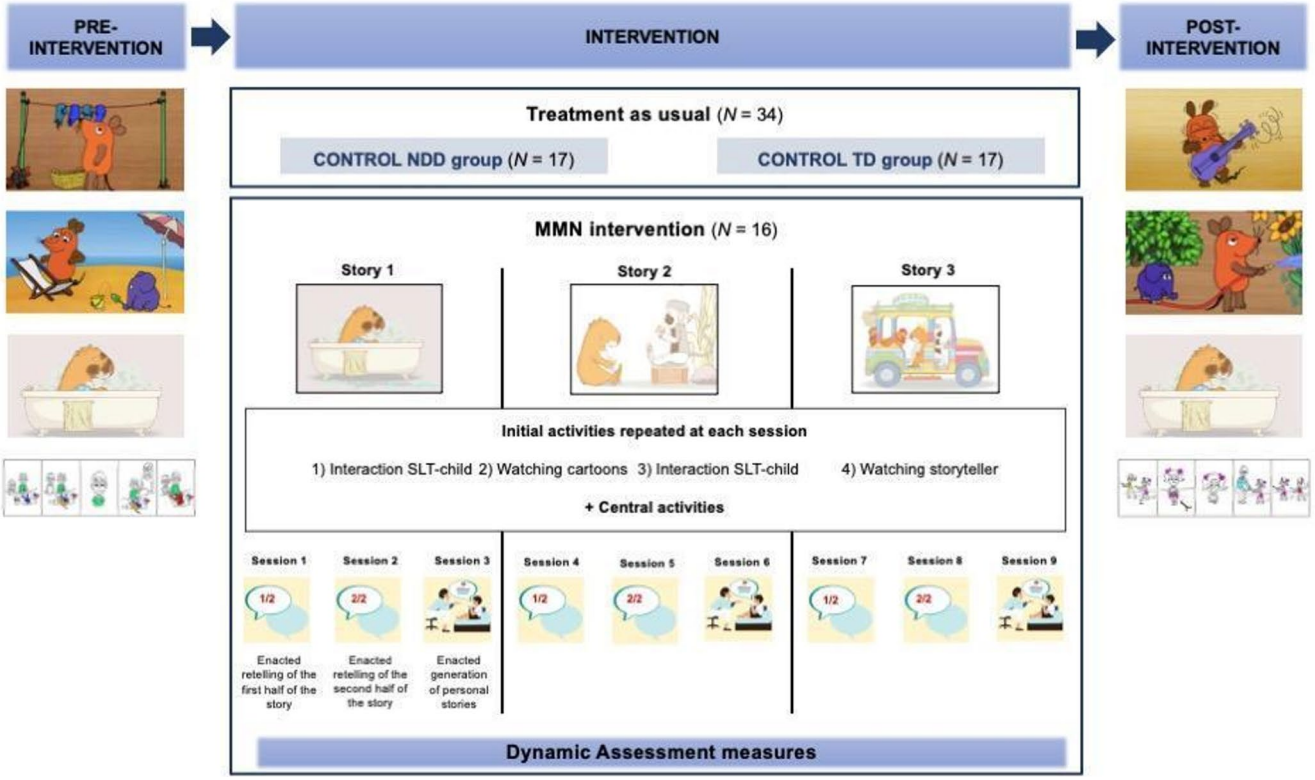


**Fig. 1** Study design and schematic representation of the study design and intervention protocol

enacting the story. During the video, the storyteller uses different gestures, facial expressions and bodily movements. A central interaction-based activity follows these four initial activities where the child actively engages with the speech-language therapist. The first and second sessions' central activity involves an enacted retelling of the story using a question-and-answer sequence about the different macrostructure and perspective-taking elements. In the first sequence, the interventionist gives a first oral model of the question-answer sequence while enacting the target element using representational gestures and facial expressions. After this, the sequence is repeated so that the child and therapist can retell the story together: the therapist asks the target question, the child answers, and then they both multimodally enact each element at the same time. The difference between the first two sessions is that the story is split into two: the first session focuses on the first half of the story, whereas the second session targets the second half of the story. The central activity of the third session of each story consists of a personal story generation that is related to the theme of each story, so that the child can relate what happened during the fictional stories to their own personal experiences. First, the interventionist produces and enacts a model personal story, followed by a question-and-answer sequence; second, the interventionist asks for the child's own personal story generation and enactment. This three-session procedure is repeated for the three stories. After each intervention session, the protocol also includes the administration of two short narrative tasks to assess children's learning process throughout the intervention period. The speech-language therapist administers two short DA measures (see *Session-by-session Dynamic Assessment outcome measures* subsection). Children received one weekly intervention session throughout nine weeks, for a total of nine intervention sessions.

For it to be as natural as possible for children, the intervention was implemented by the child's usual speech-language therapist in their usual intervention context. For this, all interventionists received a 6-hour training on how to implement the intervention together with specific training on multimodal enactment (e.g., different types of gestures and their value within narrative instruction). Although during the intervention sessions they had scripted instructions on when to enact the main actions and emotions and when to encourage the child to enact them, they were always advised to enact the stories as naturally as possible.

Each interventionist implemented the intervention with at least one child, although some implemented it with more than one child ($M=2$; $SD=1.2$; *range* = 1–4). After each intervention session, they were asked to complete a two-item treatment fidelity checklist indicating the duration of the session (i.e., *Please indicate in numbers the total exact*

*duration of the intervention session*) and whether they had followed the intervention protocol in its entirety using a yes-no format (i.e., *Overall*, *did you follow the intervention protocol?*). The average duration of the sessions was 24.53 min ($SD = 9.89$), and on 97.32% of the occasions, they affirmed that they had followed the intervention. An additional treatment fidelity assessment was conducted by an undergraduate research assistant and either the first and third authors. Provided that each interventionist was asked to video-record a minimum of three intervention sessions, 30% of the sessions (i.e., 48 sessions) were evaluated using a 1–7 Likert scale that asked the degree to which the professional (1) followed the intervention procedure and (2) used the recommended strategies of the intervention (i.e., (a) verbal modeling, (b) feedback, (c) audiovisual support, and (d) story enactment). This external evaluation showed that interventionists followed the intervention protocol ($M = 6.83$, $SD = 0.56$), and used the set of recommended strategies ($M = 6.41$, $SD = 1.1$). Inter-rater reliability was calculated for these external fidelity assessments using Gwet's Agreement Coefficient 2 (AC2, Gwet, 2012) and showed almost perfect agreement between the coders (following the steps: AC2 = 0.916 [95% CI, 0.841 to 0.991], $p < .001$; using the recommended strategies: AC2 = 0.871 [95% CI, 0.814 to 0.928], $p < .001$).

## Treatment as Usual

Children in the control groups did not receive the MMN intervention. Children in the control NDD group continued with their usual individualized intervention sessions at the speech therapy level, whereas children in the control TD group were receiving only regular school instruction. Children's usual speech-language therapists and teachers in these control groups reported that during the target intervention weeks no oral narrative activities were conducted. Rather, they worked on other abilities: (a) learning vocabulary, establishing routines, and initiating mathematical thinking at the school level (control TD), and (b) phonological awareness, syntactic structure and oral conversation at the speech-therapy level (control NDD).

