Besser, Nadine; Linberg, Anja; Dornheim, Dorothea; u. a.

Fostering toddlers’ numeracy and mathematical language skills through a professional development intervention on interaction quality in toddler classrooms

Date of secondary publication: 02.05.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-1079743

Primary publication

Besser, Nadine; Linberg, Anja; Dornheim, Dorothea; u. a. (2025): Fostering toddlers’ numeracy and mathematical language skills through a professional development intervention on interaction quality in toddler classrooms, in: Early childhood research quarterly : the scholarly journal of the National Association for the Education of Young Children (NAEYC), Amsterdam [u.a.]: Elsevier Science, Vol. 72, pp. 44–55, doi: 10.1016/j.ecresq.2025.02.004.

Early Childhood Research Quarterly 72 (2025) 44–55

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

journal homepage: www.elsevier.com/locate/ecresq

Research Paper

Fostering toddlers' numeracy and mathematical language skills through a professional development intervention on interaction quality in toddler classrooms

Nadine Besser [c,*], Anja Linberg [b,1], Dorothea Dornheim [a,2], Sabine Weinert [a,3], Hans-Günther Roßbach [a,4], Simone Lehrl [c,5]

[a] *Department of Developmental Psychology, Otto-Friedrich-University of Bamberg, Germany*
[b] *Department of Social Monitoring and Methodology, German Youth Institute (DJI), Germany*
[c] *Department of Education, University of Education Weingarten, Germany*

A R T I C L E   I N F O

Keywords:
Numeracy skills
Mathematical language
Toddlers
High-quality interactions
Domain-specific
Intervention

A B S T R A C T

The first years of life are crucial for children's concurrent and subsequent development in various domains. Given the recognized importance of high-quality adult-child interactions for promoting young children's development during this vulnerable phase, supporting such interactions is essential. This study examines the impact of a 9-week domain-specific mathematical interaction training for early childhood education and care (ECEC) teachers on children's development of numeracy skills and mathematical language from age two to four years (N = 408 children). Specifically, it investigates whether the children in groups with teachers trained math-specific, would achieve better outcomes on numeracy skills and mathematical language than a control group who followed their usual program, and, to test for the math-specific effect, to a control-intervention group who received a general interaction training. Moreover, effects on math outcomes were contrasted to general language outcomes to cross-check the training's specific impact. Results showed significant improvements in numeracy skills, in mathematical and general language over time. The math-specific intervention had particularly strong and lasting effects on numeracy and mathematical language, while the general intervention specifically improved general language compared to the control group. The two intervention groups differed significantly in their math and language skills at posttest, but these differences were reduced at follow-up testing. These findings underscore the effectiveness of targeted professional development (PD) training in fostering children's early numeracy skills and mathematical language. By systematically varying intervention content and controlling for environmental factors, this study provides valuable insights into the benefits of math-specific training for ECEC teachers.

Mathematical competencies are crucial for many individual and occupational activities in an increasingly technological and information-based society (OECD, 2023; Ropohl et al., 2018). It is well known that mathematical competencies develop early and are predictive of later mathematical performance (e.g., Lehrl et al., 2020; Nguyen et al., 2016). Thus, mathematical competencies of very young children seem to lay the foundation for later mathematical learning. Besides the so-called emergent numeracy skills (e.g., counting or cardinal number knowledge, Purpura & Lonigan, 2013), which are considered most predictive of later mathematical performance (Nguyen et al., 2016), mathematical

* Corresponding author at: University of Education Weingarten, Kirchplatz 2, 88250 Weingarten, Germany.
E-mail addresses: nadine.besser@uni-bamberg.de, nadine.besser@ph-weingarten.de (N. Besser), linberg@dji.de (A. Linberg), dorothea.dornheim@uni-bamberg.de (D. Dornheim), sabine.weinert@uni-bamberg.de (S. Weinert), hans-guenther.rossbach@uni-bamberg.de (H.-G. Roßbach), simone.lehrl@ph-weingarten.de (S. Lehrl).
[1] German Youth Institute (DJI), Nockherstr. 2, 81541 Munich, Germany.
[2] Otto-Friedrich-University of Bamberg, Markusplatz 3, 96050 Bamberg, Germany.
[3] Otto-Friedrich-University of Bamberg, Markusplatz 3, 96050 Bamberg, Germany.
[4] Otto-Friedrich-University of Bamberg, Markusplatz 3, 96050 Bamberg, Germany.
[5] University of Education Weingarten, Kirchplatz 2, 88250 Weingarten, Germany.

https://doi.org/10.1016/j.ecresq.2025.02.004
Received 15 August 2024; Received in revised form 8 January 2025; Accepted 4 February 2025
Available online 13 February 2025

language skills seem to be another important sub-domain in the course of explaining mathematical development (e.g., Peng et al., 2020; Purpura et al., 2023; Turan & De Smedt, 2022). Empirical evidence shows that children's early numeracy skills and their mathematical language differ significantly from the age of three (e.g., Clements & Samara, 2007; Dubowy et al., 2008; Purpura & Reid, 2016). Therefore, fostering numeracy skills and mathematical language under the age of three seems to be a reasonable approach to counteract differences in early mathematical competencies that build a foundation for a strong understanding of numbers and mathematical concepts and prepare children for the requirements of later life.

There is evidence that children's development of numerical and mathematical language skills is fostered by high-quality interactions between early childhood education and care (ECEC) teachers and children (e.g., Burchinal et al., 2014; Lehrl et al., 2016; Purpura, Napoli, et al., 2017). Studies indicate that domain-specific interactions, such as talking about mathematical content during shared activities, promote particularly children's numeracy skills and their mathematical language (King & Purpura, 2021; Purpura, Napoli, et al., 2017) whereas more general interactions which are related to children's learning, such as dialogic reading of picture books, enhances children's general language skills (e.g., Grolig et al., 2020; Pillinger & Vardy, 2022). However, the quality of both math-specific and general interactions in early childcare is often in the medium to low range in many countries (e.g., Germany: Bücklein et al., 2017; United States: Hamre & Pianta, 2007; Australia: Siraj et al., 2023). Moreover, teachers' beliefs are less positive towards the stimulation of mathematics compared to other developmental domains such as language or emotions (e.g., Costa et al., 2023; Jenßen, 2021). Consequently, children might lack adequate learning experiences in early childcare, especially in the mathematical domain. Professional development (PD) training, understood as a systematic process aiming to ensure that caregivers are qualified to work with children and their families, addresses these issues as it is linked to high interaction quality (Egert et al., 2018). This intervention study therefore investigates whether the participation of teachers in a training on math-specific interaction quality (math-focused intervention group—MIG) can promote the development of children's numeracy skills and mathematical language. To examine the specific effectiveness of the training, we compare the effects of the math-specific intervention to the effects of a control-intervention (general intervention group—GIG), focusing on general high-quality interactions. Moreover, in order to crosscheck possible specific effects, we contrast the effects on mathematical outcomes to the effects on children's general language outcomes (assuming that the math-specific interaction training promotes numerical skills and mathematical language in particular, while the general training fosters general language comprehension).

