



This postprint was originally published as:

Grolig, L., Cohrdes, C., Tiffin-Richards, S. P., & Schroeder, S. (2020). Narrative dialogic reading with wordless picture books: A cluster-randomized intervention study. *Early Childhood Research Quarterly*, 51, 191–203. <https://doi.org/10.1016/j.ecresq.2019.11.002>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, nontransferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. By using this particular document, you accept the above-stated conditions of use.

Provided by:

Max Planck Institute for Human Development - Library and Research Information
library@mpib-berlin.mpg.de

Narrative dialogic reading with wordless picture books: A cluster-randomized intervention study

Lorenz Grolig^{*1}, Caroline Cohrdes², Simon P. Tiffin-Richards³, Sascha Schroeder⁴

Max Planck Research Group Reading Education and Development (REaD), Max Planck Institute for Human Development, Berlin, Germany

^{*}Corresponding author at: Reading Education and Development (REaD), Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail address: grolig@posteo.de (L. Grolig).

Keywords:

Dialogic reading intervention, Wordless picture books, Comprehension Inference, Narrative, Follow-up

Abstract:

Shared reading has the potential to promote a wide range of language skills that are important for reading acquisition. Dialogic reading interventions in preschool facilitate the acquisition of vocabulary and narrative production skills, but it is unclear (a) whether dialogic reading can also foster inferential and literal narrative comprehension and (b) whether intervention effects are maintained until the beginning of formal reading instruction. To close these two gaps, we designed and conducted a low-dose narrative dialogic reading intervention with wordless picture books. On the child care center level, 201 German preschoolers ($M_{\text{age}} = 5;5$ years) were randomly assigned to the dialogic reading group, an alternative treatment group, or a no treatment group. Hierarchical linear models showed positive effects of dialogic reading on inferential and literal narrative comprehension and on vocabulary depth and breadth. The effect on inferential narrative comprehension was maintained five months after posttest. Overall, our findings indicate that even a small amount of narrative dialogic reading has small, albeit mostly short-term effects on narrative comprehension and vocabulary skills. We conclude that narrative dialogic reading is a promising approach for supporting the development of preschoolers' inferential skills. Long-term intervention studies are needed for the evaluation of long-term effects.

1. Introduction

Preschool children with well-developed oral language skills, such as vocabulary and narrative skills, are more likely to become proficient readers (Griffin, Hemphill, Camp, & Wolf, 2004; Lepola, Lynch, Kiuru, Laakkonen, & Niemi, 2016; Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012; Silva & Cain, 2015). A high proficiency in reading comprehension enables children to access a vast array of knowledge sources, including fiction and non-fiction books. Before they acquire reading skills, children often enjoy looking at picture books and making sense of picture sequences that form a narrative. Apart from being fun, talking about stories from books with an adult can help children to acquire oral language skills because children's books contain vocabulary and narratives that young children rarely encounter in other situations (Massaro, 2015; Montag, Jones, & Smith, 2015). Using children's books for collaborative informal learning activities (referred to as shared reading) is more effective when adults use communicative strategies that activate and scaffold children's thinking (Blewitt, Rump, Shealy, & Cook, 2009; Wasik, Hindman, & Snell, 2016; Zucker, Cabell, Justice, Pentimonti, & Kaderavek, 2013). A specific set of shared reading techniques known as dialogic reading fosters various oral language skills by encouraging children's active participation during shared reading (Lonigan, Purpura, Wilson, Walker, & Clancy-Menchetti, 2013; Whitehurst et al., 1994; Zevenbergen, Whitehurst, & Zevenbergen, 2003). In this study, we investigate the effects of a dialogic reading intervention that targets both vocabulary and narrative skills. Our study expands previous research by (a) modifying the dialogic reading intervention in the light of experimental and intervention research on shared reading effectiveness, (b) investigating intervention effects on different vocabulary and narrative skills, and (c) evaluating the maintenance of intervention gains until the beginning of formal reading instruction.

This work was supported by Forschungsfonds Kulturelle Bildung (Rat für kulturelle Bildung and Stiftung Mercator) [grant number 14-001-4]. We thank Maike Hille, Nicola Hohensee, Ann-Katrin Hosch, Iuna Hoffmann, Ye-Young Hwang, Janis Keck, Pia Linscheid, Laura Prieß, Sarah-Lena Schaefer, and Anna Walther for their assistance in conducting this study. We also thank all participating children and parents, child care centers, and teachers for supporting this study.

¹ Now at the Senate Department of Education, Youth, and Family, Berlin, Germany.

² Now at the Department of Epidemiology and Health Monitoring, Robert Koch Institute, Berlin, Germany.

³ Now at the Department of Psychology, University of Würzburg, Germany.

⁴ Now at the Department of Educational Psychology, University of Göttingen, Germany.

1.1. *Fostering oral language skills with picture books and extratextual talk*

Picture books possess special linguistic and graphic features that make them multifunctional tools for promoting oral language skills in early childhood. For example, they contain more low frequency words (and illustrations of those words) than average child-directed speech and thus facilitate the acquisition of vocabulary knowledge (Massaro, 2015; Montag et al., 2015). Also, the stories in picture books often follow an episodic narrative structure that consists of the following story grammar elements: an exposition, an initiating event, which is followed by an internal response and an internal plan, leading to a resolution attempt, and a reaction to the outcome of this attempt (Stein & Glenn, 1979). Discussing such stories facilitates the acquisition of narrative knowledge (Lever & Sénéchal, 2011; Zevenbergen et al., 2003). When adults and children look at a picture book, this allows a self-paced co-construction of meaning between children and adults during which adults can adjust their scaffolding activities, such as explanations of concepts and relationships, to the children's level of comprehension (van Kleeck, Gillam, Hamilton, & McGrath, 1997). This adaptability to children's needs renders picture books particularly suitable for learning a broad range of oral language skills.

Picture books with text have often been used in shared reading studies with typically developing children (Wasik et al., 2016). Wordless picture books, by contrast, have mostly been used in interventions targeting oral language skills of children with special needs due to language impairments, autism, or deafness (Arizpe, 2013). For typically developing pre-readers, using wordless picture books with a clear-cut, linear plot is likely to support early narrative skills because such books offer a sequence of pictures that form a narrative without relying on decoding skills (Paris & Paris, 2003). In addition, children and teachers produce more diverse words (higher lexical diversity) during shared reading with wordless picture books than with text-based storybooks (Chaparro-Moreno, Reali, & Maldonado-Carreño, 2017). By contrast, teachers produce longer sentences with text-based storybooks than with wordless picture books (Chaparro-Moreno et al., 2017), which is probably due to written sentences in storybooks being longer than spoken sentences in child-directed speech (Cameron-Faulkner & Noble, 2013; Montag, 2019). Therefore, using wordless picture books during shared reading is likely to be more effective in fostering vocabulary skills and less effective in fostering grammatical skills in comparison to using storybooks with text.

When adults ask comprehension questions about concepts and events of a story, often referred to as extratextual talk (Blewitt et al., 2009), this can significantly increase the effects of shared reading. Children learn more words during shared reading when they are engaged in extensive extratextual talk that includes literal and inferential comprehension questions (i.e., questions about explicit and implicit story information, respectively; Zucker et al., 2013) and when the cognitive demand level of comprehension questions is adequate for them (Blewitt et al., 2009; Mascareno, Snow, Deunk, & Bosker, 2016). Low demand questions aim to establish a basic understanding of the vocabulary and the story line. By contrast, high demand questions encourage children to draw inferences about vocabulary and story elements. If low demand questions are followed by high demand questions, this has a stronger effect on children's vocabulary than low demand questions alone (Blewitt et al., 2009). To target a broad range of oral language skills, the intervention in this study includes low demand questions, which target literal narrative comprehension (e.g., characters, setting, and events), and high demand questions, which target inferential narrative comprehension (e.g., inferencing thoughts, dialogs, and feelings). Previous studies have shown that using both question types in shared reading is important for the development of inferential and literal narrative comprehension skills in preschool because both explain a substantial unique proportion of reading comprehension differences in elementary school (e.g., Silva & Cain, 2015).

1.2. *Effects of dialogic reading on vocabulary and narrative skills*

Numerous studies conducted under controlled laboratory conditions have shown that extratextual talk has positive effects on children's word learning (Flack, Field, & Horst, 2018). By contrast, there are less intervention studies that investigated the effects of extratextual talk in ecologically more valid child care settings over a timespan of several weeks or months. Dialogic reading is an intervention in which adults and children engage in a conversation over picture books. One main aim of dialogic reading is to facilitate the acquisition of oral language skills in preschool because these skills predict how well children acquire reading comprehension skills in elementary school (Griffin et al., 2004; Lepola et al., 2012, 2016). Adults support children in making sense of the concepts and events in the story (Whitehurst et al., 1994), which facilitates the acquisition of vocabulary and narrative skills (Lever & Sénéchal, 2011; Whitehurst et al., 1999; Zevenbergen et al., 2003). Dialogic reading encourages the active participation of children in the shared reading situation and supports their oral language skills by using open-ended questions among other communicative strategies (Zevenbergen & Whitehurst, 2003). Previous dialogic reading studies have used children's books that contained narratives in text form and some illustrative pictures, and thus adults actually read to children during the original intervention (e.g., Whitehurst et al., 1999; Zevenbergen et al., 2003). In the present study, mostly wordless picture books were used because they facilitate children's narrative and vocabulary learning activities even more than storybooks with text (Chaparro-Moreno et al., 2017).

1.2.1. *Vocabulary breadth and vocabulary depth*

Traditionally, dialogic reading interventions have targeted, among oral language skills, vocabulary breadth and depth. Vocabulary breadth refers to the number of words a child knows. This vocabulary skill is often measured by presenting illustrations of objects and asking children to name them, thus estimating whether a child has basic (superficial) knowledge of words (e.g., What is this? This is an igloo). By contrast, vocabulary depth refers to the number of semantic features known for familiar words, that is, the essential conceptual components of a word's meaning. This vocabulary skill is often measured through definition production tasks (Ouellette, 2006), thus estimating the refined (deep) knowledge about the meaning of a word (e.g., What is a shoe? A shoe keeps feet warm and protects them).