## Outcome Measures

### Pre- and Post-intervention Outcome Measures

Children's oral narrative skills were evaluated before and after the intervention in all groups using a narrative retelling task involving four stories. Two of the four stories were short (~ 50 *s*) wordless video cartoons about a mouse and his elephant friend from the German series "Die Sendung mit der Maus". The third story corresponded to the first story trained during the intervention, which was also a wordless cartoon (~ 2.30 *min*). The fourth story was a comic-like sequence of five images from the CUBED Assessment (Petersen & Spencer, 2016). The three untrained stories (first, second, and fourth) were changed from pre-intervention to post-intervention to ensure that it was the first time the child saw that story, while the trained story was maintained post-intervention.

The procedure was as follows: the child individually watched the cartoon, and the experimenter then asked him/her to retell the story without having access to any visual material of the story. This process was repeated for the next two wordless cartoons. As for the comic-like sequence, the experimenter first told the story while referring to each picture, and then asked the child to retell it while looking at the set of pictures. Children were tested individually in a silent room at their school or speech therapy center. The task was administered and coded by a research assistant for TD children and by the first author and a certified psychologist for NDD children. Children were tested pre-intervention (around three and four weeks prior to the beginning of the intervention phase and post-intervention (around one and two weeks after the end of the intervention phase).

Each narrative retelling was coded for narrative macrostructure and narrative perspective using two different 0–6 codings. The macrostructure coding was adapted from Demir et al.'s (2014) coding for evaluating the child's introduction of the different macrostructural elements in the retelling. Children's narrations were scored a '0' if there was no descriptive sequence, a '1' if there was one descriptive sequence but no temporal sequence, a '2' if there was an action sequence (e.g., the main character and a problem), a '3' if it lacked two or more macrostructure elements (character, problem, attempt, solution, final), a '4' if it lacked one macrostructure element, a '5' if all macrostructure elements were included, and '6' if all macrostructural elements were included together with details about the story. Narrative perspective-taking coding was adapted from Dodd et al. (2011), considering whether the child introduced emotional terms, causal relations of emotions, and mental terms. Narrative perspective-taking was scored a '0' if no emotional or mental terms were included, a '1' if one emotion was included, a '2' if two or more emotions were included, a '3' if one emotion and its cause was included, and a '4' if two or more emotions and the cause of at least two emotions were included. Additional scores of '+ 1' and '+ 2' were added if the retelling included one or two mental terms (e.g., *thinking*, *realizing*, *willing*, *wanting*).

Inter-rater reliability was calculated for the narrative coding of the retelling task administered pre- and post-intervention with data from 16 participants (8 TD and 8 NDD), corresponding to 32% of the data. An undergraduate

speech-language therapy student received 45 min of training on how to code participants' responses, and then annotated the data. Cohen's kappa (weighted kappa) was used to calculate inter-rater reliability. The results (*N* of responses = 128) showed high agreement for macrostructure (κ = 0.86) and perspective-taking (κ = 0.70).

### Session-by-session Dynamic Assessment Outcome Measures

To assess children's ability to learn narrative macrostructure and perspective-taking throughout the intervention, two DA measures were incorporated at the end of each intervention session: a narrative retelling task and a set of comprehension questions that incorporated graduated prompting. Graduated prompting is a scaffolded approach used to support students' learning, providing further guidance on how to respond to the task. DA measures were administered and coded by the same interventionist at the end of each intervention session. Interventionists received 1-hour training before the intervention period on how to implement and code the DA measures.

Figure 2 shows a visual representation of the procedure used by the practitioner to obtain the two DA measures. As for the first measure, the child's narrative retelling was obtained using a set of supporting images of the trained story. The interventionist coded the child's retelling for narrative macrostructure and perspective-taking using the coding criteria described in the *Pre- and post-intervention outcome measures* subsection. Provided that during the first intervention session for each story only the first half of the story was trained, the macrostructure score of the narrative retelling in sessions 1, 4 and 7 could only be a maximum of 4 (as a score of 5 or 6 implied that the child introduced all the elements). After the child's retelling, the interventionist asked a set of explicit graduated comprehension questions about the main macrostructure and perspective-taking elements of the trained story with different graduated support prompts (see Supplementary Material 4 for a list of the comprehension questions asked for each story). She first asked an explicit open question to the child. If the child answered correctly, the response was coded as 'correct' and they moved on to the next question. If the child answered incorrectly or did not answer, the response was coded as 'incorrect', and the interventionist would give more support prompts for the child to answer correctly, namely a two-choice question or by showing images representing the two choices.

## Statistical Analyses

Statistical analyses were performed using R software. First, to calculate whether there were improvements from pre- to post-intervention, four Linear Mixed-Effects (LME) models were performed using the *lme4* package (Bates et al., 2015). First, two models were run to calculate the average