1. The importance of early numeracy skills for mathematical development and the specific role of mathematical language

Early numeracy skills include three interrelated domains: numbering (e.g., verbal counting, counting objects), relations (e.g., number comparison, number identification), and arithmetic operations (e.g. addition, subtraction) (Purpura & Lonigan, 2013). Numerous studies have demonstrated that early numeracy skills are predictive of both numerical and overall mathematical development including mathematical problem solving. For example, numeracy skills at the age of 4 to 6, including numbering and relations, predict mathematical performance in elementary school (Krajewski & Schneider, 2009; Niklas & Schneider, 2017). Moreover, Nguyen et al. (2016) demonstrated that the numeracy skills, including numbering, relations and arithmetic operations of 4-year-olds predict mathematical performance in 5th grade even when controlling for other mathematical domains, such as measurement and data, geometry, and patterning. Studies focusing on younger age groups to examine the early relationships between numeracy skills and mathematical development report for instance that number discrimination in two-year-olds predicts counting and arithmetic at the age of four (Ceulemans et al., 2015). Moreover, Seitz and Weinert (2022) demonstrated even earlier predictive relationships between numeracy skills and later mathematical competencies: Children's performance in a numerical habituation task at the age of 1 ½ predicted their mathematical competencies at the age of 4. Besides the described early numeracy skills that include counting skills, identification of numerals, cardinality or calculation, the children's proficiency in mathematical language, e.g., knowing math-specific vocabulary like "tall", "a little bit", "under" or "small" has been repeatedly shown to be of particular importance, as findings suggest that mathematical language is more predictive of children's numeracy skills than general language (e.g., Purpura & Logan, 2015; Toll & van Luit, 2014). Mathematical language refers to terms and concepts required for understanding and solving mathematical problems (Han & Ginsburg, 2001). There is evidence that mathematical language and numeracy skills are related as early as the age of three (e.g., Hornburg et al., 2018; Kung et al., 2019; Purpura et al., 2019). Furthermore, intervention studies indicate that increasing children's exposure to mathematical language positively influences their numeracy skills as well as their broad mathematical competencies (Payne, 2024; Purpura, Napoli, et al., 2017; Purpura et al., 2019, 2023). Since numeracy skills and mathematical language are closely related, fostering both skills at an early age has the potential to provide children with the best possible development right from the start.

2. Quality of teacher-child interactions and effects on early numeracy skills and mathematical language

Research suggests that high-quality interactions between teachers and children foster children's competence development (e.g., Burchinal et al., 2014; Lehrl et al., 2016). Interactions between teachers and children are considered the centerpiece of process quality in childcare (Mashburn et al., 2008), whereby those interactions are often described as being part of either general (e.g., general cognitive and/or social-emotional support) or domain-specific (e.g., mathematical stimulation) processes (Kluczniok & Roßbach, 2014; Sylva et al., 2011). General high-quality interactions include teacher's emotional warmth, sensitivity, and their general support for language and learning (Beckh & Becker-Stoll, 2016; La Paro et al., 2012). While high-quality general interactions are assumed to promote children's cognitive-verbal and socio-emotional development in general, high-quality domain-specific interactions are aimed at promoting specific areas of development such as mathematics (Trawick-Smith et al., 2016), based on the assumption that specific outcomes require specific interactions (Bornstein, 2017). Activities such as dialogic reading (e.g., Whitehurst et al., 1994) or playing board games (e.g., Siegler & Ramani, 2009) provide an opportunity to shape high-quality interactions. Dialogic reading is an interactive activity in which not only the adult but also the child takes an active role (Whitehurst et al., 1994). A number of intervention studies have demonstrated that the enhancement of dialogic reading is a powerful tool that allows opportunities for high-quality interaction between teachers and children resulting particularly, but not exclusively, in advanced language competencies in children (e.g., Pillinger & Vardy, 2022; Towson et al., 2023). However, an increasing number of studies also investigate the relationship between dialogic reading and numeracy and mathematical language development in children; yet, these studies focus on math-related dialogic reading sessions conducted by specifically qualified research assistants rather than teachers (Purpura, Napoli, et al., 2017; Purpura et al., 2023; Wijns et al., 2023). As children spend a large amount of time in ECEC (OECD, 2022), it is important to train teachers to shape high-quality interactions. This is particularly true as these studies conducted with children from the age of three onwards demonstrate that, especially, children's mathematical language can be stimulated through math-specific dialogic reading. However, the results regarding numeracy skills remain unclear, as associations are both found (e.g., Purpura, Napoli, et al., 2017) and not found (e.g., Wijns et al., 2023). The quality of interaction in its various facets can be improved through professional development (PD) training. Theoretical models imply a multi-stage path from teachers' PD to improved children's outcomes (Egert et al., 2018). It is assumed that PD changes the actual observable performance (e.g., interactions, activities) of the teachers via latent dispositions (e.g., skills, knowledge, attitudes/orientations; Buysse et al., 2009; Snyder et al., 2012) which leads to an increase in children's mathematical and/or other competencies via an improved quality of interaction (Brunsek et al., 2020; Egert et al., 2018). Nevertheless, to our knowledge, no study included PD in their intervention and trained teachers on dialogic reading strategies with a specific focus on mathematical stimulation, considered as a way of shaping domain-specific high-quality interactions.

In addition to math-related dialogic reading, playing board games in combination with talking about mathematical content seems to be a promising method for supporting the development of children's numeracy skills (Gasteiger & Moeller, 2021; Payne, 2024; Ramani & Siegler, 2008). For example, Gasteiger and Moeller (2021) demonstrated that playing conventional board games with traditional dot dice (1-6) over a 4-week period improved the counting skills and conceptual subitizing ability of 4 to 6 year old children. Moreover, Payne (2024) found that playing board games over an 8-week period with four-year-olds was beneficial for their numeracy development, particularly, although non-significant, when mathematical language was emphasized during play. However, in these studies again, the children played the games mostly with their parents or with research assistants and none of these studies included mathematical language as an outcome. Importantly, within this context, there is no study that has investigated the importance of the specificity of the training by offering and comparing a training in math-specific as well as a training in general interaction quality.