Many studies with preschoolers found intervention effects at posttest (e.g., Lever & Sénéchal, 2011; Whitehurst et al., 1994), but only few studies have evaluated the maintenance of effects after preschool. Whitehurst et al. (1999) found a small effect on vocabulary breadth at the end of kindergarten, one year after posttest. Another study found, in contrast, that a small short-term effect on vocabulary breadth was not maintained at the beginning of first grade, three months after posttest (Sim, Berthelsen, Walker, Nicholson, & Fielding-Barnsley, 2014). Moreover, no previous study has investigated the maintenance of dialogic reading effects on vocabulary depth. Overall, there is scarce and diverging evidence regarding the maintenance of dialogic reading effects on vocabulary breadth and depth.

1.2.2. *Narrative production and narrative comprehension*

Whereas many studies have investigated intervention effects of dialogic reading on vocabulary skills (Mol, Bus, & de Jong, 2009), few have investigated intervention effects on narrative skills (Lever

& Sénéchal, 2011; van Kleeck, Vander Woude, & Hammett, 2006; Zevenbergen et al., 2003). With respect to narrative production skills, Zevenbergen et al. (2003) found that after children participated in a dialogic reading intervention, they included more inferences (internal states and dialog) in their narratives than children in a no treatment group. Similar to this result, Lever and Sénéchal (2011) found that children in a dialogic reading group produced more story grammar units (inferential and literal information combined) in their posttest narratives than children in an alternative treatment group. Both studies show that dialogic reading fosters narrative production skills, but they did not differentiate between intervention effects on inferential versus literal narrative skills or investigate the maintenance of effects after posttest.

It is important to distinguish between literal and inferential narrative skills because both are necessary for the formation of a coherent situation model of a story (Kintsch, 1998). From a pedagogical perspective, it is important to ask young children literal questions in addition to inferential questions because it encourages their active participation and facilitates inferences, which are based on literal comprehension (van Kleeck, 2008). Intervention effects on narrative comprehension skills were investigated in a study in which a small sample of children with language impairments participated in a one-on-one dialogic reading intervention with scripted comprehension questions (van Kleeck et al., 2006). The intervention resulted in significant posttest gains in literal and inferential narrative comprehension (increase of 13% for each measure) in comparison to a no treatment group (van Kleeck et al., 2006). It is unclear, however, whether dialogic reading also promotes literal and inferential narrative comprehension in a small group setting with typically developing children, and whether such effects would be maintained after posttest.

1.3. Do narrative comprehension questions facilitate the acquisition of narrative skills?

Using comprehension questions during shared reading training sessions helps young children to acquire a framework that supports their narrative comprehension and their creation of coherent mental story representations (e.g., Morrow, 1984; van den Broek, Kendeou, Lousberg, & Visser, 2017). Even though previous studies have shown that dialogic reading has a positive effect on narrative skills, less is known about how children acquire these skills during dialogic reading. As the questions used in dialogic reading activities were not specifically designed to foster narrative skills (Lever & Sénéchal, 2011; van Kleeck et al., 2006; Zevenbergen et al., 2003), children presumably acquired narrative knowledge through incidental learning when they processed the stories with the help of an instructor. By contrast, it is unclear whether using specific narrative comprehension questions that guide children through the narrative structure of stories during the dialogic reading intervention improves their narrative production skills. Two experimental studies with five-year-olds investigated the immediate effects of narrative comprehension questions on narrative production in a shared reading setting (Silva & Cain, 2017; Silva, Strasser, & Cain, 2014). Both studies found that answering narrative comprehension questions about a wordless picture book story prior to looking at and narrating the story of the same book had a positive effect on the quality of children's narratives: If children first answered narrative comprehension questions and then narrated the story, this resulted in a higher amount of inferential and literal information in children's narratives in comparison to the experimental condition in which children first narrated the story and then answered narrative comprehension questions. Thus, using narrative comprehension questions had a positive effect on the subsequent narrative production. Presumably, the questions functioned as cues for story grammar elements, which subsequently helped the children to produce these elements in their narratives. The two studies, however, do not provide conclusive evidence regarding the question whether this experimental effect is due to a questions-induced change of narrative schemata or rather due to comprehension questions momentarily guiding children's attention to relevant information (Silva & Cain, 2017; Silva et al., 2014). Therefore, the dialogic reading routine in this study includes a set of narrative comprehension questions that is used with every story book, with the aim of fostering both narrative comprehension and production skills over an extended period of time. If this repeated exposure to narrative comprehension questions results in a persistent transformation of children's narrative schemata, then children's narrative productions should include more literal and inferential information even if their narration is not preceded by guiding comprehension questions.

1.4. The present study

The current study investigates effects of a narrative dialogic reading intervention on narrative and vocabulary skills by using a cluster-randomized pre-post between-subjects design with two follow-up measurement points. This study compares oral language skills in a dialogic reading group with the development of oral language skills in a music intervention group (alternative treatment) and a no treatment group between the beginning of the final preschool year and the end of Grade 1. Based on the small to medium effect sizes reported in meta-analyses about the effects of interactive reading interventions on vocabulary learning (Mol et al., 2009; Noble et al., 2019), we expect small to medium short-term intervention effects. To account for individual differences and effects of shared reading activities outside our intervention, we control for children's storybook exposure, verbal short-term memory, and nonverbal intelligence in our analyses (Grolig, Cohrdes, Tiffin-Richards, & Schroeder, 2019; Silva & Cain, 2015). The present study focuses on three related research questions.

First, we examine whether dialogic reading with mostly wordless picture books is effective in promoting inferential and literal narrative comprehension in a small group setting with typically developing children, and whether such effects are maintained after posttest. Based on previous findings (Lever & Sénéchal, 2011; van Kleeck et al., 2006; Zevenbergen et al., 2003), we expect dialogic reading to improve inferential and literal narrative comprehension skills.

Second, we investigate whether exposure to narrative comprehension questions during the intervention also improves children's inferential and literal narrative production skills. As the intervention directly targets narrative comprehension skills, we expect that intervention effects on narrative comprehension are stronger than transfer effects on narrative production.

Third, we evaluate effects of dialogic reading on vocabulary breadth and vocabulary depth, and the maintenance of these effects until the beginning of formal reading instruction, one year after posttest. As dialogic reading with five-year-olds not only targets vocabulary breadth, but also aims to deepen the understanding of words (Zevenbergen & Whitehurst, 2003), we expect similar positive effects on both outcome measures. We have no specific expectations regarding the maintenance of effects due to mixed results in previous dialogic reading studies (Sim et al., 2014; Whitehurst et al., 1999).

2. Method

2.1. Participants and procedure

We recruited 201 children at the beginning of their last preschool year from 32 child care groups (90 girls; $M_{\text{age}} = 5$ years 5

months, $SD = 4.4$ months) that were located at 15 child care centers in Berlin, Germany. Parental consent was obtained for all participating children. Parents completed a questionnaire in which they indicated whether their child would start school in the next year. They also provided information about the socioeconomic background (ISEI-08 scores; Ganzeboom & Treiman, 1996), indicating a middle class sample, $M = 60.6$ ($SD = 16.6$). All children spoke German, and 54 of them (27%) spoke an additional language. The most frequent languages were Turkish (8), English (7), French (5), Polish (4), Russian (4), and Spanish (4).

As we aimed to investigate the maintenance of intervention effects on the development of oral language skills in Grade 1, only those children were included in the study whose parents had indicated school entry in the next year. Age-standardized pretest scores in a word definition task, a nonverbal intelligence test, and a verbal short-term memory test indicated that all children were within 2.5 standard deviations of the age norm mean for the three measures (see Section 5.3).

Children were randomly assigned on the child care center level to one of three conditions: dialogic reading, alternative treatment, or no treatment. Data were collected at four measurement points: pretest at the beginning of the final preschool year (September and October 2015); posttest at the ending of the final preschool year (May and June 2016); first-follow-up at the beginning of Grade 1 (October and November 2016); second follow-up at the ending of Grade 1 (May and June 2017). At pretest and posttest, children were individually tested at their child care center in two assessment sessions, each lasting for 40 min. Children received a small present after each testing session. At pretest, children also completed a nonverbal intelligence test in small groups of up to six children. For the first and the second follow-up in Grade 1, the children participated in a 60 min session at our test rooms. At first and second followup, children received a small present; parents received 15 Euros as compensation.

Overall, we tested 201 children at pretest, 193 children at posttest, 146 children at first follow-up, and 135 children at second follow-up. The drop out was 33% between pretest and second follow-up, which is comparable to the drop out in developmental studies that span about two years (Sénéchal & LeFevre, 2002). Between pretest and posttest, eight children (4%) left their child care centers and did not participate in any further measurement points. Between posttest and the first follow-up measurement, there was further attrition (23%). In Germany, there are no strict regulations regarding school entry at a certain age, and many parents make use of the option to keep their child in child care for an additional year. Even though all parents had indicated that their children would start school after posttest, the parents of 24 children had decided that their children would not enter Grade 1. In addition, 23 children had moved to another part of Germany or were not available for assessment. Between the first and the second follow-up measurement point, there was a further drop out of eleven children (6%) who were not available for assessment. We tested whether there were systematic differences between children who completed the study and children who dropped out. A logistic regression showed that older children were more likely to complete the study than younger children, $\chi^2(1) = 5.27$, $p < .05$. By contrast, treatment, $\chi^2(2) = 4.43$, $p = .10$, control variables, all $\chi^2(1) < 3.47$, all $ps > .06$, and vocabulary depth, $\chi^2(2) = 0.38$, $p = .54$, were not related to the likelihood of completing the study.

2.2. Treatments

The current study reports data from the MusiCo project that investigated the effects of a language and a music intervention during the last preschool year before school entry (Schroeder & Grolig, 2018). The language intervention combined dialogic reading with a phonological awareness training. In the present study, we focus on the effects of dialogic reading on narrative and vocabulary skills. In each intervention, children participated twice a week in sessions that lasted for 45 min. Each intervention program comprised a total of 42 sessions over six months.