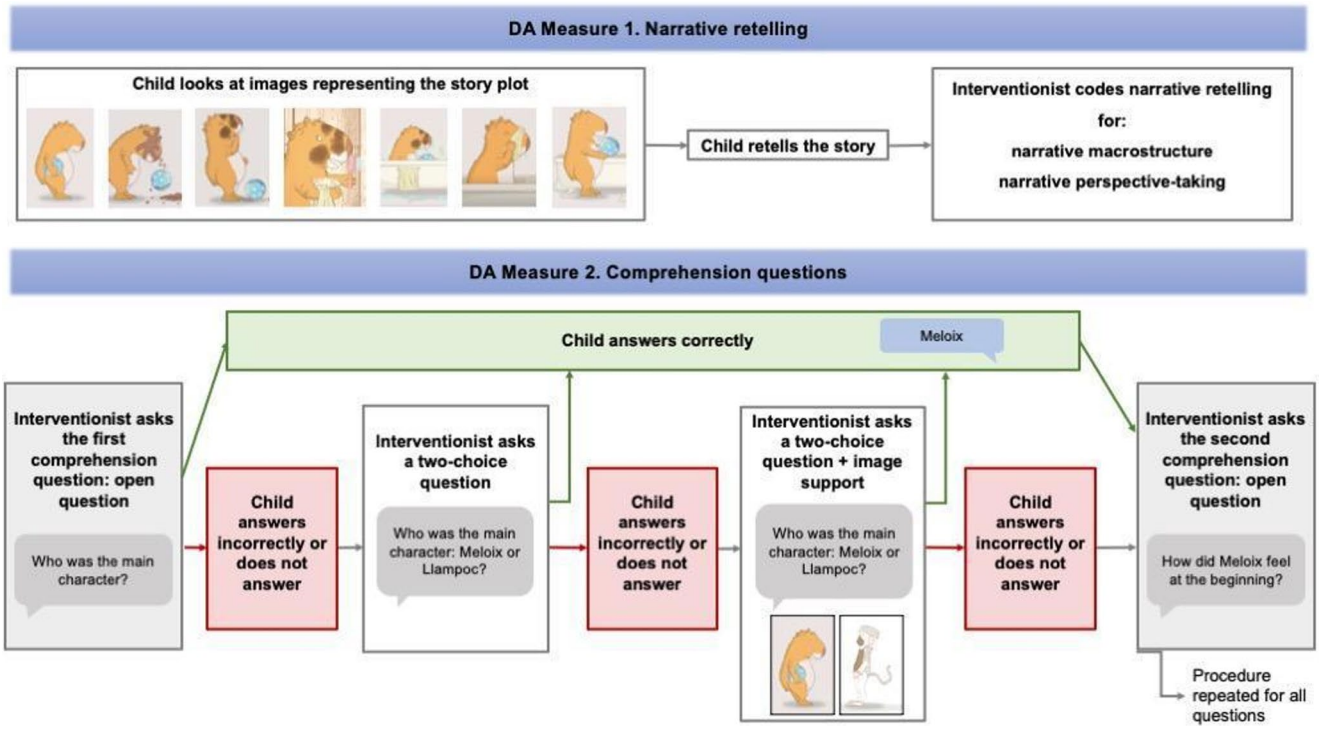


**Fig. 2** Visual representation of the procedure used to obtain the dynamic assessment measures

scores for narrative macrostructure and narrative perspective-taking for the three untrained stories. Second, two additional models were run using the scores corresponding to the trained story. The models used the narrative macrostructure score or the narrative perspective-taking score as the dependent variable. Two fixed factors were included in the models: Test (2 levels: pre-intervention and post-intervention) and Group (3 levels: Experimental NDD, Control NDD and Control TD), as well as their two-way interaction. The random-effects structure included by-Participant varying intercepts, which was determined using the *performance* package (Lüdecke et al., 2021) in R. Finally, post-hoc pairwise comparisons were run for all significant main effects and interactions using Bonferroni correction with the *emmeans* package (Lenth, 2021), together with a measure of effect size (via Cohen's *d*).

As for the session-by-session DA measures, we first ran two repeated-measures analysis of variance (ANOVA) models to examine whether there was an improvement throughout the intervention sessions. The dependent variable was the score for narrative macrostructure and narrative perspective-taking that each child received in each session. Session was included as a fixed factor, with six levels for narrative macrostructure score (Session 2, Session 3, Session 5, Session 6, Session 8 and Session 9) and nine levels for narrative perspective-taking (Sessions 1–9). Narrative macrostructure included only six levels, as the first session of each cartoon was excluded (see *Session-by-session Dynamic Assessment outcome measures* for details). Participant was included as a random factor. Post-hoc pairwise comparisons were performed using the *rstatix* package (Kassambara, 2023). Additionally, two separate LME models were used to evaluate children's ability to answer comprehension questions about the story and the support prompts they needed to answer. The first model assessing children's answers to the open questions included Score (the average number of correct and incorrect responses) as the dependent variable and Session (9 levels: Sessions 1–9), Response Type (2 levels: correct, incorrect), and Narrative Measure (2 levels: macrostructure, perspective-taking) as fixed factors. The second model assessing the support prompts needed to answer correctly included Score (the average number of support prompts) as the dependent variable and Response (4 levels: open question, two-choice question, two-choice question + image support, no correct response reached) as a fixed factor. Both models included by-Participant varying intercepts. The *emmeans* package was used to calculate post-hoc comparisons together with Cohen's *d* for the effect size.

Finally, to determine whether the support prompts that children needed could be significant predictors of later performance, a linear regression model was run. The linear regression model included the Narrative Score (i.e., the average narrative macrostructure score for untrained stories at post-test) as the dependent variable and Support Prompts (i.e., the average number of support prompts needed throughout the sessions) as a predictor.

## Results

### Gains in Narrative Macrostructure and Narrative Perspective-taking Comparing Between Pre- and Post-intervention

For the descriptives of all the narrative macrostructure and narrative perspective-taking measures at pre- and post-intervention, see Table 3. The model for narrative macrostructure skills for untrained stories showed a main effect of Group ($\chi^2(2) = 30.19$, $p < .001$), which indicated that, regardless of testing time, the Control TD and the Experimental group had significantly higher scores than the Control NDD group. The

**Table 3** Descriptives (M, SD and Range) for pre- and post-intervention assessments

| | | | Control TD | | Control NDD | | Experimental NDD | |
|---|---|---|---|---|---|---|---|---|
| Variable | Story | Measure | PRE | POST | PRE | POST | PRE | POST |
| Narrative macrostructure | Trained story | *M* | 4.35 | 3.71 | 1.82 | 2.76 | 2.63 | 4.00 |
| | | *SD* | 0.70 | 0.92 | 1.78 | 1.35 | 1.02 | 1.32 |
| | | *Range* | 3–5 | 2–6 | 0–5 | 0–5 | 1–4 | 1–6 |
| | Untrained stories | *M* | 4.29 | 4.04 | 2.22 | 2.29 | 2.81 | 3.77 |
| | | *SD* | 1.14 | 0.63 | 1.41 | 1.31 | 1.09 | 1.19 |
| | | *Range* | 0.67–5.67 | 3–5.33 | 0–4.67 | 0–4.33 | 1–4 | 1.67–5.67 |
| Narrative perspective-taking | Trained story | *M* | 0 | 0 | 0.24 | 0.06 | 0 | 0.31 |
| | | *SD* | 0 | 0 | 0.75 | 0.24 | 0 | 0.60 |
| | | *Range* | – | – | 0–3 | 0–1 | – | 0–2 |
| | Untrained stories | *M* | 0.39 | 0.06 | 0.29 | 0.02 | 0.26 | 0.23 |
| | | *SD* | 0.40 | 0.13 | 0.44 | 0.08 | 0.29 | 0.23 |
| | | *Range* | 0–1.33 | 0–0.33 | 0–1.67 | 0–0.33 | 0–0.67 | 0–0.67 |

See Supplementary Material 5 for this table separating participants with autism and DLD

**Table 4** Results of significant post-hoc comparisons for narrative macrostructure (untrained stories)

| Effect | Comparison | Estimate | SE | df | t | *p* | d |
|---|---|---|---|---|---|---|---|
| Group | Control TD > Control NDD | 1.91 | 0.35 | 47 | –5.49 | < .001 | –2.46 |
| | Experimental > Control NDD | –1.04 | 0.35 | 47 | –2.93 | .016 | –1.33 |
| Test * Group | Pre-intervention: Control TD > Control NDD | –2.08 | 0.40 | 72.4 | –5.25 | < .001 | –2.67 |
| | Pre-intervention: Control TD > Experimental | 1.48 | 0.40 | 72.4 | 3.68 | .004 | 1.91 |
| | Post-intervention: Control TD > Control NDD | –1.75 | 0.40 | 72.4 | –4.41 | < .001 | –2.24 |
| | Post-intervention: Experimental > Control NDD | –1.48 | 0.40 | 72.4 | –3.67 | .004 | –1.90 |
| | Experimental: Post-intervention > pre-intervention | 0.96 | 0.28 | 47 | 3.49 | .01 | 1.23 |

For each comparison, ">" indicates that the first element of the comparison had higher scores (better performance) than the second

interaction between Test and Group was also found to be significant ($\chi^2(2) = 10.63$, $p < .005$). Post-hoc comparisons showed first that at pre-intervention the Control TD group had significantly higher scores than the Control NDD and the Experimental groups. While the significant difference between the two control groups was maintained at post-intervention, we observed that the Experimental group significantly improved from pre-intervention to post-intervention and outperformed the Control NDD group at post-intervention. See Table 4 for the statistical report of all post-hoc comparisons and Fig. 3 for a visual representation of the results. As for narrative perspective-taking, results only showed a main effect of Test ($\chi^2(1) = 15.14$, $p < .001$), indicating that scores were significantly higher at pre-intervention than at post-intervention (see Table 5 for the results), regardless of Group. No significant effects were found for Group ($p = .468$) and for the two-way interaction ($p = .067$).