2.1. Study objectives

Previous studies on the promotion of children's numeracy skills and mathematical language show that promising approaches seem to involve initiating and engaging in high-quality math-related interactions with the children, which can take place in activities like math-specific dialogic reading and playing board games including the use of mathematical language. However, it is important to note that most previous studies investigating the promotion of children's numeracy skills and mathematical language either included specially trained research assistants or instructed parents to support the children. Furthermore, the different approaches used to promote numeracy skills and mathematical language are only examined separately and not combined into a PD training for teachers. Accordingly, there is a lack of well-controlled intervention studies with teachers to investigate the causal relationship between PD training on math-related interaction quality and the development of early numeracy skills and mathematical language before the start of formal schooling (Nelson et al., 2024). Furthermore, most of these earlier studies examined the effects of the intervention on numeracy skills and mathematical language in children from the age of three onward. However, since individual differences in numeracy skills and mathematical language can be found and measured already at the age of two and since more and more children under the age of three are being cared for in ECEC (OECD, 2022), it is important to consider this early phase as well. Therefore, we integrated promising strategies for shaping math-related high-quality interactions characterized by sensitivity and cognitive-verbal stimulation of math development in activities like dialogic reading and playing board games into PD training for teachers working with two-year-old children and evaluated the short- and long-term effectiveness of such training—as compared to a general interaction quality training and a control group—for the development of children's numeracy skills, their mathematical language and—to contrast the possible specific effects—their general language between the ages of two and four. In particular, in the intervention study, we expected the participation of teachers in PD training on math-related interaction quality to have positive effects on the development of children's numeracy skills and mathematical language compared to the other groups, while we expected the participation in a general interaction quality training to be particularly effective in promoting general language development. By forming two intervention groups with a different focus in terms of training content (math-specific interaction quality and general interaction quality, respectively) while applying the same training methods and a control group without training, we examined whether the effects were particularly attributable to participation in math-specific training.

2.2. Present study

To investigate the effect of a math-specific intervention on children's early numeracy skills and mathematical language development, we set up a design that, in addition to the classical control condition without intervention, systematically varied the content of the intervention (math-specific vs. general). This was necessary to rule out the possibility that a training on overall warm, sensitive and general cognitive-verbal stimulating interactions would be "enough" to foster the development of numeracy skills and mathematical language. Moreover, besides numeracy skills and mathematical language, we also assessed the children's general language skills. As briefly outlined above, we expect a math-specific PD intervention to impact particularly children's numeracy and mathematical language development, while the general interaction quality intervention should impact children's general language skills, but less (or even not at all) their mathematical development. Such differential training effects would highlight the specificity of training effects. Thus, we hypothesize the math-specific intervention group to demonstrate higher numeracy and mathematical language progress than children both in the control group and in accordance to the specificity principle by Bornstein (2017) also compared to the general intervention group. For the contrasting domain, general language development, we expect the general intervention group to be particularly effective compared to the math-specific intervention group. However, both intervention groups should outperform the control groups, as they were both trained in content regarding language promoting strategies. As children in Germany move from "Krippe" (childcare 0-3 years) to "Kindergarten" (childcare 3-6 years) at around the age of three and are therefore no longer cared for by the trained teachers after the posttest, children cannot profit from the trained teacher's environment any longer. Instead, in line with the principle of self-productivity (Cunha & Heckman, 2007), which follows the idea that skills acquired in one period persist into future periods, and that skills are self-reinforcing, we assume that the initial boost that the children experience as a result of the teachers' participation in the training will persist over the course of the year between posttest and follow-up testing.

3. Methods

3.1. Study design

The present study is part of the intervention study EarlyMath project (Linberg, Lehrl, Dornheim, Weinert, & Roßbach, 2020, Lehrl et al., 2024) funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), which examines several family and institutional factors influencing the development of numeracy skills and mathematical language of children between the ages of two to four years (Fig. 1). Since 2020 the project investigated N = 95 groups from Bavaria (Germany) in a 2-cohort design (intervention period cohort 1 (N = 50): 12/2020 to 04/2021, cohort 2 (N = 45): 03/22 to 06/22) with a pre- and posttest presented before and after the intervention (impact period: 2 months), respectively, and a follow-up one year later. The teachers (one per childcare center) were assigned quasi-randomized to one of four experimental conditions based on structural characteristics of the

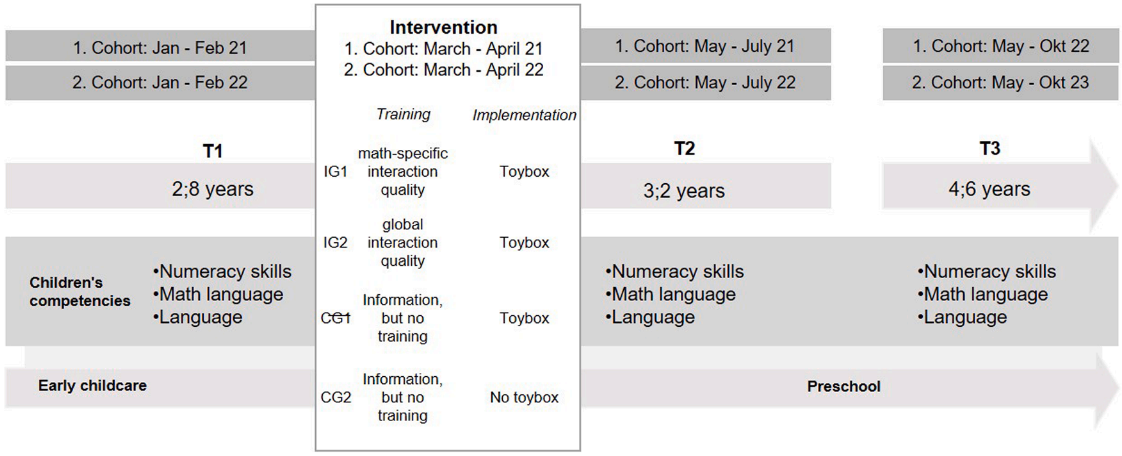

Fig. 1. Design of the early math study.
Notes. T1 – T3: Measurement Points 1 to 3

childcare centers (proportion of children with a migration background, socio-economic background of the children [low, medium, high; rated by the teachers], region [suburbs, medium-sized city, large city], number of children per group). As preschool teachers and their corresponding children were randomly assigned to the study conditions and the differences in competencies are therefore by chance, we combined the two cohorts. Descriptive information on the analysis variables by group assignment and cohort can be found in the online supplement (Tables S1-S3)

During a nine-week intervention phase, teachers in the two intervention groups received two days of intensive training on math-specific (MIG; $N_{\text{children}} = 131$) or general (GIG; $N_{\text{children}} = 141$) interaction quality, which was complemented by an implementation period and reflection sessions (see below); teachers in the two control groups (control group 1 - CG1; $N_{\text{children}} = 70$; and control group 2 - CG2; $N_{\text{children}} = 66$) received only an information session on interaction quality. All groups except CG2 also received a toy box before the posttest while CG2 received it after the posttest. This was done to exclude the possibility that providing teachers with stimulating materials is enough to improve the quality of interaction. As the children in both control groups did not differ in their performance, both groups were combined (CG; $N_{\text{children}} = 136$) to obtain nearly even group sizes for all study conditions (see online supplement Table S4).