The overall potential treatment time of the language intervention program was 31 h and 30 min (42 sessions that lasted for 45 min). The dialogic reading part lasted for 20 min in each session, with the remaining 25 min in each session being used for warmup and phonological awareness games. To maintain the children's interest in the narrative comprehension questions, they were used only in 36 of the 42 intervention sessions (three books, each being used twice during the intervention, contained target vocabulary but not narratives). Therefore, the overall potential dosage of dialogic reading was 14 h, but the overall potential exposure to narrative comprehension questions during dialogic reading was only 12 h.

The interventions were administered in small groups of four to eleven children (language intervention: $M = 8.5$, $SD = 2.1$; music intervention: $M = 7.3$, $SD = 1.5$). Interventions were administered by research assistants who had an academic background and professional experience in language pedagogy or clinical linguistics (dialogic reading) or music pedagogy (alternative treatment). All research assistants were professionally trained to work with young children and had worked in a child care setting prior to the study.

2.2.1. Dialogic reading

We selected 19 age-appropriate picture books (16 narrative books plus two wimmelbooks⁵ and one rhyming book to provide a more diverse program to the children and maintain their interest in the training; see supplemental materials, Table S1). Only the narrative books had a clear storyline and contained no or very few words. Wordless picture books were selected because they allow children to engage in narrative activities without relying on the decoding abilities of adults (Paris & Paris, 2003), which in turn should support their acquisition of narrative skills. Books were also selected on the basis that they contained some of the targeted vocabulary that was assessed by a standardized test of vocabulary depth (vocabulary subtest of the German Wechsler Preschool and Primary Scale of Intelligence; Petermann, 2009). To sustain interest in the books, each book was used in two sessions that were three months apart. All children were exposed to the same picture books and were asked the same standardized narrative comprehension questions.

The intervention used the dialogic reading principles developed by Whitehurst and colleagues (Whitehurst et al., 1994; Zevenbergen et al., 2003). In contrast to previous dialogic reading research, this intervention focused on story-specific questions that were designed to foster the inferential and literal comprehension of story elements. The story-related questions were the same type of questions that were used in the narrative task for the assessment of narrative comprehension (Paris & Paris, 2003; see Section 5.3 and supplemental materials). Half of them were high demand questions that targeted inferential story grammar elements (dialog, feelings, thoughts, prediction, theme), and half of them were low demand questions that targeted literal story grammar elements (characters, setting, initiating event, problem identification, resolution). These story-specific questions were complemented by the general communicative strategies of dialogic reading used with five-year-olds (Zevenbergen & Whitehurst, 2003), such as open-ended prompts, "wh"-questions (e.g., *what*, *where*, *why*), and distancing prompts that relate story contents to children's life experiences.

⁵ Wimmelbooks are large-format, wordless picture books that are very popular in Germany. The pages of a wimmelbook are entirely covered by small-scale scenes in which some items can be found.

In each session, these basic dialogic reading questions were used to facilitate a basic understanding of the story. Next, instructors asked the story-specific narrative comprehension questions to highlight the story elements. To standardize the intervention procedure, the manual included scripts for each book that included (a) questions about target vocabulary (labeling, definitions, relating concepts to children's experiences) and (b) the inferential and literal narrative comprehension questions about story-specific content.

2.2.2. Alternative treatment

This intervention was based on an early music education program (Nykrin, Grüner, & Widmer, 2007) and focused on tonal discrimination, rhythm repetition, and synchronization. Further information on this intervention and results regarding music abilities are reported in Cohrdes, Grolig, and Schroeder (2019).

2.3. Measures

2.3.1. Narrative skills

We adapted a narrative task by Paris and Paris (2003) that has previously been used for the assessment of literal and inferential narrative skills (Lepola et al., 2012; Silva & Cain, 2015; Silva et al., 2014). Several studies have demonstrated that there are substantial concurrent and predictive correlations of literal and inferential oral narrative comprehension scores with literal and inferential reading comprehension scores (Silva & Cain, 2015; Lepola et al., 2012; Paris & Paris, 2003), confirming that the task has a high validity with respect to early reading comprehension skills. To avoid memory effects and preserve the children's interest in the task, we used a different picture book at each measurement point (see supplemental materials, Table S2; Krahn, 1979; Mayer, 1969, 1973, 1975). All four books have been used in narrative development studies with young children (Paris & Paris, 2003; Silva & Cain, 2015; Lepola et al., 2012) and were equivalent in that they were (a) wordless picture books with black and white drawings, (b) contained about two thirds of literal main points and one third of inferential main points, and (c) had a clear-cut storyline.

The narrative task consisted of three parts: First, a picture walk during which children looked through the book and endeavored to comprehend the story; second, a narrative production task; third, narrative comprehension questions. As we were interested in the specific effects of dialogic reading, separate scores for inferential and literal narrative skills were calculated both for the production task and the comprehension task.

Narrative production. After the picture walk, children were asked to tell the story with the book as a prompt. The stories were recorded via headsets and transcribed afterwards. Children were told that a fictional character (the mascot of the study) was sick at home, and that he could hear the narration of the child via the headset, but that he could not see the pictures of the book. Thus, children were encouraged to tell an engaging, comprehensive story that did not omit details that were visually present in the book. Before coding children's data, main points of each story were empirically determined in a pilot study (Yuill & Oakhill, 1991; see supplemental materials, Tables S4–S7, for details). We collected 30 oral narratives from university students for each book, transcribed them, and determined which story grammar units (Paris & Paris, 2003; Stein & Glenn, 1979) were mentioned by 60% or more of them (see Wright, Capilouto, Srinivasan, & Fergadiotis, 2011, for methodological details).

Inferential production units included dialog, feeling, thought, and motive. Literal production units included character, setting, initiating event, problem identification, event, and resolution. These main points were used for the coding of the children narratives. In the supplemental materials, Table S2 provides an overview of the number of main points (for details, see Tables S4–S7). For each literal or inferential main point that was included in their narratives, children received one point. A second rater coded 20% of the narratives. Raters were blind to the treatment condition of the children. The intra-class correlation (*ICC2*) between first and second rater was calculated as a measure of interrater reliability. At all four measurement points, the interrater reliability was good for the inference score (pretest: *ICC2* = .97; posttest: *ICC2* = .96; first follow-up: *ICC2* = .96; second follow-up: *ICC2* = .94). For the literal score, interrater reliability was also good at each measurement point (pretest: *ICC2* = .98; posttest: *ICC2* = .95; first follow-up: *ICC2* = .96; second follow-up: *ICC2* = .94). Discrepancies between the raters were resolved through discussion.

Narrative comprehension. After completion of the narrative production task, children answered 10 narrative comprehension questions adapted from Paris and Paris (2003). For the pretest, we used an adaptation by Silva and Cain (2015); for the first follow-up, we used an adaptation by Lepola et al. (2012). For the posttest and the second follow-up, we developed two new adaptations based on Paris and Paris (2003). The questions and coding schemes are provided in the supplemental materials. Five questions targeted inferential comprehension, and thus tapped implicit information that required inferencing (dialog, feelings, thoughts, prediction, and theme). Five questions targeted literal comprehension, and thus tapped information explicitly stated in the picture book (characters, setting, initiating event, problem identification, and resolution). Answers were transcribed and scored on the basis of the coding scheme by Paris and Paris (2003). Children received one point if their answer contained the requested element (e.g., plausible thoughts). If the answer also contained an elaboration of the element (e.g., the cause of the thoughts), children received an additional point. To determine the interrater reliability (Cohen's kappa), a second rater scored 20% of children's responses. At all four measurement points, the interrater reliability was good for the inferential comprehension items (pretest: κ s = .68–.84; posttest: κ s = .72–.82; first follow-up: κ = .68–.79; second follow-up: κ s = .66–.93) and good for the literal comprehension items (pretest: κ s = .72–.95; posttest: κ s = .76–.96; first follow-up: κ s = .71–.87; second follow-up: κ s = .68–.86). Discrepancies between the raters were resolved through discussion.

2.3.2. Vocabulary

Breadth. To assess intervention effects on vocabulary breadth, we developed a picture naming task. Similar to the procedure reported by Lever and Sénéchal (2011), we selected 15 pictures from the wordless picture books that were used in the dialogic reading intervention. Each picture represented a familiar concept. Children were asked to name the objects by target words (see supplemental materials, Table S3). Target words were assumed to be unfamiliar to five-year-olds prior to the dialogic reading intervention because they were low in word frequency according to the childLex corpus (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015), normalized lemma frequency/million: $M = 7.41$ ($SD = 6.32$). If children did not produce the target word (e.g., *igloo*) but a similar word (e.g., "house"), they were asked to produce alternative labels for the object until they produced the target word or until they could not think of any more alternative labels. Children received one point for each target word and zero points for a similar word. Due to time constraints, the picture naming task was not administered at the first follow-up measurement point. The internal consistency was adequate (pretest: $\alpha = .71$; posttest: $\alpha = .76$; second follow-up: $\alpha = .75$).

Depth. Children completed a standardized test of vocabulary depth in which they were asked to explain 14 concepts (vocabulary subtest of the German Wechsler Preschool and Primary Scale of Intelligence; Petermann, 2009). Children received up to two points

for each item. Answers were scored according to the test manual. Two points were awarded for one or two essential semantic features, or at least two important, but not essential semantic features that demonstrated that the child had a comprehensive understanding of the concept. One point was awarded for an important, but not essential semantic feature, or for an appropriate example that demonstrated that the child had a proper, but not comprehensive understanding of the concept. The reported split-half reliability is good ($r = .88$).

2.3.3. Control variables

Storybook exposure. To assess the amount of storybook exposure prior to the dialogic reading intervention, children completed an auditory version of the title recognition test for preschoolers (TRT; Grolig, Cohrdes, & Schroeder, 2017), which is an adaptation of the title recognition test by Sénéchal, LeFevre, and Lawson (1996). The TRT consists of 20 real titles and 10 foils. Children indicated via computer button press whether they recognized a title. A corrected score was computed by subtracting the proportion of selected foils from the proportion of selected titles. None of the books used in the TRT was used for the vocabulary breadth task or in the dialogic reading intervention. In this sample, the corrected score varied considerably ($M = .37$, $SD = .21$). Children recognized on average about eight titles (range: zero to 20 titles). The internal consistency was acceptable ($\alpha = .68$).