Similar findings were found for the models with the trained story, for both narrative macrostructure and perspective-taking (see Table 3 above for the descriptive statistics). The model showed significant main effects of Group ($\chi^2(2) = 22.20$, $p < .001$) and Test ($\chi^2(2) = 10.29$, $p = .001$). A significant two-way interaction ($\chi^2(2) = 26.72$, $p < .001$) was obtained for the macrostructure model. The interaction showed that at pre-intervention, the Control TD group had significantly higher scores than the Control NDD group and the Experimental group, while no significant differences were found between the two groups with NDD children ($p = .592$). Additionally, results indicated that the Experimental group significantly improved from pre-intervention to post-intervention, and that at post-test, scores were significantly higher than those of the Control NDD group. See Table 6 for the statistical report of the post-hoc comparisons. As for narrative perspective-taking skills for the trained story, results showed a significant two-way interaction between Test and Condition ($\chi^2(2) = 6.80$, $p = .033$), but post-hoc comparisons were not statistically significant, showing no differences between the groups at the two testing times.

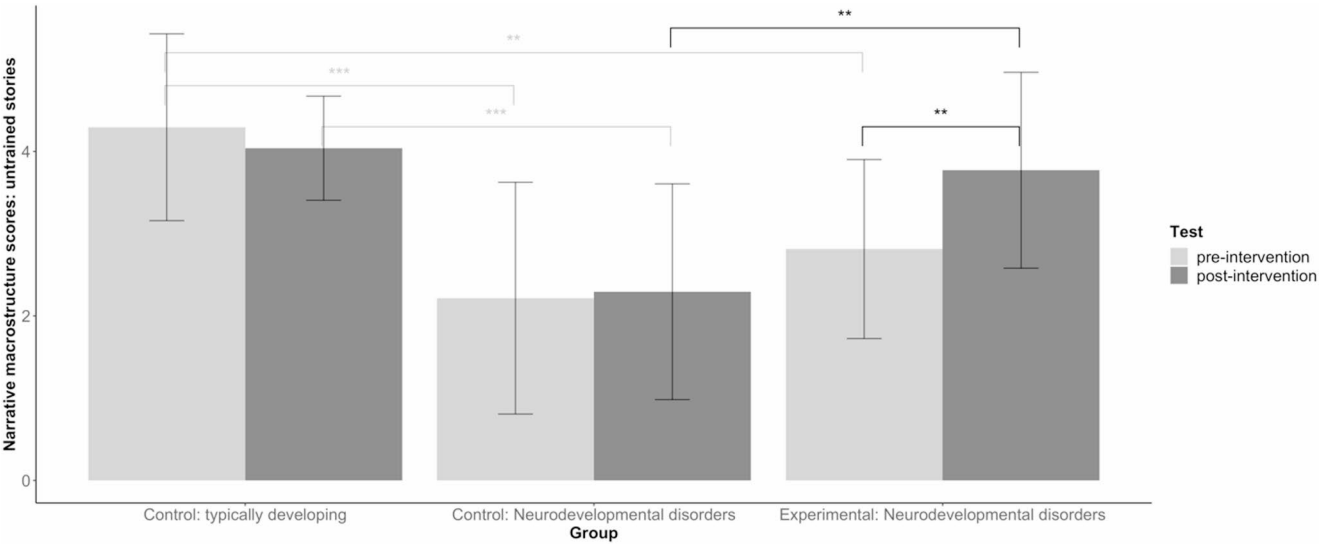


**Fig. 3** Mean narrative macrostructure scores for the untrained stories broken down by Test and Group. Note. *: $p \leq .05$; **: $p \leq .01$; ***: $p \leq .001$

**Table 5** Results of significant post-hoc comparisons for narrative perspective-taking (untrained stories)

| Effect | Comparison | Estimate | SE | df | t | *p* | d |
|---|---|---|---|---|---|---|---|
| Test | Pre-intervention > Post-intervention | –0.21 | 0.06 | 47 | –3.82 | < .001 | –0.77 |

For each comparison, ">" indicates that the first element of the comparison had higher scores (better performance) than the second

## Gains in Narrative Macrostructure and Narrative Perspective-taking Using Dynamic Assessment Measures

As mentioned before, two basic DA measures were collected at the end of each intervention session, namely a narrative retelling and a set of comprehension questions with graduated prompting. Concerning the session-by-session analysis of the DA narrative retellings, the ANOVA model evaluating narrative macrostructure showed a main effect of Session ($F(5, 70) = 2.53$, $p = .036$), showing that there was a significant improvement from Session 2 to Session 3 and from Session 2 to Session 5. No significant differences were found for the rest of the sessions, suggesting that the improvement was maintained throughout the remaining intervention sessions. The model for narrative perspective-taking also showed a main effect of Session ($F(8, 114) = 3.68$, $p < .001$), which reported significant improvements from Session 1 to Session 6, from Session 1 to Session 8, and from Session 1 to Session 9, showing that it took more time for children to acquire this skill. See Table 7 for the descriptive statistics of all sessions, Table 8 for the statistical report of significant post-hoc comparisons, and Fig. 4 for a representation of these improvements.

Second, regarding the analysis of the comprehension prompting questions that defined the support needed by the children, the first LME model evaluating their ability to answer open questions correctly or incorrectly showed a significant main effect of Response Type ($\chi^2(1) = 30.92$, $p < .001$), indicating that overall there were significantly more correct answers (60.06%) than incorrect answers (39.94%). No main effects were found for Session ($p = 1$)

**Table 6** Results of significant post-hoc comparisons for narrative macrostructure (trained story)

| Effect | Comparison | Estimate | SE | df | t | *p* | d |
|---|---|---|---|---|---|---|---|
| Group | Control TD > Control NDD | –1.74 | 0.37 | 47 | –4.69 | < .001 | –2.06 |
| | Experimental > Control NDD | –1.02 | 0.38 | 47 | –2.71 | .028 | –1.21 |
| Test | Post-intervention > pre-intervention | 0.56 | 0.17 | 47 | 3.30 | .002 | 0.66 |
| Test * Group | Pre-intervention: Control TD > Control NDD | –2.53 | 0.43 | 73.2 | –5.99 | < .001 | –3.01 |
| | Pre-intervention: Control TD > Experimental | 1.73 | 0.43 | 73.2 | 4.03 | .001 | 2.05 |
| | Post-intervention: Experimental > Control NDD | –1.24 | 0.43 | 73.2 | –2.88 | .047 | –1.47 |
| | Experimental: Post-intervention > pre-intervention | 1.38 | 0.30 | 47 | 4.62 | < .001 | 1.63 |

For each comparison, ">" indicates that the first element of the comparison had higher scores (better performance) than the second