3.2. Participants

In Germany, childcare is typically organized in groups for children under the age of 3 (crèche or crèche groups) and for children aged 3 to 6 (preschool) (with some overlap due to available places in the subsequent institution). As the target age of children was around 2;6 years, early childcare centers in Bavaria (Germany) that cared for children under the age of three were contacted for recruitment. Only one teacher was recruited per center. The teachers were mainly female (93.3 %), on average 36 years old (*SD* = 11 years) and most of them had 5 years of vocational training (66.3 %). Vocational training of ECEC teachers in Germany combines vocational school and supervised internship and focuses specifically on the needs of children aged 0 to 6 years in non-academic, play-oriented settings. However, the period of the narrow preschool years (3 to 6) is in most cases the center of the curriculum in which math tends to be a marginal topic. Typically, in Bavaria (Germany), about 12 children per group are cared for aged from zero to three years. About 3 to 6 children per group fell in our target age range (2;6 - 2;11). The teachers handed the information material and consent forms to all the parents in their group who had children of the target age. An average of four children (range: 1-8) per early childcare center participated in the project, after the parents had given their informed consent. Participation was voluntary and the children's agreement was asked before the playful test sessions. The sample (N = 408) was tested at three measurement times. At the first measurement time, the children were on average M = 31.40 months old (*SD* = 3.32, range = 20.00-41.00 months). The sex of the children was approximately equally distributed (52.8 % male). At home, most children spoke exclusively German (77.7 %), only a few spoke at least one other language besides German(20.4 %), and only eight children (1.9 %) spoke exclusively a language other than German at home. Parents were mostly high educated (71.5 % with at least one parent holding a college degree), while under one-third-of the parents indicated vocational education (28.5 %). Further descriptive information on child and teacher-related variables can be found in the online supplement (Tables S5 and S6).

3.3. Procedure of the assessment of children's numeracy skills, their mathematical language and their general language comprehension

Trained examiners tested the children at three measurement times: at pre- and posttest with an average of 4 months between testing, and again at a follow-up testing one year after the posttest. The individual tasks to assess the children's numeracy skills, their mathematical language and their general language comprehension were conducted on at least two consecutive days in sessions of about 30 to 40 min each. The appointments were coordinated with the participating teachers and usually took place in the morning after breakfast and/or the morning circle, in which the examiners were already participating, allowing the children to meet the examiners outside of the test situation. The examiners were students who either already had a bachelor's degree in psychology, pedagogy, or educational science or were about to complete it. Before conducting the tasks with the children, each examiner completed a six-hour online training session[6] with video feedback conducted by the project's research assistants. The examiners were blind to the intervention group assignment of the teachers and children, respectively.

3.4. Measures

Within the framework of the EarlyMath project, a variety of numerical, verbal and general cognitive competencies of the children were assessed. All tasks and tests were implemented in a playful manner with

[6] Due to the COVID-19 protection policies at that time, we had to switch to a digital training for the first cohort. We continued to provide the same online training for the second cohort after the protection policies were removed, in order to keep the impact the same.

the help of a stuffed animal (a sheep named Woolly), which was supposed to motivate and focus the children's attention. Table 1 provides a descriptive summary of the aggregated scales (including number of items and internal consistency).

3.4.1. Numeracy skills

We operationalized the construct of numeracy skills by summing up the scores achieved per child on four subscales described hereafter and then dividing them by the highest possible score (Pretest: M = .27, *SD* = .12; posttest: M = .31, *SD* = .14; follow up: M = .55, *SD* = .14). Confirmatory factor analyses revealed a unidimensional factor (Chi-square $\chi^2 = 255.49$, $p < .001$; CFI = .94; TFI = .82; RMSEA = .15; SRMR = .06). Children's counting skills were assessed using a *verbal counting task* and a *counting objects task*. The *Give-a-number task* (Give-N; Sarnecka & Carey, 2008; Wynn, 1990) and the *Point-to-x task* (Wynn, 1992) were used to measure the development of the children's number concept.

Verbal counting: The children were asked twice in succession to count as far as they could (starting at 1). In the beginning, Woolly counted up to 3 to help them. Similar to the counting objects task, the following number word series are scored: (respectively up to) 1, 2, 3, 5, 7, 10, 15, 20, 30, >30, i.e., the child receives one point for each of the lengths of the number word series named here. For example, three points were given for reciting the number word series up to 3, and six points were given for counting up to 10.

Counting Objects: The children were asked to help Woolly the sheep to count the stones that Woolly had collected. Ten different stripes with glass nuggets (quantities: 1, 2, 3, 5, 7, 10, 15, 20, 30, >30) were presented to the child one by one, starting with three glass nuggets). If this item was solved correctly, the next stripes were presented in ascending order until the child either worked through all the items or the dropout criterion (two wrong stripes in a row) was reached. If the initial item was solved incorrectly, the child was presented first with one glass nugget and then with two glass nuggets. In this case, the task continued until either the dropout criterion was reached or the child completed all items. The child received one point for each correctly solved item.

Give-a-number task: In the Give-N task (Sarnecka & Carey, 2008; Wynn, 1990), the children were supposed to feed Woolly with a certain amount of food (wooden cubes) by placing the required number of wooden cubes on a plate. At the beginning, Woolly demonstrated an example item (1 piece of food). Then, in succession, the numbers 1 to 7 were each asked twice, though the same numbers did not directly follow each other, except number 1 (1, 1, 2, 3, 2, 3, 4, 5, 4, 5, 6, 7, 6, 7). The task was finished when the children had either completed all the items or the dropout criterion (three consecutive wrong items) was reached. For each correctly given number, the children received one point.

Point-to-x task (Wynn, 1992): The children were asked to identify a specific quantity of objects from two presented quantities. On picture cards, two quantities of identical objects were presented in a vertical arrangement (e.g., one cat vs. two cats). Quantities 1 to 8 (1 vs. 2, 2vs. 3, 3. vs. 4, etc.) were each asked twice with different objects. Each item was presented in ascending order until either the children completed all items or the dropout criterion (four consecutive items incorrect) was reached. For each quantity correctly identified by tapping at it, the children received one point.

3.4.2. Mathematical language

The children's mathematical language was assessed using a German short-version of the Preschool Assessment of the Language of Mathematics (PALM; Purpura & Logan, 2015). In this task, which deliberately excludes numerals, children were presented with various terms of mathematical language particularly unspecified quantitative terms (e.g., "a bit", "more") and spatial prepositions (e.g., "up", "in front") to either identify on picture cards (e.g., "Where are more dots?") or perform with wooden cubes (e.g., "Take away many blocks."). The EarlyMath team translated the PALM (Purpura & Logan, 2015) into German and expanded it by adding geometric items about shapes, bodies and size

Table 1
Descriptive information on the analysis variables (sum scores[1]).