Nonverbal IQ. The raw scores of the subtests Classifications, Matrices, and Completing Sequences from the Culture Fair Intelligence Test (CFT 1-R; German version: Weiß & Osterland, 2013) were added to form a single scale. Age-standardized scores were used in all analyses. In this sample, the average standardized nonverbal IQ, $M = 48.6$ ($SD = 5.7$), was similar to the mean norm score, $M = 50$ ($SD = 10$). The reported re-test reliability of the combined subtests is good ($r = .94$).

Verbal short-term memory. Children completed a standardized digit span forward test (BUEVA; Esser, 2002). Age-standardized scores were used in all analyses. In this sample, the average standardized verbal short-term memory, $M = 49.9$ ($SD = 8.3$), was similar to the mean norm score, $M = 50$ ($SD = 10$). The reported re-test reliability is good ($r = .90$).

2.4. Treatment fidelity

To ensure that the dialogic reading intervention and the alternative intervention were implemented as planned, each intervention followed a manual. The manual of the dialogic reading intervention prescribed the sequence in which the books were used and included scripted questions about target vocabulary, literal narrative comprehension, and inferential narrative comprehension. To assess treatment fidelity, instructors completed a short questionnaire directly after each of the 42 sessions. In line with early literacy research (e.g., Guo et al., 2016), we assessed adherence to the intervention program on the group level (5 = *very good*, 4 = *good*, 3 = *satisfactory*, 2 = *unsatisfactory*, 1 = *poor*), dosage (number of sessions attended by each child), and participant responsiveness on the individual level (quantity and quality of child involvement in the program: 5 = *very good*, 4 = *good*, 3 = *satisfactory*, 2 = *unsatisfactory*, 1 = *poor*). Adherence to the intervention program was good (dialogic reading: $M = 4.5$, $SD = 0.3$; alternative treatment: $M = 3.9$, $SD = 0.9$). On average, children attended 33.9 (81%) of the 42 sessions (dialogic reading: $M = 34.4$, $SD = 3.8$; alternative treatment: $M = 33.6$, $SD = 5.3$). All books used in the 42 dialogic reading sessions contained target vocabulary. By contrast, only in 36 of the 42 dialogic reading sessions, narrative comprehension questions were asked. Thus, the effective mean treatment exposure to narrative comprehension questions during dialogic reading was 9 h and 45 min whereas the effective mean treatment exposure to vocabulary during dialogic reading was 11 h and 30 min. Participant responsiveness was good regarding involvement quantity (dialogic reading: $M = 4.1$, $SD = 0.4$; alternative treatment: $M = 4.2$, $SD = 0.4$) and involvement quality (dialogic reading: $M = 4.3$, $SD = 0.3$; alternative treatment: $M = 4.1$, $SD = 0.5$).

2.5. Statistical analysis

All analyses were conducted with the software R (R Core Team, 2017). Due to technical problems or illnesses, there were a few missing data values in the oral language variables at each measurement point (pretest: 1.3%; posttest: 2.2%; first follow-up: 2.5%; second follow-up: 4.1%; see Table 1). Measurement points were nested in children, and children were nested in child care groups. We fitted 3-level hierarchical linear models (Raudenbush & Bryk, 2002) with children and child care group as random effects. The models included treatment, time, and their interaction as fixed effects. We controlled for storybook exposure, nonverbal IQ, and verbal short-term memory in all models. The categorical variables treatment and time were effect-coded. To reduce nonessential multicollinearity, the continuous variables storybook exposure, nonverbal IQ, and verbal short-term memory were centered.

Models were estimated with the lmer function of the R package lme4 (Bates, Mächler, Bolker, & Walker, 2015). Effects were tested by using the Anova function of the car package (type 3 model comparisons; Fox & Weisberg, 2011). Significant interactions between treatment and time were explored by first testing the simple main effect of treatment (Fox & Weisberg, 2011) at each measurement point. If the simple main effect of treatment was significant at a measurement point, this effect was further explored using post hoc comparisons between treatments, for which we applied the glht function of the multcomp package (Hothorn, Bretz, & Westfall, 2008) by using cell means coding and single degree-of-freedom contrasts. Effect sizes that are based on standard deviations (e.g., Cohen's d) are often not informative for practical considerations (Wasik et al., 2016). Consequently, we report absolute differences (Δ) in proportion correct estimates as effect size measure. This allows us to compare the magnitude of effects between different language measures. Finally, to assess the fit of the models, we estimated the variance explained by fixed effects (marginal R^2) and the variance explained by fixed effects and random effects (conditional R^2) as reported in Nakagawa and Schielzeth (2012) by using the r.squaredGLMM function of the MuMIn package (Bartón, 2017).

3. Results

3.1. Descriptive results

Table 1 summarizes the results of the oral language outcome variables at the four measurement points. Children's vocabulary skills increased considerably between pretest and second followup. For example, at pretest, children mentioned on average about 50% of the semantic features of target vocabulary, and they labeled on average 33% of low frequency words correctly. By contrast, at second follow-up, they mentioned on average about 70% of the semantic features of target vocabulary, and they labeled on average 66% of low frequency words correctly.

3.2. Treatment effects

We investigated the effects of the dialogic reading intervention on the development of oral language skills by fitting a hierarchical linear model for each of the six language variables. Preliminary analyses showed that all interaction terms between treatment and control variables were not significant and are therefore not reported in the final models, which are summarized in Table 2. The

Table 1

Descriptive statistics for language outcome measures by treatment and measurement point.

Treatment	Pretest			Posttest			Follow-up I			Follow-up II		
	M	(SD)	n	M	(SD)	n	M	(SD)	n	M	(SD)	n
	Inferential narrative comprehension											
Dialogic reading	.46	(.19)	68	.69	(.22)	66	.68	(.19)	57	.78	(.18)	48
Alternative treatment	.47	(.18)	64	.54	(.18)	60	.54	(.16)	39	.72	(.17)	35
No treatment	.47	(.17)	65	.54	(.23)	60	.57	(.16)	46	.78	(.15)	41
	Literal narrative comprehension											
Dialogic reading	.52	(.22)	68	.73	(.17)	65	.73	(.18)	53	.81	(.13)	50
Alternative treatment	.56	(.21)	64	.60	(.21)	61	.61	(.15)	40	.76	(.14)	35
No treatment	.49	(.23)	65	.63	(.23)	60	.63	(.18)	47	.79	(.11)	41
	Inferential narrative production											
Dialogic reading	.27	(.15)	68	.40	(.16)	66	.39	(.14)	56	.45	(.15)	53
Alternative treatment	.27	(.15)	64	.38	(.18)	62	.38	(.17)	40	.48	(.19)	35
No treatment	.23	(.16)	65	.36	(.18)	60	.38	(.16)	47	.47	(.15)	41
	Literal narrative production											
Dialogic reading	.44	(.16)	68	.53	(.16)	66	.54	(.17)	56	.69	(.14)	53
Alternative treatment	.44	(.14)	64	.52	(.16)	62	.54	(.14)	40	.64	(.13)	35
No treatment	.38	(.17)	65	.48	(.15)	60	.55	(.17)	47	.63	(.09)	41
	Vocabulary depth											
Dialogic reading	.45	(.19)	69	.53	(.18)	67	.61	(.16)	58	.68	(.15)	56
Alternative treatment	.49	(.18)	66	.46	(.16)	64	.59	(.13)	41	.68	(.13)	37
No treatment	.47	(.18)	66	.48	(.20)	61	.61	(.17)	47	.68	(.16)	41
	Vocabulary breadth											
Dialogic reading	.32	(.19)	69	.56	(.19)	67				.60	(.22)	56
Alternative treatment	.29	(.19)	66	.38	(.20)	64			N/A	.56	(.19)	38
No treatment	.29	(.20)	66	.39	(.23)	62				.54	(.19)	41

Note. Proportion correct scores are displayed. Vocabulary breadth data were not collected at follow-up I.

Table 2

Hierarchical linear models of language outcomes with treatment, time, their interaction, and control variables as fixed effects and child care group and children as random intercepts.

	df	Narrative comprehension				Narrative production				Vocabulary			
		Inferential		Literal		Inferential		Literal		Depth		Breadth	
		χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
<i>Fixed effects</i>													
Intercept	1	2362.46	<.001	4733.88	<.001	1262.93	<.001	3019.50	<.001	1547.58	<.001	756.28	<.001
<i>Experimental variables</i>													
Treatment	2	10.13	<.01	7.05	.03	0.82	.66	3.25	.20	0.16	.92	6.53	.04
Time	3	221.62	<.001	172.50	<.001	153.81	<.001	206.95	<.001	280.69	<.001	343.61	<.001
Treatment x Time	6	32.76	<.001	16.08	.01	4.00	.68	7.48	.28	16.22	<.01	38.21	<.001
<i>Control variables</i>													
Storybook exposure	1	6.01	.01	7.09	<.01	0.05	.82	8.00	<.01	12.67	<.01	42.31	<.001
Verbal STM	1	1.14	.28	1.88	.17	0.18	.67	1.02	.31	5.95	.02	38.21	<.001
Nonverbal IQ	1	4.45	.03	3.05	.08	2.15	.14	5.78	.01	5.42	.02	0.40	.53
<i>Random effects</i>													
Children	1	22.92	<.001	5.77	.02	12.63	<.01	23.52	<.001	141.55	<.001	123.72	<.001
Child care groups	1	3.02	.08	0.10	.75	6.44	.01	4.38	.04	10.11	<.01	5.84	.02
Marginal R ²		.30		.26		.18		.27		.27		.39	
Conditional R ²		.46		.34		.33		.43		.67		.76	

Note. STM, Short-term memory. Tests are based on Type III sum of squares and χ^2 values with Kenward–Roger df.

fixed effects in the models explained a substantial amount of variance in narrative and vocabulary skills (between 18% and 39%). The random effects also explained a substantial amount of variance. The variation between children within a child care group was more pronounced than the variation between children from different child care groups (see Table 2).