**Table 7** Descriptives (M, SD and Range) for the session-by-session narrative retelling

| | | Session | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Narrative macrostructure | *M* | – | 3.33 | 4.13 | – | 4.44 | 4.40 | – | 4.19 | 3.86 |
| | *SD* | – | 1.40 | 1.60 | – | 1.41 | 1.92 | – | 1.72 | 1.99 |
| | *Range* | – | 0–5 | 0–6 | – | 2–6 | 1–6 | – | 1–6 | 1–6 |
| Narrative perspective-taking | *M* | 0.56 | 1.07 | 0.93 | 1.20 | 1.06 | 1.53 | 0.80 | 2.00 | 1.64 |
| | *SD* | 0.51 | 1.33 | 1.16 | 1.21 | 1.06 | 0.92 | 0.77 | 1.26 | 1.08 |
| | *Range* | 0–1 | 0–4 | 0–4 | 0–4 | 0–4 | 0–3 | 0–2 | 0–4 | 0–3 |

**Table 8** Results of significant post-hoc comparisons for the session-by-session narrative retelling

| Measure | Comparison | t statistic | df | *p* |
|---|---|---|---|---|
| Narrative macrostructure | Session 2 to Session 3 | –4.20 | 13 | .015 |
| | Session 2 to Session 5 | –4.00 | 14 | .020 |
| Narrative perspective-taking | Session 1 to Session 6 | –4.58 | 14 | .015 |
| | Session 1 to Session 8 | –4.37 | 15 | .020 |
| | Session 1 to Session 9 | –4.16 | 15 | .040 |

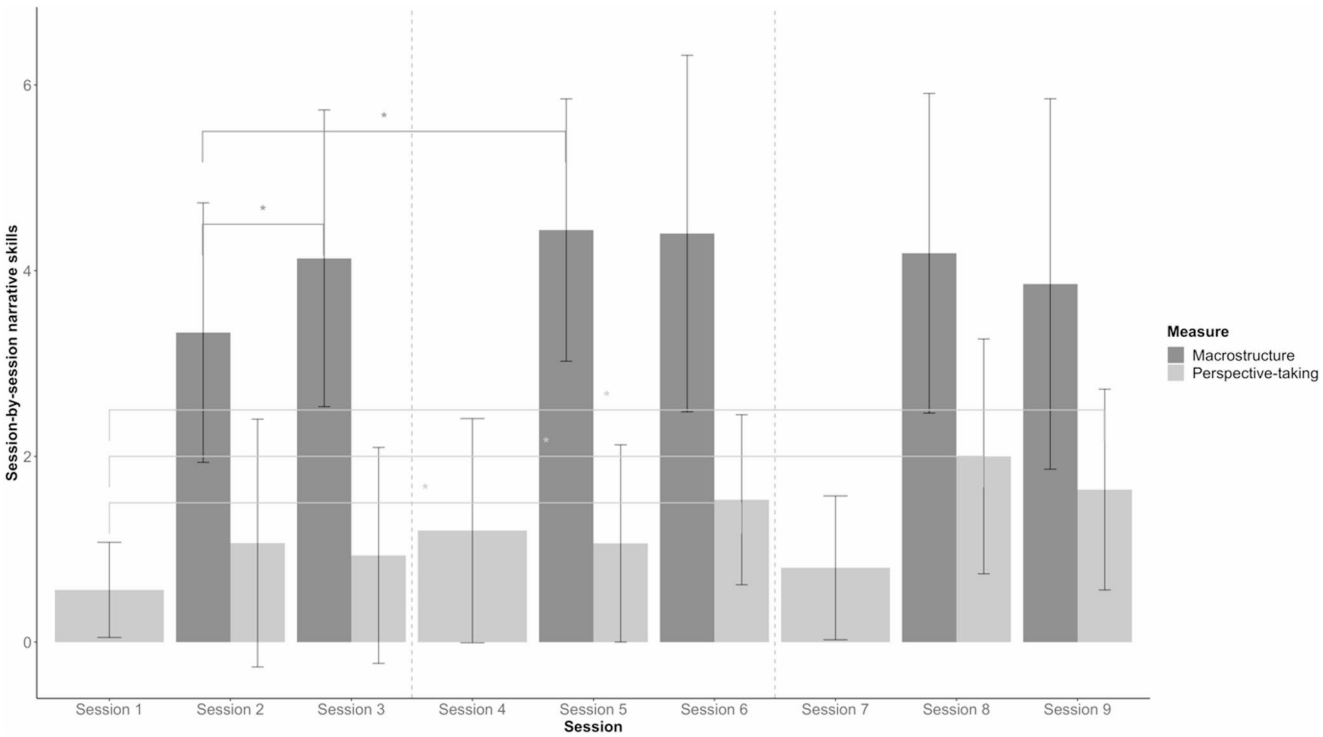


**Fig. 4** Mean narrative macrostructure and perspective-taking scores for the session-by-session narrative retelling task broken down by Narrative Measure and Session. Note. *: $p \leq .05$; **: $p \leq .01$; ***: $p \leq .001$

or Narrative Measure (i.e., macrostructure, perspective-taking) ($p = .998$). The second LME evaluating the support prompts needed to answer correctly showed a main effect of Response ($\chi^2(3) = 93.02$, $p < .001$) which showed that most children could reach a correct response after the hierarchy of support prompts, such as the open question or the two-choice question. It also showed that children could significantly answer more often at the first prompt (i.e., open question) (60.68%) than at the other graduated prompts, such as two-choice question (27.12%) or two-choice question + image support (11.97%), independently of whether it was a macrostructure or perspective-taking question, and that in very few occasions (0.23%) they could not reach a correct response. All in all, these findings suggest that although children needed different support prompts, overall they could answer all questions correctly. See Table 9 for the statistical report of significant post-hoc comparisons.

Finally, the linear regression model to assess whether the average number of graduated support prompts could predict narrative gains at post-test showed that the model accounted for 52.12% of the variance ($R^2 = 0.52$, $F(1,14) = 17.33$, $p < .001$), suggesting that the average number of support prompts children needed to answer correctly was a significant predictor of children's narrative macrostructure score on the untrained stories at the post-intervention assessment ($\beta = -0.89$, $p < .001$).

## Discussion

The current study evaluated the efficacy of the MMN intervention on a group of children with NDD. In particular, the intensive version of the intervention was used, which crucially incorporates a set of validated instructional strategies

**Table 9** Results of significant post-hoc comparisons for the graduated prompting comprehension questions

| Comparison | Estimate | SE | df | t | *p* | d |
|---|---|---|---|---|---|---|
| First step | | | | | | |
| correct > incorrect | 20.1 | 3.62 | 325 | 5.54 | < .001 | 0.58 |
| Second step | | | | | | |
| open question > no correct response | –60.4 | 6.66 | 45 | –9.08 | < .001 | –3.21 |
| two-choice question > no correct response | –26.9 | 6.66 | 45 | –4.04 | .001 | –1.43 |
| open question > two-choice question | 33.6 | 6.66 | 45 | 5.04 | < .001 | 1.78 |
| open question > two-choice question + image support | 48.7 | 6.66 | 45 | 7.32 | < .001 | 2.59 |

For each comparison, ">" indicates that the first element of the comparison had higher scores (better performance) than the second

including a strong multimodal component, and it is naturalistic (implemented by the child's usual speech-language therapist). Specifically, we aimed to assess its effectiveness by comparing children's oral narrative performance before and after the intervention through narrative retellings and during the intervention through DA measures. As a result of the intervention, we found, first, that children who received the intervention significantly improved their narrative macrostructure skills (for both the trained and the untrained stories), but not their narrative perspective-taking skills. More precisely, when looking at the session-by-session learning process, we observed that narrative macrostructure was learned during the first sessions of the intervention, while it took longer for children to learn narrative perspective-taking skills. Our results showed that even though children could answer questions about the story, around 40% of the time they needed extra support prompts, such as open questions, two-choice questions, or image support to answer the questions correctly. Crucially, the amount of support that children needed was a significant predictor of their narrative macrostructure performance in the post-intervention assessment.