	Items			N			M			SD			Min			Max			Internal consistency ρ_{KR20}		
	Pre	Post	Follow	Pre	Post	Follow	Pre	Post	Follow	Pre	Post	Follow	Pre	Post	Follow	Pre	Post	Follow	Pre	Post	Follow
Numeracy skills	48	55	55	343	351	209	12.99	17.13	29.35	6.45	8.19	10.53	0	0	6	35	41	55	.84	.89	.98
Verbal Counting	10	10	10	341	349	209	2.11	3.08	5.95	2.14	2.35	2.02	0	0	0	9	9	10	.83	.84	.93
Counting objects	10	10	10	342	350	209	1.97	3.03	5.48	1.91	2.12	2.08	0	0	0	7	8	9	.78	.80	.92
Give-N	14	17	17	339	350	209	3.61	4.65	8.94	1.73	2.31	4.46	0	0	0	11	15	17	.67	.78	.96
Point-to-X	14	18	18	307	331	208	5.98	6.80	9.02	3.15	4.13	5.14	0	0	0	12	17	18	.73	.84	.96
PALM	21	31	39	315	332	208	8.91	13.97	26.73	3.53	4.62	.8.27	0	0	0	19	26	39	.67	.71	.98
TROG-D/ SETK-2	44	44	36	317	326	208	12.96	17.10	15.50	7.23	8.25	7.27	0	0	0	33	37	30	.91	.93	.96

Give-N = Give-a-number task (Sarnecka & Carey, 2008; Wynn, 1990); Point-to-*X* = Point-to-x task (Wynn, 1992); PALM = Preschool Assessment of the Language of Mathematics (PALM; Purpura & Logan, 2015); TROG-*D* = short-form of the German Version of the Test for Reception of Grammar (Fox-Boyer et al., 2020; Weinert & Ebert, 2013); SETK-2 = Sprachentwicklungstest für zweijährige Kinder [Language development test for two-year-old children] (Grimm, 2016).

[1] Note that the sum-scores are partially based on different numbers of items as some tasks were extended by more difficult items due to the age-related skill development of the children. In the analyses, this was taken into account.

(for number of items and internal consistency see Table 1). The children completed all items regardless of the correctness of the solutions; there was no dropout criterion. The children received one point for each correctly solved item. The analyses are based on the scores achieved by the children, which were then divided by the highest possible score (Pretest: M = .41, *SD* = .15; posttest: M = .45, *SD* = .14; follow up: M = .68, *SD* = .15).

3.4.3. General language comprehension

The comprehension of morpho-syntactic structures of the German language was assessed using the German Version of the Test for Reception of Grammar (TROG-D; Fox-Boyer et al., 2020 shortened; see Weinert & Ebert, 2013). The children were asked to select the picture out of four colored pictures that best matched the sentence presented by the examiner. The items were ranked in ascending order of difficulty. The children received one point for each correctly solved item. The analyses are based on the mean values of the composite scale (Pretest: M = .29, *SD* = .15; posttest: M = .39, *SD* = .17; follow up: M = .43, *SD* = .14). As the TROG-D (Fox-Boyer et al., 2020 shortened; see Weinert & Ebert, 2013) is only normed from the age of three, the test was supplemented at the first two measurement times with a comparable item set suitable for two year old children from the Sprachentwicklungstest für zweijährige Kinder [Language development test for two-year-old children] (SETK-2; Grimm, 2016) to avoid floor effects.

3.4.4. Control variables

Since the professional experience of teachers has an influence on the way in which teacher-child interactions are shaped (Kuger et al., 2016; LoCasale-Crouch et al., 2007), we controlled for teacher's working experience in years. Moreover, we controlled for the age and sex of the children and for some family background characteristics, as these variables are closely related to the development of numeracy skills and mathematical language (Kung et al., 2019). The family background characteristics were indicated by the variables home language (other language than German / exclusively German), mother's education level (vocational / college degree), number of children and single parent (yes / no). The information was provided by the teachers (work experience) and the parents (child-related items) via questionnaires.

3.5. Intervention

We designed four experimental conditions (two intervention groups and two control groups) in which we varied the content of training and the support of the implementation. In the intervention conditions, two training sessions aimed to improve math-specific (MIG) and general (GIG) interaction quality, respectively, of teachers. As practical application and practice of the content is an essential part of effective training (e.g., Egert et al., 2018), we have integrated this as a key component of the training. Likewise, we placed a strong emphasis on linking theoretical knowledge with practical application throughout the training and practiced high-quality general interaction behavior using the method of dialogic picture book reading (e.g., Grolig et al., 2020; Pillinger & Vardy, 2022; Wijns et al., 2023). Please note, however, that in this study dialogic picture book reading is not understood as an exercise but as a principle of support through which sensitive and general cognitive-verbal stimulating interactions can be shaped. Consequently, the teachers were trained in how to shape a picture book situation in which they perceive the child's perspective, for example, pick up on and supplement the child's verbal expressions and stimulate the child's thinking and general language learning by linking the storybook experiences with the children's everyday life.

The project's three research assistants conducted the training sessions, which were separated in nine training modules (for more information on the intervention and the training content, see online supplement Table S7). The training sessions each lasted 6 h (plus breaks) over two days with a one-week interval, during which the initial training content could already be practiced in pedagogical everyday situations. The methodological structure of the training sessions in MIG and GIG was the same. First, each module began with a short theoretical input, considering the prior knowledge of the participating professionals to create a common knowledge base (Archambault et al., 2022; Chen et al., 2018). Then, by using the dialogic picture book reading method, best-practice examples in video vignettes were analyzed collaboratively using observation sheets developed specifically for the project, thus training the professional perception of interactions of the participating teachers (Major & Watson, 2018). Subsequent role-plays in small virtual groups supported the transfer into practice (e.g., Archambault et al., 2022). For both intervention groups, the initial focus was on the concepts and importance of emotional support and general cognitive-verbal stimulation as central dimensions of pedagogical interactions. For the teachers in MIG only, these concepts were expanded by addressing the importance and practicing the use of mathematical language. In the later modules, the dialogic reading strategies that were addressed and practiced in modules 2-5, were transferred to playing board games (module 6) and interactions in everyday pedagogical situations (module 7). The content of the modules focused exclusively on mathematics in MIG (size and measurement; counting, numbers and arithmetic; space, shape and time; patterns and relationships). In GIG, various contents were addressed (environment and nature; mathematics; physical activity; social relationships) which, as it was a general training, also included small parts of mathematics but without placing a specific focus on mathematics (approximately only 30 min mathematics in GIG vs. about 4 h in MIG).

The teachers of both control groups (CG1 and CG2) attended a 90-minute online information session in which content on math-specific and general interaction quality was presented. However, unlike the intervention groups, the content was more general, without opportunities for practical application, and no additional exchange followed the information session.

To extend the interaction training relating to the presence of high-quality materials, shortly before the start of the intervention, the teachers of all groups except CG2 received a toy box with standardized content, including picture books and board games, being particularly, but not exclusively, suitable for stimulating mathematical content. The teachers in both, MIG and GIG, used some of the picture books and board games from the toy boxes during the training sessions. We systematically varied the provision of toy boxes within the control groups (CG1 received the toy box before the posttest, CG2 afterwards) to exclude the possibility that it would be "sufficient" to provide teachers with appealing and high-quality children's books and play materials. However, the control groups did not differ in terms of children's skills at posttest, thus, to increase statistical power, we collapsed both groups (see study design).

An eight-week implementation phase after the intervention was intended to ensure that the teachers from MIG and GIG also implemented the training content in their everyday pedagogical work. The implementation tools included self-reflection using observation sheets, reminders of the picture books and board games from the toybox, newsletters with various stimuli and ideas for high-quality activities, as well as self-designed "pocket cards", which were small cards with interaction strategies and specific prompts for the picture books and board games from the toybox for spontaneous use in interaction situations.