Fig. 1 shows the models' estimates for each treatment at the four measurement points. To analyze the effects of the dialogic reading intervention, we explored significant interactions between treatment groups and time by using post hoc contrasts. We tested whether treatment groups (dialogic reading, alternative treatment, no treatment) differed at pretest, posttest, first follow-up, and second follow-up.

3.2.1. Narrative skills

Inferential narrative comprehension. There were significant main effects of time and treatment, and the interaction between treatment and time was also significant (see Table 2 and Fig. 1a). The simple main effect of treatment was not significant at pretest, $\chi^2(2) = 0.15, p = .93$, but significant at posttest, $\chi^2(2) = 23.30, p < .001$. At posttest, dialogic reading outperformed alternative treatment ($\Delta = 16\%, t = 4.14, p < .001$) and no treatment condition ($\Delta = 16\%, t = 4.15, p < .001$). Alternative treatment and no treatment condition did not differ at posttest ($\Delta = 0\%, t = -0.10, p = .92$). At first follow-up, the simple main effect of treatment was significant, $\chi^2(2) = 13.67, p < .01$, with the same results pattern as for posttest: Dialogic reading outperformed alternative treatment ($\Delta = 15\%, t = 3.44, p < .001$) and no treatment condition ($\Delta = 11\%, t = 2.73, p < .01$). Alternative treatment and no treatment condition did not differ at first follow-up ($\Delta = -4\%, t = -0.85, p = .39$). The simple main effect of treatment was not significant at second follow-up, $\chi^2(2) = 3.56, p = .17$.

At posttest, children in the dialogic reading group made between one and two inferences more than children in the comparison

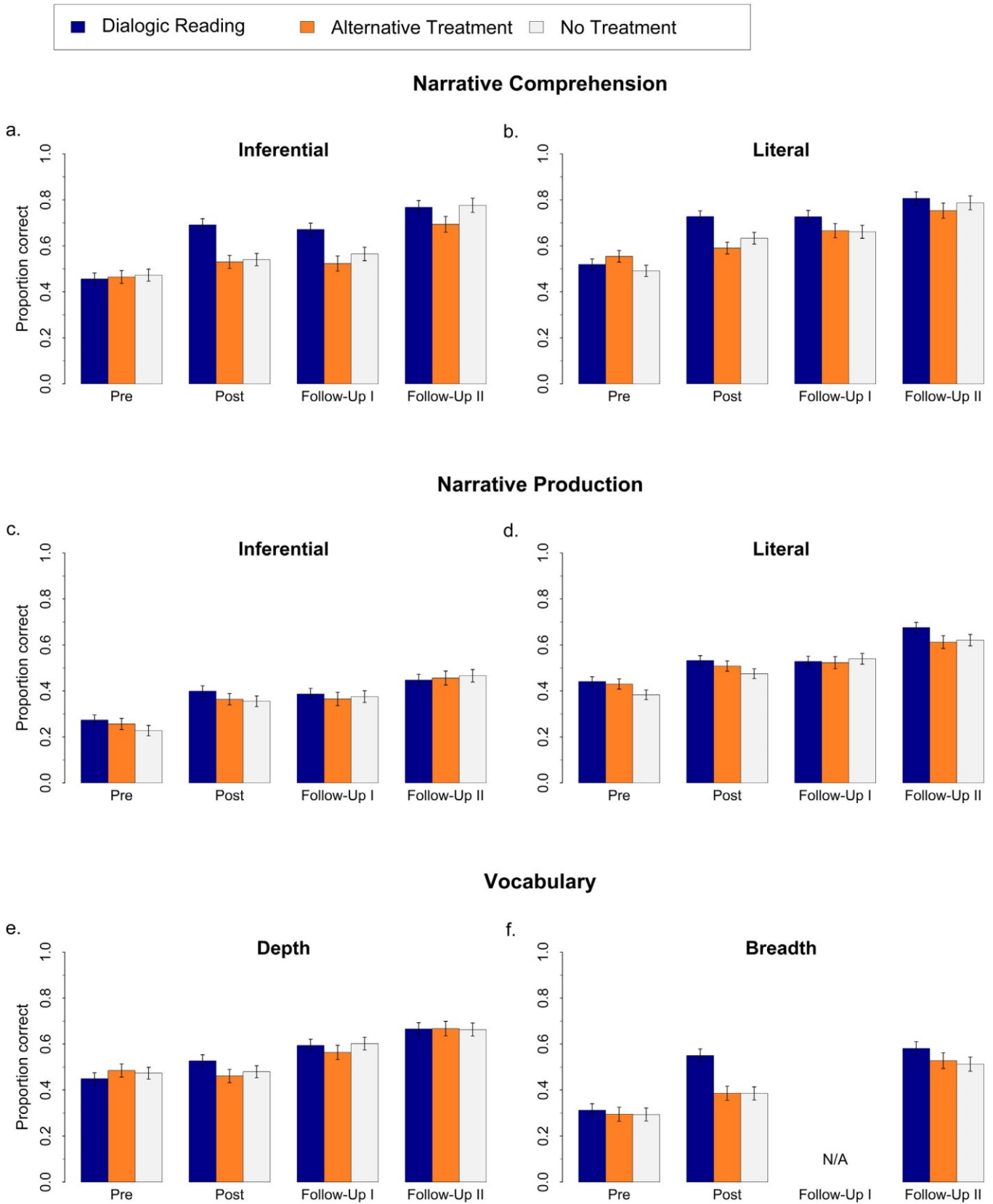


Fig. 1. Mean proportion correct estimates in narrative and vocabulary skills by treatment and measurement point. Error bars show standard errors.

groups, who made about five out of ten inferences. Between posttest and first follow-up, the dialogic reading group maintained its advantage over the comparison groups. At second follow-up, the comparison groups had caught up with the dialogic reading group. In sum, the results indicate that dialogic reading had a positive effect on inferential comprehension skills that was maintained five months after the posttest.

Literal narrative comprehension. There was a significant main effect of time, and the interaction between treatment and time was also significant (see Table 2 and Fig. 1b). Post hoc comparisons showed that the simple main effect of treatment was not significant at pretest, $\chi^2(2) = 2.73$, $p = .16$. By contrast, the simple main effect of treatment was significant at posttest, $\chi^2(2) = 14.52$, $p < .001$. At posttest, dialogic reading outperformed alternative treatment ($\Delta = 13\%$, $t = 3.69$, $p < .001$) and no treatment condition ($\Delta = 9\%$, $t = 2.62$, $p < .01$). Alternative treatment and no treatment condition did not differ at posttest ($\Delta = -4\%$, $t = -1.06$, $p = .29$). Further, the simple main effect of treatment was not significant at first follow-up, $\chi^2(2) = 3.75$, $p = .15$, and at second follow-up, $\chi^2(2) = 2.72$, $p = .26$.

At posttest, children in the dialogic reading group included on average one literal comprehension aspect more than children the comparison groups, who included on average six out of ten literal comprehension aspects. At first follow-up, the comparison groups had caught up with the dialogic reading group. In sum, the results indicate that dialogic reading had a short-term effect on literal comprehension skills.

Inferential narrative production. There was a significant main effect of time, but no significant interaction between treatment and time (see Table 2 and Fig. 1c). This indicates that dialogic reading did not have an effect on the proportion of inferential main points included in the story.

Literal narrative production. There was a significant main effect of time, but no significant interaction between treatment and time (see Table 2 and Fig. 1d). This indicates that dialogic reading did not have an effect on the proportion of literal main points included in the story.

3.2.2. Vocabulary skills

Vocabulary depth. There was a significant main effect of time and a significant interaction between treatment and time (see Table 2 and Fig. 1e). Post hoc comparisons showed that the simple main effect of treatment was not significant at pretest, $\chi^2(2) = 0.90$, $p = .64$, at posttest, $\chi^2(2) = 4.28$, $p = .12$, at first follow-up, $\chi^2(2) = 0.75$, $p = .69$, and at second follow-up, $\chi^2(2) = 0.34$, $p = .84$. Thus, none of the simple main effects of treatment were significant. As the omnibus interaction effect between treatment and time was significant and the simple main effect of treatment approached significance at posttest, we investigated the nature of this interaction by comparing the pretest–posttest change in scores between treatment groups. Dialogic reading showed a stronger growth than alternative treatment $\Delta = 10\%$, $t = 3.73$, $p < .001$) and no treatment condition ($\Delta = 7\%$, $t = 2.61$, $p < .01$). Growth did not differ between alternative treatment and no treatment condition ($\Delta = 3\%$, $t = 1.08$, $p = .28$).

Between pretest and posttest, the dialogic reading group learned on average one word definition more than the comparison groups, which produced on average between six and seven out of 14 word definitions. At first follow-up, the comparison groups had caught up with the dialogic reading group. In sum, the results indicate that dialogic reading had a short-term effect on vocabulary depth.

Vocabulary breadth. There were significant main effects of time and treatment, and the interaction between treatment and time was also significant (see Table 2 and Fig. 1f). Post hoc comparisons showed that the simple main effect of treatment was not significant at pretest, $\chi^2(2) = 1.78$, $p = .41$. By contrast, the simple main effect of treatment was significant at posttest, $\chi^2(2) = 22.25$, $p < .001$. Dialogic reading outperformed alternative treatment ($\Delta = 16\%$, $t = 3.93$, $p < .001$) and no treatment condition ($\Delta = 17\%$, $t = 4.15$, $p < .001$) at posttest. Alternative treatment and no treatment condition did not differ at posttest ($\Delta = 0\%$, $t = 0.02$, $p = .99$). At second follow-up, the simple main effect of treatment was not significant, $\chi^2(2) = 2.93$, $p = .23$.

Overall, the dialogic reading group learned low frequency words faster than the children in the comparison groups. At posttest, children in the dialogic reading group knew on average three low frequency words more than children in comparison groups, who knew on average between five and six out of 15 low frequency words. The comparison groups caught up with the dialogic reading group between posttest and second follow-up. In sum, the results indicate that dialogic reading had a short-term effect on vocabulary breadth.