## MMN Effects on Narrative Macrostructure

The findings of this study contribute to the existing narrative intervention literature by providing evidence of the effectiveness of individualized multimodal-based narrative instruction on macrostructure skills. We believe that a variety of factors have contributed to the effectiveness of the MMN intervention. First, the MMN complies with the characteristics of effective interventions described in the review by Favot and colleagues (2021), which are also in line with some of the principles of narrative intervention described by Spencer and Petersen (2020), such as the use of visual materials like story icons, modeling from the speech-language therapist, and retelling the story during each intervention session. Second, the MMN not only incorporates these characteristics but also systematically includes multimodal strategies, such as the use of story enactment to visually represent the main actions and emotions of the story, which have been found to be boost children's language skills (e.g., Dargue et al., 2021; Frey & Lüke, 2023). Finally, the MMN was co-created with a big group of professionals (see Florit-Pons et al., 2025b), which ensured that it integrated professionals' needs and real-life intervention practice, confirming its applicability. The feasibility of the MMN intervention was also evaluated in a pilot study (see Florit-Pons et al., 2025a), reporting positive outcomes by practitioners implementing it and preliminary effectiveness of the intervention, which suggested that a larger implementation was feasible. We believe that these two complementary actions might have helped tailor the intervention to the Catalan context, thus making it more effective. This co-creation process was particularly relevant, as the Catalan context is characterized by an educational system that prioritizes early literacy over oral discourse (see Regional Ministry of Education, Catalan Government, 2023), which means that speech-language therapists and teachers often lack validated resources for fostering oral narrative skills in Catalan, which is the vehicular language of schools. Overall, ensuring that interventions are developed considering the implementation context makes them sensitive and adapted to the professionals' needs as well as more likely to be implemented successfully by practitioners. Therefore, this emphasizes the need to test the effectiveness of narrative interventions in different contexts: although the core intervention principles might be transferable to different languages, the contextual factors might vary across different professional settings in different countries and regions.

## MMN Effects on Narrative Perspective-taking

Nevertheless, the effectiveness of the MMN intervention in boosting narrative perspective-taking skills is not clearcut. While we observed that children were slowly learning this skill throughout the intervention using DA, this improvement was not observed during the post-intervention assessment. First, it is important to consider that children with NDD usually have difficulties comprehending and expressing their thoughts and emotions (see Fujiki & Brinton, 2017 for a review). In addition, although children can identify and verbalize basic emotions by the ages of 3–4 in emotion identification tasks (e.g., Widen & Nelson, 2022), we believe that identifying and expressing them in a narrative discourse might be more complex. Building a narrative discourse requires multiple cognitive and linguistic processes to function simultaneously. Children must memorize and organize both the story events and characters' emotions, thoughts and perspectives. They also need to find the appropriate words to describe these events and mental terms, construct correct syntactic structures, and express them cohesively within the narrative discourse while considering multiple perspectives (e.g., García-Pérez et al., 2008). In fact, the difficulty of children with NDD to identify and express the character's emotions and perspectives within a narrative discourse has already been documented in previous research (for reviews, see e.g., Baixauli et al., 2016; Norbury et al., 2014; Winters et al., 2022). Therefore, we believe that children at these early ages (particularly children with NDD) might struggle to deal with all these processes at the same time, and for this, they avoid talking about emotions and thoughts, which are less salient components within the narrative. We believe that significant effects of the intervention on perspective-taking

skills would have been found if children were older, had received more specific instruction on narrative perspective-taking (such as the one by Dodd et al., 2011), or if the intervention had been longer (e.g., included more sessions). As suggested by the results of our DA, it seemed to take longer for children to acquire narrative perspective-taking skills.

## Dynamic Assessment as a Tool for Measuring Progress

The current study has shown that using multiple DA measures throughout the intervention helps to understand children's learning abilities. Specifically, we observed that children with NDD needed different support prompts, such as two-choice questions or visuals, to answer correctly. Our findings have shown that DA, specifically the assessment of the support prompts that each child needs, predicts children's post-intervention outcomes (in line with Camilleri & Botting, 2013), indicating that it is a relevant and valid measure for intervention practice, especially narrative intervention. Overall, the inclusion of these measures in narrative interventions is useful to test the learning capability of each child during the intervention and to adapt the intervention targets and procedure according to the child's needs. Continuously assessing the amount of support prompts children need to learn a certain linguistic skill, such as narrative macrostructure or perspective-taking, will serve to predict future outcomes and adapt the intervention accordingly.

## Limitations and Directions for Future Research

It is necessary to acknowledge some relevant limitations of the present study. While our study provides valuable insights on the preliminary effectiveness of the MMN intervention, the sample size of this study was small compared to some existing large-scale randomized controlled trials. The small sample size in this study can be attributed to several factors. First, it was difficult to recruit a larger and more homogeneous sample, mostly because the experimental group required double participation (i.e., the child and the therapist implementing the intervention). Although the initial age target was children aged four to six (which is when children's ability to narrate a story starts establishing), the double participation requirement made us have to widen the age range in order to be able to recruit enough participants. Despite this, it was still difficult to recruit a sample that contained a balanced number of children with autism and children with DLD. Second, there was also a 9.4% dropout rate (see Supplementary Material 1) for children who did not meet the inclusion criteria or abandoned the study during the intervention period for various reasons, such as recurrent illness or difficulties in arranging the post-intervention session with the family. Overall, although the sample size was determined by an a priori power analysis, we acknowledge that the sample size is small, thus restricting the generalizability of the findings to a larger population and limiting the ability to draw broad conclusions. It should also be acknowledged that the power analysis was conducted specifically for the models assessing narrative gains between pre- and post-intervention, but not for the models assessing the DA, which were only administered to the experimental group. Another limitation of this study is that inter-rater reliability could not be calculated for the DA, provided that not all of the intervention sessions were video recorded and that some interventionists stopped the recordings right before administering these measures. A final limitation of this study is that no background information (e.g., education, years of experience) was collected from the speech-language therapists implementing the intervention. Finally, our study has left some questions for future research. We believe that future studies should analyze children's narrative perspective-taking skills more deeply, for instance by distinguishing between emotional and mental terms, as well as the causal relations behind those terms. This more fine-grained analysis would serve to identify which terms children used more frequently and in which area was there an improvement.