3.6. Statistical analyses

As is common in longitudinal studies, not all children in this study participated at all measurement points (for patterns of missing values in the analysis variables please see online supplement Table S8). We dealt with the resulting missing by imputing them with SPSS using the regression-based linear trend at point option, where a regression is performed on an index variable scaled from 1 to n for the existing data series and missing values are replaced by their predicted value. Afterwards the intervention effects were examined in five steps. In a first step, conducting a one-factorial analysis of variance (one-way ANOVA), it was ensured that the children's performances did not differ between groups at the first measurement. This was followed by a mixed-model ANOVA for both numeracy skills and mathematical language to examine whether the children's performances differed after the intervention both in terms of time (three-level within-participant factor) and in terms of their intervention group assignment (three-level between-participant factor) thereby focusing on the hypothesized interaction effect between these two factors (within-between interaction). We also examined whether the results of the mixed-model ANOVA vary when controlling for work experience of the teachers and age and sex of the child as well as their family background characteristics (indicators: home language, mother's education level, number of siblings, single parent). Furthermore, we conducted the mixed-model ANOVAS separately for the individual time periods (pretest vs. posttest, pretest vs. follow-up test, and posttest vs. follow-up test) to identify whether a differential training effect was observed mainly at posttest or also for the follow-up test. Finally, a-priori contrasts (MIG vs. CG; GIG vs. CG; MIG vs. GIG) were conducted to test the hypotheses that a math-specific intervention is of special importance for the development of numeracy skills and mathematical language in contrast to the general intervention and the control group. In order to crosscheck possible specific effects, we contrast the effects on mathematical outcomes to the effects on children's general language outcomes, and therefore additionally use general language skills (receptive grammar) as outcome variable. The main analyses were performed with R (version 4.4.1) using the afex package (Singmann et al., 2012). The effect sizes are reported in partial η^2 (η_p^2), with $\eta_p^2 \sim .01$ being considered as small, $\eta_p^2 \sim .06$ as medium and $\eta_p^2 \sim .14$. as large (Cohen, 1988).

4. Results

Prior analyses confirmed that neither the numeracy skills ($F(2,408) = 0.49$, $p = .615$, $\eta_p^2 = .00$), nor the mathematical language ($F(2,408) = 0.83$, $p = .435$, $\eta_p^2 = .00$), nor the general language comprehension ($F(2,408) = 0.30$, $p = .740$, $\eta_p^2 = .00$) differed at the first measurement point depending on the group assignment. Based on this result, and since the conventional test statistic requirements (e.g., homoscedasticity, sphericity, normal distribution of the residuals) were also met, the mixed-model ANOVAS and the a-priori contrasts were conducted.

4.1. Mixed-models ANOVA across three measurement times

The mixed-models ANOVA across the three measurement times revealed a significant main effect for time for both numeracy skills ($F(2, 816) = 915.98$, $p < .001$, $\eta_p^2 = .69$) and mathematical language ($F(2, 816) = 629.54$, $p < .001$, $\eta_p^2 = .61$), such that, overall, both performances increased over time. The interaction effect between time and group assignment was also significant for both numeracy skills ($F(4, 816) = 5.33$, $p < .001$, $\eta_p^2 = .03$) and mathematical language ($F(4, 816) = 3.51$, $p < .010$, $\eta_p^2 = .02$), indicating that, for both performances, the differences in their development depend on whether children were assigned to the math-specific or general intervention group or the control group (Figs. 2 and 3).

Moreover, the interaction effects between time and group assignment remain stable when controlling for the teacher's work experience (in years) as well as age and sex of the child and their family background characteristics (indicators: home language, mother's education level, number of siblings, family type), for both numeracy skills ($F(4, 816) = 5.02$, $p < .001$, $\eta_p^2 = .02$) and mathematical language ($F(4, 816) = 3.51$, $p < .010$, $\eta_p^2 = .02$).

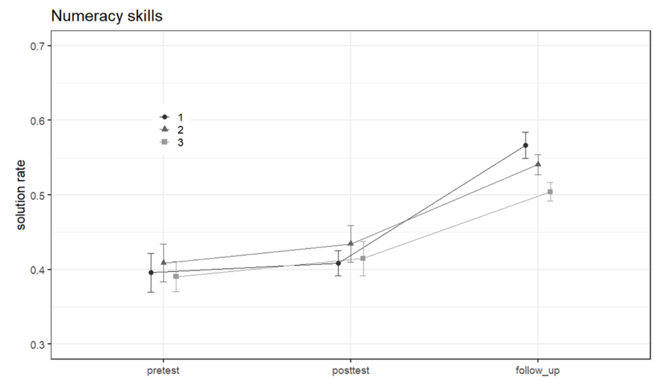

Fig. 2. Solution rates of numeracy skills by time and group.
Notes. 1 = IG1 (math), 2 = IG2 (general), 3 = CG (control)

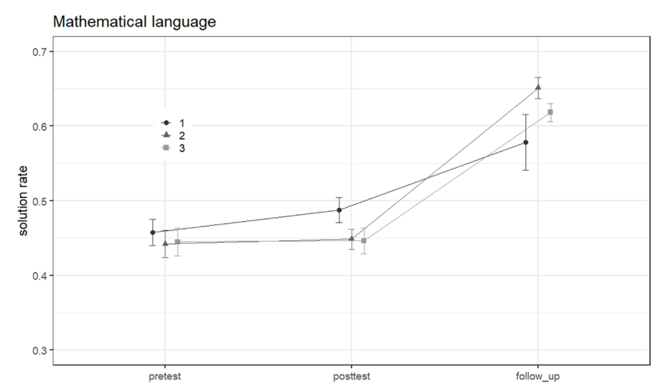

Fig. 3. Solution rates of math languages by time and group.
Notes. 1 = IG1 (math), 2 = IG2 (general), 3 = CG (control)

4.2. Pretest—posttest differences (training effect)

The mixed-models ANOVA across pre- and posttest revealed a significant main effect for time for both numeracy skills ($F(1, 408) = 31.60$, $p < .001$, $\eta_p^2 = .07$) and mathematical language ($F(1, 408) = 21.42$, $p < .001$, $\eta_p^2 = .05$), such that, overall, both performances increased from pretest to posttest. The interaction effect between time and group assignment was also significant for both numeracy skills ($F(2, 408) = 4.02$, $p < .050$, $\eta_p^2 = .03$) and mathematical language ($F(2, 408) = 4.87$, $p < .010$, $\eta_p^2 = .02$), indicating that, for both performances, the differences in development depend on whether children were assigned to the math-specific or general intervention group or the control group. The subsequently conducted a-priori contrasts revealed that the children in the math-specific intervention group, but not the children in the general intervention group, demonstrated higher performances at the posttest than the children in the control group (see Table 2). This holds for both numeracy skills (difference = 0.06, $SE = 0.02$, $p < .010$, $\eta_p^2 = .03$) and mathematical language (difference = 0.08, $SE = 0.02$, $p < .001$, $\eta_p^2 = .05$). Comparing the two intervention groups, the a-priori contrasts indicate that the children in the math-specific intervention group performed better than the children in the general intervention group. This holds for both numeracy skills (difference = 0.03, $SE = 0.02$, $p < .100$, $\eta_p^2 = .01$) and mathematical language (difference = 0.06, $SE = 0.02$, $p < .010$, $\eta_p^2 = .03$).