3.2.3. Summary of treatment effects

Altogether, the analyses showed that dialogic reading was effective in promoting many of the targeted language skills, including inferential narrative comprehension, literal narrative comprehension, vocabulary depth, and vocabulary breadth, but not narrative production skills. The treatment groups did not differ in oral language skills at pretest, and the dialogic reading group showed a stronger growth in the narrative comprehension and vocabulary measures than both comparison groups during the last preschool year. At the beginning of grade 1, the dialogic reading group still outperformed the comparison groups in inferential narrative comprehension, indicating that the positive effect on this skill was maintained five months after the posttest.

More specifically, at posttest, the difference between the dialogic reading group and the comparison groups was slightly higher for inferential narrative comprehension (about 16%) than for literal narrative comprehension (about 11%). Moreover, the difference between the dialogic reading group and the comparison groups was higher for vocabulary breadth (about 16%) than for vocabulary depth (about 8%) at posttest. At first follow-up, the difference between dialogic reading group and comparisons groups in inferential narrative comprehension (about 13%) were similar to this difference at posttest (about 16%). Thus, dialogic reading was more effective in supporting the development of inferential comprehension skills and vocabulary breadth than literal comprehension skills and vocabulary depth.

4. Discussion

This study investigated the impact of narrative dialogic reading with mostly wordless picture books on oral language skills in comparison to an alternative treatment group and a no treatment group. Narrative dialogic reading had positive effects on inferential and literal narrative comprehension skills during the last year before school entry. The effect on inferential narrative comprehension was stronger than the effect on literal narrative comprehension. Overall, effects on narrative comprehension were stronger than the impact on narrative production skills, which was not significant. Dialogic reading also benefitted the growth of vocabulary breadth and depth during the last preschool year. The effect on vocabulary breadth was stronger than the effect on vocabulary depth. At the beginning of first grade, the comparison groups had caught up with the dialogic reading group in the narrative comprehension and vocabulary skills, with one exception: The dialogic reading group still outperformed the comparison groups in inferential narrative comprehension. The comparison groups caught up by the end of first grade. In comparison to other dialogic reading interventions (e.g., Whitehurst et al., 1994; Zevenbergen et al., 2003), which lasted for

up to 38 h over one school year, the potential treatment dosage in our study was only 14 h, and the mean effective treatment dosage was 9 h and 45 min for narrative comprehension and 11 h and 30 min for vocabulary. This intervention dose had small short-term effects on narrative comprehension and vocabulary skills.

4.1. Effects on narrative skills

Inferential and literal narrative comprehension skills both benefitted from dialogic reading. The effect on narrative comprehension skills in this study had the same magnitude as the effect in a previous study (van Kleeck et al., 2006), indicating that children with typically developing language skills benefit from this approach to the same degree as children with a language impairment. In the present study, the effect on inferential narrative comprehension was slightly stronger than the effect on literal narrative comprehension. By contrast, the previous study found that literal and inferential skills benefitted equally from dialogic reading (van Kleeck et al., 2006). This discrepancy is likely due to the amount of inferential and literal questions that were asked in the respective interventions. Our intervention used the same amount of inferential and literal comprehension questions (50% of each type), whereas, in the previous study, only 30% of the questions were inferential, and 70% were literal (van Kleeck et al., 2006). The higher proportion of inferential questions in this study's intervention plausibly explains why we found a stronger effect on inferential skills than on literal skills. Possibly, the use of wordless picture books also contributed to inferential comprehension learning by facilitating children's inferential activities in comparison to storybooks with text, which were used by van Kleeck et al. (2006).

Five months after posttest, the effect on inferential narrative comprehension was maintained. This indicates that children in the dialogic reading group used narrative knowledge that they had acquired through the intervention in the narrative comprehension task. Importantly, different stories were used at the first follow-up measurement and at the posttest. Therefore, the maintenance of the effect was not due to some selective memory effect in the dialogic reading group: Children applied their narrative knowledge to similar, but new narrative contents. In contrast, the advantage in literal narrative comprehension was not maintained. Parents generally ask more questions about explicit story contents than questions about implicit story contents (Huebner & Meltzoff, 2005; van Kleeck et al., 1997). Therefore, the comparison groups probably acquired literal narrative comprehension faster than inferential narrative comprehension through shared reading at home. Taken together, the results suggest that the intervention provided unique learning opportunities for inferential skills that were not provided to the same extent by shared reading at home. Dialogic reading was effective in fostering cognitive skills that contribute substantially to children's later reading comprehension (Silva & Cain, 2015).

Previous intervention studies have found a positive effect of dialogic reading on narrative production skills (Lever & Sénéchal, 2011; Zevenbergen et al., 2003). Contrary to expectation, our intervention had no effect on narrative production skills, which can be explained by several reasons. First, the present study investigated whether actively involving children in producing narratives and answering comprehension questions would improve their ability to communicate the main points in their narratives, which would have been a transfer effect. Transfer effects of shared reading interventions are often smaller than effects on specifically targeted skills (e.g., Mol et al., 2009). Our intervention targeted narrative comprehension skills through specific, standardized questions. Thus, facilitating the production of narrative main points was not the focus of the dialogic reading intervention. Regarding that the effects on narrative comprehension were small, it is not surprising that we did not find transfer effects on narrative production.

Second, the narrative production task was particularly demanding for young children because they produced a narration based on a sequence of pictures, without any oral narrative being provided by an adult. By contrast, retelling tasks are far less demanding because the children first hear an oral narrative from an adult and then reproduce this narrative with the aid of picture prompts (e.g., Lever & Sénéchal, 2011). Possibly, a retelling task instead of a story production task would have been more sensitive for assessing differences in narrative production skills between treatment groups.

Third, whereas basic dialogic reading questions were asked by the trainers during the intervention to stimulate the oral production of the children, these questions were not asked during the assessment of narrative production skills for standardization reasons. The assessment was conducted in a one-on-one situation in which the experimenter did not participate in the narration because the aim was to assess children's narrative production skills. Therefore, children received much less feedback on their narration than during the intervention, which could have decreased their narrative production performance during the assessment in comparison to the intervention group situation. Considering that narration is inherently social, new forms of narrative assessment could use joint narration of children groups or dyads (e.g., parent and child) as a unit for analyzing intervention effects.

Crucially, the intervention in this study highlighted the importance of story grammar elements through narrative comprehension questions implicitly. Story grammar elements were not explicitly introduced as building blocks that should be used in telling stories. This strategy apparently did not benefit narrative production, which places greater demands on children's ability to structure story elements than narrative comprehension. Engaging children in narrative extratextual talk by asking comprehension questions may not be sufficient for changing the narrative schemata that underlie children's narrations. By contrast, our results suggest that positive effects of narrative comprehension questions on subsequent narrative production reported by experimental studies were probably due to an attention-guiding function of questions (Silva et al., 2014; Silva & Cain, 2017; see Lenhart, Lenhard, Vaahtoranta, & Suggate, 2019, for a discussion of the attention-guiding function of questions regarding vocabulary acquisition). Teaching story grammar by explicitly introducing, labeling and applying story categories could help preschoolers to produce better narratives. This notion is supported by a study with first graders in which explicit story grammar instruction improved both narrative comprehension and production skills (Paris & Paris, 2007). In conclusion, using picture books with a clear story line is apparently a necessary, but not sufficient condition for achieving an effect on narrative production skills. Using both implicit and explicit instruction strategies could support the development of both narrative comprehension and narrative production skills in preschool.

4.2. Effects on vocabulary skills

Vocabulary breadth and vocabulary depth increased through dialogic reading. After six months, children had learned on average three out of 15 targeted, low frequency words. In comparison, in a dialogic reading intervention with small groups that lasted for two months, children learned on average one out of 16 targeted words (Lever & Sénéchal, 2011). Thus, effects on vocabulary breadth are similar between the studies if we take into account the different durations of the interventions. Importantly, even though adding scripted narrative comprehension questions changed the intervention focus and reduced the time available for extratextual talk about vocabulary, this did not prevent the children from learning new words in this study.

Other than expected, the effect on vocabulary breadth was stronger than the effect on vocabulary depth, even though the inter-

vention targeted both skills equally. Experimental studies found that explaining word meanings during storybook reading to kindergarteners, as practiced in our intervention, leads to larger gains in their vocabulary breadth than vocabulary depth (Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009; Damhuis, Segers, & Verhoeven, 2015). Our results suggest that narrative dialogic reading with wordless picture books is particularly suitable for teaching vocabulary breadth. To achieve stronger effects on vocabulary depth, children apparently need an extended vocabulary instruction (Coyne et al., 2009).

By the beginning of first grade, children in the comparison groups had caught up with the dialogic reading group. The results (see Fig. 1) show that the dialogic reading group did not fall back in any oral language skill, but that the comparison groups were fast in catching up with them. This finding is in line with a dialogic reading study that found no maintenance of effects in a middle class sample (Sim et al., 2014). By contrast, intervention effects on vocabulary were maintained one year after a Head Start intervention with a low-income sample (Whitehurst et al., 1999). Parents with a higher socioeconomic status often engage in more early literacy activities with their children before school entry than parents with a lower socioeconomic status (Sénéchal et al., 1996). Thus, the advantage of the dialogic reading group over the comparison groups was presumably short-lived because shared reading at home also had a strong influence on their development.

4.3. Implications, limitations, and conclusions

Most effects of the low-dose intervention did not outlast the posttest, which brings up the question what amount of dialogic reading would be needed to produce more sustainable effects. According to a large survey, child care workers read on average twice a day to preschoolers for about 15 min, which sums up to about 90 h of shared reading in one academic year, of which about 72 h are spent with fiction books (Wirts, Egert, & Reber, 2017). In our study, the mean effective treatment dosage for narrative comprehension was 9.75 h (11.5 h for vocabulary), which is about 14% (16%) of the time child care workers usually spend on average reading fiction books with preschoolers (Wirts et al., 2017). To achieve substantial long-term effects, narrative dialogic reading would need to be implemented with a higher frequency and possibly a time span of two or more years (Noble et al., 2019), as was demonstrated by long-term developmental studies investigating the relation between shared reading and oral language development (e.g., Debaryshe, 1993; Farrant & Zubrick, 2013). To attain additional effects on narrative production, explicit instruction strategies could be used in addition to implicit narrative instruction (Paris & Paris, 2007).