## Clinical and Practical Implications

The current study has positive implications for speech-language pathology research and practice. First, the study has shown that adopting evidence-based interventions co-created with professionals helps ensure that the intervention is suitable for both children receiving the intervention and professionals implementing it. Second, the MMN protocol includes different educational strategies for improving oral discourse, such as verbal modeling, feedback, use of audiovisual support, and story enactment. The systematic incorporation of these strategies makes interventionists more aware of the aim and effect of each strategy and allows them to use them in other intervention contexts. In line with this, a direct and practical outcome of this study is the materialization of the MMN, readily accessible to professionals who lack scientifically validated narrative intervention materials in the Catalan language (Florit-Pons et al., 2025b). Finally, because speech-language therapists do not usually implement pre- and post-intervention assessments, as they are time-consuming, having DA can be especially helpful. The fact that the intensive implementation of the MMN intervention incorporates DA can help them quickly assess whether children are following the intervention and learning, or whether they need more support and the pace of intervention needs to be adapted.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10643-026-02190-z.

**Acknowledgements** The authors of this study would like to thank all participating families and children as well as all speech-language therapists who participated in implementing the intervention sessions (Ana Corbalán, Laia Esteban, Júlia Sais, Clara Vila, Ester Parra, Marta Solé, Maite Piqueras, Anna Llanes) and recruiting participants (Cristina Muñoz, Amparo Torras, Ana Ramos, Sívia Ferraté, Sandra Cuesta, Lourdes Rius). We also would like to acknowledge the support given by the different speech-language therapy centers (CDIAP del Maresme, CDIAP Sant Andreu de la Barca, CDIAP Balaguer, CDIAP Aspanias, Institut Guttmann, COADI) and services at schools (CREDA Tarragona - Escola Voramar, CREDA Maresme-Vallès Oriental - Escola Sinera and Escola Santa Susanna, Escola Norai) together with the participating school (Escola Milagros Consarnau). In line with this, our gratitude is also extended to all 93 speech-language therapists and preschool teachers who participated in the co-creation process of the intervention, with the support of the Regional Ministry of Education from the Catalan Government. Finally, we would also like to thank Carla Rufí, Mireia Martínez and Fernanda Pérez for their help during the data collection and Patrick Rohrer and Míriam Martínez for participating in the data reliability coding.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This research received financial support awarded by the Spanish Ministry of Science and Innovation, Agencia Estatal de Investigación and Fondo Europeo de Desarrollo Regional (PID2021-123823NB-I00); by the Generalitat de Catalunya (2021 SGR_922). Also, the first author would like to acknowledge a Ph.D. grant, awarded by the Generalitat de Catalunya (2021 FI_B 00778) and by the Department of Translation and Language Sciences (Universitat Pompeu Fabra).

**Data Availability** The data that support the findings of this study are openly available in Open Science Framework at https://osf.io/3qbmz.

## Declarations

**Ethics Approval and Consent to Partcipate** Ethical approval was obtained from the Institutional Committee for Ethical Review of Projects at the Universitat Pompeu Fabra (ref.: 228) and by the Regional Ministry of Education (Catalan Government). The procedures used in this study adhere to the tenets of the Declaration of Helsinki. Ethical approval was also obtained from specific speech therapy centers when required by the institution.
Legal guardians were provided with written information about the project and then provided written consent for their child to participate in the study, including video recordings of the testing and intervention sessions. Additionally, all families of participating children consented to anonymized clinical data for use in scientific research.

**Competing interests** All authors declare that they have no competing interests.



## References

Andrés-Roqueta, C., Flores-Builes, R., & Igualada, A. (2024). Validation of PleaseApp: a digital tool for the assessment of receptive pragmatic abilities in children with neurodevelopmental disorders. *Frontiers in Psychiatry*, 15. Article 1329022. https://doi.org/10.3389/fpsyt.2024.1329022

Astle, D. E., Holmes, J., Kievit, R., & Gathercole, S. E. (2022). Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry*, *63*(4), 397–417. https://doi.org/10.1111/jcpp.13481

Babayiğit, S., Roulstone, S., & Wren, Y. (2021). Linguistic comprehension and narrative skills predict reading ability: A 9-year longitudinal study. *British Journal of Educational Psychology*, *91*, 148–168. https://doi.org/10.1111/bjep.12353

Baixauli, I., Colomer, C., Roselló, B., & Miranda, A. (2016). Narratives of children with high-functioning autism spectrum disorder: A meta-analysis. *Research in Developmental Disabilities*, *59*, 234–254. https://doi.org/10.1016/j.ridd.2016.09.007

Bamford, C. K., Masso, S., Baker, E., & Ballard, K. J. (2022). Dynamic assessment for children with communication disorders: A systematic scoping review and framework. *American Journal of Speech-Language Pathology*, *31*(4), 1878–1893. https://doi.org/10.1044/2022_ajslp-21-00349

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLOS ONE, 11*(7), e0158753. https://doi.org/10.1371/journal.pone.0158753. Article e0158753.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-2 consortium. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Camilleri, B., & Botting, N. (2013). Beyond static assessment of children's receptive vocabulary: The dynamic assessment of word learning (DAWL). *International Journal of Language & Communication Disorders*, *48*(5), 565–581. https://doi.org/10.1111/1460-6984.12033

Dargue, N., Sweller, N., & Carter, M. (2021). Short report: Learning through iconic gesture in autism spectrum disorder. *Research in Developmental Disabilities, 115*, Article 104000. https://doi.org/10.1016/j.ridd.2021.104000

Delgado-Cruz, A., Acosta-Rodríguez, V. M., & Ramírez-Santana, G. M. (2022). Intervention in the coherence of narrative discourse in students with Developmental Language Disorder and with Typical Development. *Journal for the Study of Education and Development*, *45*(4), 774–803. https://doi.org/10.1080/02103702.2022.2096279

Delgado-Cruz, A., Acosta-Rodríguez, V. M., & Ramírez-Santana, G. M. (2024). Enseñanza de esquemas narrativos en alumnado de Educación Infantil: trastorno del desarrollo del lenguaje y desarrollo típico. *Estudios sobre Educación*, *46*, 99–120. https://doi.org/10.15581/004.46.005

Demir, Ö. E., Fisher, J. A., Goldin-Meadow, S., & Levine, S. C. (2014). Narrative processing in typically developing children and children with early unilateral brain injury: Seeing gesture matters. *Developmental Psychology*, *50*(3), 815–828. https://doi.org/10.1037/a0034322

Dickinson, D. K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research & Practice*, *16*(4), 186–202.

Dodd, J. L., Ocampo, A., & Kennedy, K. S. (2011). Perspective taking through narratives: An intervention for students with ASD. *Communication Disorders Quarterly*, *33*(1), 23–33. https://doi.org/10.1177/1525740110395014

Donolato, E., Toffalini, E., Rogde, K., Nordahl-Hansen, A., Lervåg, A., Norbury, C., & Melby-Lervåg, M. (2023). Oral language interventions can improve language outcomes in children with neurodevelopmental disorders: A systematic review and meta-analysis. *Campbell systematic reviews, 19*, Article e1368. https://doi.org/10.1002/cl2.1368. Article e1368.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Favot, K., Carter, M., & Stephenson, J. (2021). The effects of oral narrative intervention on the narratives of children with language disorder: A systematic literature review. *Journal of Developmental and Physical Disabilities*, *33*(4), 489–536. https://doi.org/10.1007/s10882-020-09763-9

Fiestas, C. E., & Peña, E. D. (2018). The dynamic assessment of narratives: A bilingual study. *Journal of Cognitive Education and Psychology*, *17*(1), 97–111. https://doi.org/10.1891/1945-8959.17.1.97