4.3. Pretest—follow-up test differences (long-term training effect)

The mixed-models ANOVA across pre- and follow-up test revealed a significant main effect for time for both numeracy skills ($F(1, 408) = 1,465.92$, $p < .001$, $\eta_p^2 = .78$) and mathematical language ($F(1, 408) =$

Table 2
Overview of performance (solution rate and standard error) on the dependent variables in the intervention groups and across the three measurement time points.

	Intervention group	Pretest	Posttest	Follow-up
Numeracy	Math group	.28 (.12)	.34 (.14)	.60 (.15)
	General group	.26 (.14)	.30 (.14)	.56 (.13)
	Control group	.27 (.12)	.28 (.13)	.51 (.13)
Math language	Math group	.43 (.14)	.49 (.13)	.72 (.17)
	General group	.41 (.15)	.44 (.14)	.70 (.11)
	Control group	.41 (.16)	.41 (.14)	.64 (.16)
General language comprehension	Math group	.29 (.14)	.38 (.17)	.44 (.15)
	General group	.29 (.15)	.42 (.17)	.45 (.15)
	Control group	.28 (.15)	.35 (.16)	.40 (.14)

872.95, $p < .001$, $\eta_p^2 = .68$), such that, overall, both performances increased from pretest to follow-up test. The interaction effect between time and group assignment was also significant for both numeracy skills ($F(2, 408) = 9.89$, $p < .001$, $\eta_p^2 = .05$) and mathematical language ($F(2, 408) = 4.22$, $p < .010$, $\eta_p^2 = .02$), indicating that, for both performances, the differences in development depend on whether children were assigned to the math-specific or general intervention group or the control group. For follow-up, the a-priori contrasts revealed that both the children in the math-specific intervention group and the children in the general intervention group demonstrated higher performances than the children in the control group. This holds for both, numeracy skills (math vs. control: difference = 0.09, $SE = 0.02$, $p < .001$, $\eta_p^2 = .06$; general vs. control: difference = 0.06, $SE = 0.02$, $p < .001$, $\eta_p^2 = .03$) and mathematical language (math vs. control: difference = 0.08, $SE = 0.02$, $p < .001$, $\eta_p^2 = .05$; general vs. control: difference = 0.06, $SE = 0.02$, $p < .010$, $\eta_p^2 = .02$). Comparing the two intervention groups, the a-priori contrasts indicate that the children in the math-specific intervention group demonstrated marginally higher numeracy skills (difference = 0.03, $SE = 0.02$, $p < .100$, $\eta_p^2 = .01$) but not mathematical language (difference = 0.02, $SE = 0.02$, $p = .221$, $\eta_p^2 < .00$) than the children in the general intervention group.

4.4. Posttest—follow-up test differences (stability of training effect)

The mixed-models ANOVA across post- and follow-up test revealed a significant main effect for time for both numeracy skills ($F(1, 408)$= 1,081.10, $p < .001$, $\eta_p^2 = .73$) and mathematical language ($F(1, 408) = 763.59$, $p < .001$, $\eta_p^2 = .65$), such that, overall, both performances increased from posttest to follow-up test. However, the interaction effect between time and group assignment was not significant for either numeracy skills ($F(2, 408) = 1.90$, $p = .152$, $\eta_p^2 = .01$) or mathematical language ($F(2, 408) = 1.79$, $p = .169$, $\eta_p^2 = .01$), indicating that between posttest and follow-up the children in all groups demonstrated comparable development.

4.5. Specificity of the training on math-specific interaction quality

To examine whether the math-specific training of childcare teachers specifically favored children's numeracy skills and their mathematical language in contrast to general language performance, we conducted mixed model ANOVAs using children's general language comprehension as outcome. The mixed-models ANOVA revealed a significant main effect for time ($F(2, 816) = 142.19$, $p < .001$, $\eta_p^2 = .26$), such that, overall, general language comprehension increased over time. The interaction effect between time and group assignment was also significant ($F(4, 816) = 2.46$, $p < .050$, $\eta_p^2 = .01$), indicating that the differences in development depend on whether children were assigned to the math-specific or general intervention group or to the control group (Fig. 4).

The a-priori contrasts revealed that at posttest the children in the general intervention group, but not the children in the math-specific intervention group, demonstrated higher performances than the children in the control group (difference = 0.07, $SE = 0.02$, $p < .010$, $\eta_p^2 = .03$). For follow-up, the a-priori contrasts revealed that both the children in the math-specific group (difference = 0.03, $SE = 0.02$, $p < .100$, $\eta_p^2 = .01$) and the children in the general group (difference = 0.05, $SE = 0.02$, $p < .010$, $\eta_p^2 = .02$) demonstrated higher performances than the children in the control group. Comparing the two intervention groups, the a-priori contrasts indicate that the children in the general intervention group demonstrated higher general language comprehension than the children in the math-specific intervention group at posttest (difference = 0.04, $SE = 0.02$, $p < .050$, $\eta_p^2 = .01$) but not at follow-up (difference = 0.01, $SE = 0.02$, $p = .454$, $\eta_p^2 < .00$).

The slightly nested structure (children in different early childcare centers) within the data (ICCs range between 0.09 and 0.18) was addressed by conducting sensitivity analyses. A simple regression analysis was performed in which the dummy-coded grouping variable was regressed on the respective outcome for posttest and follow-up testing, while clustering for the early childcare centers. The main results are consistent with those of the mixed-ANOVA and can be found in the supplement (Table S9).

5. Discussion

In this intervention study, we explored a new approach on how the participation of teachers in PD training on math-specific or general interaction quality may contribute to the early development of numeracy skills and mathematical language. Previous studies investigating the effects of high-quality sensitive and cognitive-verbal stimulating interactions, which can be shaped in activities like dialogic reading or playing board games, on the development of numeracy skills and mathematical language in children relate to investigations with parents or with specially trained research assistants rather than teachers (e.g., Purpura, Napoli, et al., 2017; Wijns et al., 2023). Moreover, those studies examined numeracy skills and mathematical language of children aged three years and older rather than including younger children. This study is the first both to include younger children from the age of two and to investigate whether participation in a PD training on math-specific interaction quality, which was specifically designed for teachers working with children between the ages of two and four, promotes children's numeracy skills and mathematical language. We particularly focused on training in math-specific interaction quality and training in general interaction quality, respectively, to a control group without training. We hypothesized that math-specific training of teachers is particularly effective and may be even needed to promote children's numeracy skills and their mathematical language. Therefore, we not only compared the development of the children in the math-specific group with the development of the children in the control group, but also with the development of the children in the general

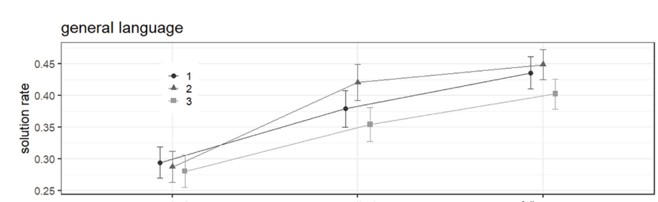

Fig. 4. Solution rates of general language comprehension by time and group. Notes. 1 = IG1 (math), 2 = IG2 (general), 3 = CG (control)

intervention group. This allows us to draw conclusions about the specificity of the effects from participating in a math-specific interaction training.