The intervention effects were not moderated by children's storybook exposure outside the intervention, or verbal short-term memory, or nonverbal intelligence. This indicates that the intervention was equally effective for all children regardless of their storybook exposure or cognitive abilities. Even though the intervention included high demand questions and targeted higher level language skills such as inferential and literal narrative comprehension, the benefit for children did not depend on prior experiences with shared book reading, indicating that narrative dialogic reading with wordless books is a low-threshold intervention.

On a theoretical level, our results shed new light on the learning mechanisms that underlie dialogic reading. The effects of narrative dialogic reading with wordless picture books on vocabulary skills had a similar magnitude as the reported effects of dialogic reading with text-based storybooks (Mol et al., 2009). This result supports the notion that learning through conversation and discussion could be the main learning mechanism behind dialogic reading, whereas the exposure to specific linguistic features of the text in children's books might be of secondary importance for vocabulary learning. To date, it is unclear which components of dialogic reading are most effective in promoting oral language skills, and there is a lack of research on the role of different book types regarding the magnitude and sustainability of the intervention effects (Zevenbergen & Whitehurst, 2003). Future studies should investigate this question by comparing effects of dialogic reading interventions with wordless picture books versus the same picture books with some pictures replaced by text passages. Additionally, from the perspective of a triad model of shared reading in which oral language learning through shared reading originates from a triangular interplay between adult, child, and book (Fletcher & Reese, 2005; van Kleeck, 2003), our results suggest that a more direct interaction of children with wordless picture books can be similarly effective as children's interaction with text-based storybooks, which is (at least partly) mediated through adults' decoding skills.

The present study has several limitations. First, the focus of this study was to evaluate the impact of a narrative dialogic reading intervention on preschoolers' oral language skills. Regarding developmental effects, data about narrative skills should be interpreted with caution. Even though the same set of questions was used at each measurement point, we did not counterbalance the four picture books. Thus, it is possible that differences in narrative skills between measurement points are not only due to developmental trends, but also due to slight differences between the picture books used. To allow strong developmental interpretations, future studies could counterbalance picture books and measurement points. Second, children were randomly assigned to a treatment group on the child care center level. Thus, condition assignment was not entirely random on the child level. As this randomization procedure is not optimal regarding sampling error, we controlled for the dependency of observations within clusters by using hierarchical linear models with both child and child care group as random effects. Future studies should assign children individually to a treatment. Third, the children in this study had already acquired age-adequate oral language skills prior to intervention, which potentially diminished the effects of our intervention. Future studies should investigate whether children with under-average oral language skills benefit to a similar or even higher degree from a narrative dialogic reading intervention. Fourth, even though the dialogic reading intervention activities and the 20 min book sessions at the child care centers were similar to regular shared reading activities in kindergarten groups, the intervention was conducted by research assistants, which reduces the ecological validity of our findings. In future studies, child care workers should be trained as instructors to maximize the ecological validity. Fifth, our intervention was not effective in promoting narrative production skills. Adding intervention elements for explicitly teaching story grammar could be more effective in promoting preschooler's narrative production skills. Future studies could implement such explicit strategies to see whether they are effective with preschool children. In particular, future studies could compare the effects of traditional dialogic reading versus narrative dialogic reading with wordless picture books to evaluate their respective effects on children's oral language. Moreover, future studies could investigate how informational books could be used for fostering children's oral language skills through dialogic reading (e.g. Bergman Deitcher, Aram, & Johnson, 2018). Sixth, due to frequent questions about figures' inner states, it is possible that our intervention also had effects in non-language domains such as socio-emotional development (e.g., social imagination; Lysaker & Miller, 2012) and visual literacy (Arizpe, 2013), which could be investigated by future studies. Seventh, the reliability of the measures used in the present study was good, with the exception of the title recognition test's internal consistency, which was slightly lower than what is conventionally

deemed to be acceptable in educational research ($\alpha \geq .70$; but see Taber, 2018, for a critical discussion of this convention).

In conclusion, the present study provides new evidence that narrative dialogic reading with wordless picture books has similar positive effects on narrative comprehension skills as it has on vocabulary skills, underlining the versatile potential of picture book reading for the acquisition of different oral language skills. Five months after the intervention, the dialogic reading group still outperformed the comparison groups in inferential narrative comprehension, suggesting that the learning opportunities provided by narrative dialogic reading were more effective in supporting the development of this skill than the learning opportunities in other educational (or familial) settings. Our study shows that narrative comprehension questions and wordless picture books are promising modifications of the original dialogic reading intervention that contribute to enhancing the intervention's positive effects on different oral language skills.

Appendix A. Supplementary materials

Supplementary material related to this article can be found, in the online version, at <https://doi.org/10.1016/j.ecresq.2019.11.002> and <http://hdl.handle.net/21.11116/0000-0005-565F-1>.

References

- Arizpe, E. (2013). Meaning-making from wordless (or nearly wordless) picture books: What educational research expects and what readers have to say. *Cambridge Journal of Education*, 43, 163–176. <http://dx.doi.org/10.1080/0305764x.2013.767879>
- Bartón, K. (2017). MuMIn: Multi-model inference. Retrieved from <https://cran.r-project.org/package=MuMIn>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. <http://dx.doi.org/10.18637/jss.v067.i01>
- Bergman Deitcher, D., Aram, D., & Johnson, H. (2018). Does book genre matter? Boys' and girls' word learning from narrative and informational books in the preschool years. *Journal of Research in Reading*, 42, 193–211. <http://dx.doi.org/10.1111/1467-9817.12266>
- Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, 101, 294–304. <http://dx.doi.org/10.1037/a0013844>
- Cameron-Faulkner, T., & Noble, C. (2013). A comparison of book text and Child Directed Speech. *First Language*, 33, 268–279. <http://dx.doi.org/10.1177/0142723713487613>
- Chaparro-Moreno, L. J., Reali, F., & Maldonado-Carreño, C. (2017). Wordless picture books boost preschoolers' language production during shared reading. *Early Childhood Research Quarterly*, 40, 52–62. <http://dx.doi.org/10.1016/j.ecresq.2017.03.001>
- Cohrdes, C., Grolig, L., & Schroeder, S. (2019). The development of music competencies in preschool children: Effects of a training program and the role of environmental factors. *Psychology of Music*, 47(3), 358–375. <http://dx.doi.org/10.1177/0305735618756764>
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *Elementary School Journal*, 110, 1–48. <http://dx.doi.org/10.1086/598840>
- Damhuis, C. M. P., Segers, E., & Verhoeven, L. (2015). Stimulating breadth and depth of vocabulary via repeated storybook readings or tests. *School Effectiveness and School Improvement*, 26, 382–396. <http://dx.doi.org/10.1080/09243453.2014.965181>
- Debaryshe, B. D. (1993). Joint picture-book reading correlates of early oral language skill. *Journal of Child Language*, 20, 455–461. <http://dx.doi.org/10.1017/s030500090008370>
- Esser, G. (2002). *BUEVA. Basisdiagnostik umschriebener Entwicklungsstörungen im Vorschulalter [Basic assessment of developmental disorders in preschool]*. Göttingen, Germany: Hogrefe.
- Farrant, B. M., & Zubrick, S. R. (2013). Parent-child book reading across early childhood and child vocabulary in the early school years: Findings from the Longitudinal Study of Australian Children. *First Language*, 33, 280–293. <http://dx.doi.org/10.1177/0142723713487617>
- Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The effects of shared storybook reading on word learning: A meta-analysis. *Developmental Psychology*, 54, 1334–1346. <http://dx.doi.org/10.1037/dev0000512>
- Fletcher, K. L., & Reese, E. (2005). Picture book reading with young children: A conceptual framework. *Developmental Review*, 25, 64–103. <http://dx.doi.org/10.1016/j.dr.2004.08.009>
- Fox, J., & Weisberg, S. (2011). *An (R) companion to applied regression (2nd ed.)*. Thousand Oaks, CA: Sage.
- Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research*, 25, 201–239. <http://dx.doi.org/10.1006/ssre.1996.0010>
- Griffin, T., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, 24, 123–147. <http://dx.doi.org/10.1177/0142723704042369>
- Grolig, L., Cohrdes, C., & Schroeder, S. (2017). Der Titelreognitionstest für das Vorschulalter (TRT-VS): Erfassung des Lesevolumens von präkonventionellen Lesern und Zusammenhänge mit Vorläuferfertigkeiten des Lesens [The Title Recognition Test for Kindergarteners (TRT-VS): Assessment of pre-conventional readers' print exposure and its relations to precursors of reading]. *Diagnostica*, 63, 309–319. <http://dx.doi.org/10.1026/0012-1924/a000186>
- Grolig, L., Cohrdes, C., Tiffin-Richards, S., & Schroeder, S. (2019). Effects of preschoolers' storybook exposure and literacy environments on lower level and higher level language skills. *Reading and Writing: An Interdisciplinary Journal*, 32, 1061–1084. <http://dx.doi.org/10.1007/s11145-018-9901-2>
- Guo, Y., Dymia, J. M., Logan, J. A. R., Justice, L. M., Breit-Smith, A., & Kaderavek, J. N. (2016). Fidelity of implementation for an early-literacy intervention: Dimensionality and contribution to children's intervention outcomes. *Early Childhood Research Quarterly*, 37, 165–174. <http://dx.doi.org/10.1016/j.ecresq.2016.06.001>
- Huebner, C. E., & Meltzoff, A. N. (2005). Intervention to change parent-child reading style: A comparison of instructional methods. *Journal of Applied Developmental Psychology*, 26, 296–313. <http://dx.doi.org/10.1177/1468798409356987>
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50, 346–363. <http://dx.doi.org/10.1002/bimj.200810425>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.
- Krahn, F. (1979). *Robot-bot-bot*. New York, NY: Dutton.
- Lenhart, J., Lenhart, W., Vaahoranta, E., & Suggate, S. (2019). The effects of questions during shared-reading: Do demand-level and placement really matter? *Early Childhood Research Quarterly*, 47, 49–61. <http://dx.doi.org/10.1016/j.ecresq.2018.10.006>
- Lepola, J., Lynch, J., Kiuru, N., Laakkonen, E., & Niemi, P. (2016). Early oral language comprehension, task orientation, and foundational reading skills as predictors of grade 3 reading comprehension. *Reading Research Quarterly*, 51, 373–390. <http://dx.doi.org/10.1002/rrq.145>
- Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4–6-year-old children. *Reading Research Quarterly*, 47, 259–282. <http://dx.doi.org/10.1002/rrq.020>
- Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, 108, 1–24. <http://dx.doi.org/10.1016/j.jecp.2010.07.002>
- Lonigan, C. J., Purpura, D. J., Wilson, S. B., Walker, P. M., & Clancy-Menchetti, J. (2013). Evaluating the components of an emergent literacy intervention for preschool children at risk for reading difficulties. *Journal of Experimental Child Psychology*, 114, 111–130. <http://dx.doi.org/10.1016/j.jecp.2012.08.010>
- Lysaker, J. T., & Miller, A. (2012). Engaging social imagination: The developmental work of wordless book reading. *Journal of Early Childhood Literacy*, 13, 147–174. <http://dx.doi.org/10.1177/1468798411430425>
- Morrow, L. M. (1984). Reading stories to young children: Effects of story structure and traditional questioning strategies on comprehension. *Journal of Reading Behavior*, 16, 273–288. <http://dx.doi.org/10.1080/10862968409547521>
- Mascareno, M., Snow, C. E., Deunk, M. I., & Bosker, R. J. (2016). Language complexity during read-alouds and kindergartners' vocabulary and symbolic understanding. *Journal of Applied Developmental Psychology*, 44, 39–51. <http://dx.doi.org/10.1016/j.appdev.2016.02.001>
- Massaro, D. W. (2015). Two different communication genres and implications for vocabulary development and learning to read. *Journal of Literacy Research*, 47, 505–527. <http://dx.doi.org/10.1177/1086296x15627528>
- Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Press.
- Mayer, M. (1973). *Frog on his own*. New York, NY: Dial Press.
- Mayer, M. (1975). *One frog too many*. New York, NY: Dial Press.
- Moi, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, 79, 979–1007. <http://dx.doi.org/10.3102/0034654309332561>
- Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First Language*, 39(5), 527–546. <http://dx.doi.org/10.1177/0142723719849996>
- Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, 26, 1489–1496. <http://dx.doi.org/10.1177/0956797615594361>
- Nakagawa, S., & Schielzeth, H. (2012). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142. <http://dx.doi.org/10.1111/j.2041-210x.2012.00261.x>
- Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C. F., Gobet, F., et al. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, 28. <http://dx.doi.org/10.1016/j.edurev.2019.100290>. Advance online publication
- Nykrin, R., Grüner, M., & Widmer, M. (2007). *Musik und Tanz für Kinder [Music and dance for children]*. Mainz, Germany: Schott.

- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98, 554–566. <http://dx.doi.org/10.1037/0022-0663.98.3.554>
- Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly*, 38, 36–76. <http://dx.doi.org/10.1598/rrq.38.1.3>
- Paris, A. H., & Paris, S. G. (2007). Teaching narrative comprehension strategies to first graders. *Cognition and Instruction*, 25, 1–44. <http://dx.doi.org/10.1080/07370000709336701>
- Petermann, F. (2009). *Wechsler Preschool and Primary Scale of Intelligence – III, German version*. Frankfurt, Germany: Pearson.
- R Core Team. (2017). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*. Thousand Oaks, CA: Sage.
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods*, 47, 1085–1094. <http://dx.doi.org/10.3758/s13428-014-0528-1>
- Schroeder, S., & Grolig, L. (2018). Fostering Language Skills in Preschool Using Musical Training: Opportunities and Challenges. In *Research on Impacts of Arts Education: German-Dutch Perspectives*. Rat für Kulturelle Bildung e.V. Retrieved from <https://www.lkca.nl/~media/kennisbank/publicaties/2018/n-s/ publikation%20german-dutch%20colloquium.pdf>
- Sénéchal, M., & LeFevre, J.-A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, 73, 445–460. <http://dx.doi.org/10.1111/1467-8624.00417>
- Sénéchal, M., LeFevre, J. A., & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of Educational Psychology*, 88, 520–536. <http://dx.doi.org/10.1037/0022-0663.88.3.520>
- Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology*, 107, 321–331. <http://dx.doi.org/10.1037/a0037769>
- Silva, M., & Cain, K. (2017). The use of questions to scaffold narrative coherence and cohesion. *Journal of Research in Reading*, 42(1), 1–17. <http://dx.doi.org/10.1111/1467-9817.12129>
- Silva, M., Strasser, K., & Cain, K. (2014). Early narrative skills in Chilean preschool: Questions scaffold the production of coherent narratives. *Early Childhood Research Quarterly*, 29, 205–213. <http://dx.doi.org/10.1016/j.ecresq.2014.02.002>
- Sim, S., Berthelsen, D., Walker, S., Nicholson, J. M., & Fielding-Barnsley, R. (2014). A shared reading intervention with parents to enhance young children's early literacy skills. *Early Child Development and Care*, 184, 1531–1549. <http://dx.doi.org/10.1080/03004430.2013.862532>
- Stein, N., & Glenn, C. (1979). *An analysis of story comprehension in elementary school children*. In R. O. Freeble (Ed.), *Advances in discourse processing, Vol. 2: New directions in discourse processing* (pp. 53–120). Norwood, NJ: Ablex.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48, 1273–1296. <http://dx.doi.org/10.1007/s11165-016-9602-2>
- van den Broek, P., Kendeou, P., Lousberg, S., & Visser, G. (2017). Preparing for reading comprehension: Fostering text comprehension skills in preschool and early elementary school children. *International Electronic Journal Of Elementary Education*, 4, 259–268. Retrieved from <https://www.iejee.com/index.php/IEJEE/article/view/223>
- van Kleeck, A. (2003). Research on book sharing: Another critical look. In A. van Kleeck, S. Stahl, & E. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 271–320). Mahwah, NJ: Erlbaum. <http://dx.doi.org/10.4324/9781410607355>
- van Kleeck, A. (2008). Providing preschool foundations for later reading comprehension: The importance of and ideas for targeting inferencing in storybook-sharing interventions. *Psychology in the Schools*, 45, 627–643. <http://dx.doi.org/10.1002/pits.20314>
- van Kleeck, A., Gillam, R. B., Hamilton, L., & McGrath, C. (1997). The relationship between middle-class parents' book sharing discussion and their preschoolers' abstract language development. *Journal of Speech, Language, and Hearing Research*, 40, 1261–1271. <http://dx.doi.org/10.1044/jslhr.4006.1261>
- van Kleeck, A., Vander Woude, J., & Hammett, L. (2006). Fostering literal and inferential language skills in Head Start preschoolers with language impairment using scripted book-sharing discussions. *American Journal of Speech-Language Pathology*, 15, 85–95. [http://dx.doi.org/10.1044/1058-0360\(2006\)009](http://dx.doi.org/10.1044/1058-0360(2006)009)
- Wasik, B. A., Hindman, A. H., & Snell, E. K. (2016). Book reading and vocabulary development: A systematic review. *Early Childhood Research Quarterly*, 37, 39–57. <http://dx.doi.org/10.1016/j.ecresq.2016.04.003>
- Weiß, R. H., & Osterland, J. (2013). *CFT 1-R. Grundintelligenztest Skala 1: Revision [Culture fair intelligence test – Revised German version]*. Göttingen, Germany: Hogrefe.
- Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology*, 86, 542–555. <http://dx.doi.org/10.1037/0022-0663.86.4.542>
- Whitehurst, G. J., Zevenbergen, A. A., Crone, D. A., Schultz, M. D., Veltling, O. N., & Fischel, J. E. (1999). *Outcomes of an emergent literacy intervention from Head Start through second grade*. *Journal of Educational Psychology*, 91, 261–272.
- Wirts, C., Egert, F., & Reber, K. (2017). Early literacy in deutschen Kindertageseinrichtungen. Eine Analyse der Häufigkeit von Literacy-Aktivitäten im Kita-Alltag [Early literacy in German early child education and care: Analysis on the frequency of early literacy activities during a day]. *Forschung Sprache*, 2, 96–106. Retrieved from <https://www.forschungssprache.eu>
- Wright, H. H., Capilouto, G. J., Srinivasan, C., & Fergadiotis, G. (2011). Story processing ability in cognitively healthy younger and older adults. *Journal of Speech, Language, and Hearing Research*, 54, 900–917. [http://dx.doi.org/10.1044/1092-4388\(2010\)09-0253](http://dx.doi.org/10.1044/1092-4388(2010)09-0253)
- Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge, UK: Cambridge University Press.
- Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic reading: A shared picture book intervention for preschoolers. In A. van Kleeck, S. Stahl, & E. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 177–200). Mahwah, NJ: Erlbaum. <http://dx.doi.org/10.4324/9781410607355>
- Zevenbergen, A. A., Whitehurst, G. J., & Zevenbergen, J. A. (2003). Effects of a shared-reading intervention on the inclusion of evaluative devices in narratives of children from low-income families. *Journal of Applied Developmental Psychology*, 24, 1–15. [http://dx.doi.org/10.1016/s0193-3973\(03\)00021-2](http://dx.doi.org/10.1016/s0193-3973(03)00021-2)
- Zucker, T. A., Cabell, S. Q., Justice, L. M., Pentimonti, J. M., & Kaderavek, J. N. (2013). The role of frequent, interactive prekindergarten shared reading in the longitudinal development of language and literacy skills. *Developmental Psychology*, 49, 1425–1439. <http://dx.doi.org/10.1037/a0030347>