Florit-Pons, J., Igualada, A., & Prieto, P. (2025a). Evaluating the feasibility and preliminaryeffectiveness of a multi-tiered multimodal narrative intervention program for preschool children. *Language, Speech, and Hearing Services in Schools, 56*(1), 17–41. https://doi.org/10.1044/2024_LSHSS-24-00048

Florit-Pons, J., Prieto, P., & Igualada, A. (2025b). Co-creation of a narrative interventionprogram for speech–language pathology and educational settings. *American Journal of Speech-Language Pathology, 34*(6), 3279–3299. https://doi.org/10.1044/2025_AJSLP-25-00099

Frey, N., & Lüke, C. (2023). Iconic gestures support novel word learning in multilingual students with SLCN in classrooms. *Child Language Teaching and Therapy*, *39*(2), 119–134. https://doi.org/10.1177/02656590231166919

Fujiki, M., & Brinton, B. (2017). Social communication intervention for children with language impairment. In R. McCauley, M. E. Fey, & R. Gillam (Eds.), *Treatment of language disorders,* 421–449. Brookes Publishing. http://www.brookespublishing.com/treatment-of-language-disorders-in-children

García-Pérez, R. M., Hobson, R. P., & Lee, A. (2008). Narrative role-taking in autism. *Journal of Autism and Developmental Disorders*, *38*, 156–168. https://doi.org/10.1007/s10803-007-0379-z

Gwet, K. L. (2012). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among multiple raters*. Advanced Analytics Press.

Hessling, A., & Schuele, C. M. (2020). Individualized narrative intervention for school-age children with specific language impairment. *Language Speech and Hearing Services in Schools*, *51*(3), 687–705. https://doi.org/10.1044/2019_LSHSS-19-00082

Kassambara, A. (2023). rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.2. https://rpkgs.datanovia.com/rstatix/

Kaufman, A. S., & Kaufman, N. L. (1990). Kaufman Brief Intelligence Test. Pearson. https://www.pearsonclinical.es/es-es/Store/Evaluaciones-Profesionales/Inteligencia-6-Neuropsicolog%C3%ADa/Test-Breve-de-Inteligencia-de-KAUFMAN/p/P100200030?srsltid=AfmBOoomHMJcDiLtBdnrSbBQtk4ssxSpxyTUhiMNznbYfv5TaImSY7nw&tab=descripcin

Lenth, R. V. (2021). emmeans: Estimated marginal means, aka least-squares means (R Package Version 1.5.5-1). https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), Article 3139. https://doi.org/10.21105/joss.03139

McGregor, K. K. (2020). How we fail children with developmental language disorder. *Language Speech and Hearing Services in Schools*, *51*(4), 981–992. https://doi.org/10.1044/2020_LSHSS-20-00003

Miller, L., Gillam, R. B., & Peña, E. (2001). *Dynamic assessment and intervention: Improving children's narrative abilities*. PRO-ED.

Norbury, C. F., & Bishop, D. V. (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders*, *38*(3), 287–313. https://doi.org/10.1080/1368231000010813

Norbury, C. F., Gemmell, T., & Paul, R. (2014). Pragmatics abilities in narrative production: A cross-disorder comparison. *Journal of Child Language*, *41*(3), 485–510. https://doi.org/10.1017/S0305000913000007X

Orellana, C. I., Wada, R., & Gillam, R. B. (2019). The use of dynamic assessment for the diagnosis of language disorders in bilingual children: A meta-analysis. *American Journal of Speech-Language Pathology*, *28*(3), 1298–1317. https://doi.org/10.1044/2019_AJSLP-18-0202

Peña, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T. (2006). Dynamic assessment of school-age children's narrative ability: an experimental investigation of classification accuracy. *Journal of Speech Language and Hearing Research*, *49*(5), 1037–1057. https://doi.org/10.1044/1092-4388(2006/074)

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech Language and Hearing Research*, *57*(6), 2208–2220. https://doi.org/10.1044/2014_JSLHR-L-13-0151

Petersen,D. B., & Spencer, T. D. (2016). *CUBED*. Language Dynamics Group. https://languagedynamicsgroup.com/cubed/

Regional Ministry of Education, Catalan Government (2023). Decret (21/2023), de 7 de febrer, d'ordenació dels ensenyaments de l'educació infantil [Decree (21/2023), of February 7, on the organization of early childhood education]. *Diari Oficial de la Generalitat de Catalunya*, *8851*, de9defebrerde2023. https://portaldogc.gencat.cat/utilsEADOP/PDF/8851/1955221.pdf

Siller, M., Swanson, M. R., Serlin, G., & George, A. (2014). Internal state language in the storybook narratives of children with and without autism spectrum disorder: Investigating relations to theory of mind abilities. *Research in Autism Spectrum Disorders*, *8*(5), 589–596. https://doi.org/10.1016/j.rasd.2014.02.002

Spencer, T. D., & Petersen, D. B. (2018). Bridging oral and written language: An oral narrative language intervention study with writing outcomes. *Language Speech and Hearing Services in Schools*, *49*(3), 569–581. https://doi.org/10.1044/2018_LSHSS-17-0030

Spencer, T. D., & Petersen, D. B. (2020). Narrative intervention: Principles to practice. *Language Speech and Hearing Services in Schools*, *51*(4), 1081–1096. https://doi.org/10.1044/2020_LSHSS-20-00015

Whalon, K., Henning, B., Jackson, E., & Intepe-Tingir, S. (2019). Effects of an adapted story grammar intervention on the listening comprehension of children with autism. *Research in Developmental Disabilities*, *95*, Article 103507. https://doi.org/10.1016/j.ridd.2019.103507

Widen, S. C., & Nelson, N. L. (2022). Differentiation and Language Acquisition in Children's Understanding of Emotion. In D. Dukes, A. C. Samson, & E. A. Walle (Eds.), *The Oxford Handbook of Emotional Development*, 174–187. Oxford Academic. https://doi.org/10.1093/oxfordhb/9780198855903.001.0001

Wiig, E. H., Secord, W. A., & Semel, E. (2009). *Clinical Evaluation of Language Fundamentals–Preschool (CELF-2)*. Pearson. https://www.pearsonassessments.com/en-us/Store/Professional-Assessments/Developmental-Early-Childhood/Clinical-Evaluation-of-Language-Fundamentals-Preschool-2-Spanish/p/100000167?srsltid=AfmBOoraghXkETw1U80c5TUuBnNAjg1E8pLlnqCdGTVDjExoHgnIwopD

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals–fifth edition (CELF-5)*. Pearson. https://www.pearsonclinical.es/es-es/Tienda/Evaluaciones-Profesionales/Lenguaje/Evaluaci%C3%B3n-Cl%C3%ADnica-de-los-Fundamentos-del-Lenguaje-5/p/P100200019?srsltid=AfmBOorEeCzDKgrg6ok_XQBeUSsY_85G3fNiYmKRVCBhLl_CFHZ6K3nm

Winters, K. L., Jasso, J., Pustejovsky, J. E., & Byrd, C. T. (2022). Investigating narrative performance in children with developmental language disorder: A systematic review and meta-analysis. *Journal of Speech Language and Hearing Research*, *65*(10), 3908–3929. https://doi.org/10.1044/2022_JSLHR-22-00017