Overall, we found that children in all groups improved both their numeracy skills and their mathematical language over time. This is in line with other studies reporting age-related changes in math skills across preschool age (e.g., Gasteiger & Moeller, 2021). The present study highlights that these developments are significantly influenced by early domain-specific interaction quality. The effects of PD training in terms of the children's numeracy skills and mathematical language were significantly more striking in the group in which the teachers received math-specific training. As expected, the pretest-posttest comparisons revealed that the children in the math-specific intervention group, but not the children in the general intervention group, demonstrated higher numeracy skills and more advanced mathematical language than the children in the control group at post-test, hinting to a specific effect on math development. This specific training effect becomes even more evident as the comparison between the two intervention groups demonstrated that at posttest the children in the math-specific intervention group performed better in both numeracy skills and mathematical language compared to children in the general intervention group. This is consistent with Payne (2024) results, in which four-year-old children in two intervention groups played identical games, with one group also focusing on supporting the children's mathematical language during play. This math-language group achieved the greatest gains in numeracy skills, although non-significant. The present study goes beyond these findings, as it realized a follow up measure, which indicates, that even one year later, both the children in the math-specific and the general intervention group demonstrated higher numeracy skills and higher mathematical language compared to the children in the control group. Even more, children in the math specific intervention group outperformed children in the general intervention group in terms of their numeracy skills. However, one might say, that the math specific training is limited to promotion in numeracy only, as with respect to overall language development the general PD training, including general cognitive-verbal stimulation with a focus on various educational topics showed the assumed effect at posttest: the children in the general intervention group, but not the children in the mathematical intervention group, demonstrated higher general language comprehension than the children in the control group. Practicing high-quality interactions during dialogic reading and playing games throughout various educational domains (general intervention group) seems to boost general language development more compared to practicing such high-quality interactions only with the restricted mathematical focus (mathematical intervention group). It might be, that teachers in the general intervention group increased their variety of verbal input—-which has been shown to be of crucial importance in children's language development (e.g., Anderson et al., 2021)—more than teachers in the mathematical intervention group which in turn resulted in higher gains in language development. However, this effect seems to be short-term, as the differences between the math-specific and the general intervention group concerning general language development disappeared. Even more, children in the math-specific intervention group catch up in terms of their general language ability so that differences to the control group become apparent. One explanation for these findings might be that it is not sufficient to provide professionals with theoretical information on how to design high-quality interactions, but that practical training and reflection on their own practices are important for the transfer into everyday pedagogical practice. Moreover, the specific differences between the two intervention groups support the assumption that specific training is important. Yet, we cannot rule out that the performance differences between the intervention and control groups could also result from the different duration of the intervention (multi-day training for the IGs vs. 2-hour information session for CG). However, since the math-specific intervention group had the highest effect sizes, even compared to the general intervention group, which was the same length, it seems very unlikely that a very non-specific training of the same length as the other interventions would have had the same effects.

In sum, we could conclude that the interventions have an effect on their particular outcome—short-term: Training on math-specific interaction quality influences both numeracy skills and mathematical language, while training on general interaction quality particularly influences the children's general language comprehension. However, in the longer term, we find a slight favor of the math-specific intervention as it boosts math specific outcomes *and* language development compared to settings without the PD training. Please note, however, that the comparison between the two intervention groups conducted in this study was a rigorous test, as the general training also covered mathematical content to a small extent. This overlap in content may be the reason why the effects in terms of numeracy skills were only marginally significant between the two IGs at both measurement times.

One explanation that the children in the general intervention group caught up between post- and follow-up testing in numeracy skills and mathematical language could be that the training on general interaction quality was a rather indirect form of support. The teachers were also trained on the importance of sensitivity and cognitive-verbal stimulation while reading picture books or playing board games, but in contrast to the math-specific group, they were trained on general topics (environment and nature; mathematics; physical activity; social relationships) of child development rather than exclusively on math-specific topics (size and measurement; counting, numbers and arithmetic; space, shape and time; patterns and relationships) as in the math-specific group, which resulted in an increase in general language comprehension for the children in the general intervention group. Since general language and math skills are closely related, especially at around the age of three (Peng et al., 2020), the general language advantage that the children in the general intervention group had may have helped them to access math content more quickly in the long term, which may also have contributed to an improvement in numerical skills and math language. This assumption would tie in with Cunha and Heckmann (2007) hypotheses, that skills are self-reinforcing. Moreover, there is also evidence that the links between mathematical and general language skills are not exclusively direct, but rather mediated by mathematical language (e.g., Purpura, Logan, et al., 2017). Applying this approach, the increase in numeracy skills among the children in the math-specific interaction group led to an increase in mathematical language (additionally promoted by the mathematical language input of the teachers), which in turn influenced the children's general language comprehension in the long term. Beyond the explanatory mechanisms that relate to children's skills development, we also know that teachers' participation in our intervention study had a positive influence both on the math-specific interaction quality during free-play observed by observers at posttest and on the math-specific interaction quality rated by the teachers themselves (Baron et al., 2025). This was particularly true for the teachers who were in the math-specific intervention group. In other words, the teachers in the math-specific intervention group performed more math activities during free-play (e.g., playing games with a number cube) with the children at posttest compared to the teachers in the general and the control group. Therefore, the increase in math-specific interaction quality in the math-specific intervention group might be one of the mechanisms that contributed to an improvement of both numeracy skills and mathematical language, particularly for the children in the math-specific intervention group. However, more research is needed to draw conclusions about these mechanisms. Therefore, we are currently analyzing observational data from the pre- and posttest to confirm whether and how well the teachers have implemented what they have learned in the training sessions.

Our findings indicate an opportunity to promote children's numeracy skills and (mathematical) language from a young age. Such skills, in turn, form the basis for their later mathematical learning and can thus prepare children for later challenges right from the start. The PD described in this study is a comparatively short training of qualified teachers, which results in promoting effects in both the numeracy skills and mathematical language of young children. Considering that the professionalization of teachers also includes a continuous learning commitment, including PD (NAEYC & NCTM, 2010), such training programs can be supportive in improving the interaction quality of teachers in order to provide children with high-quality learning experiences. However, a country-specific focus must be taken into account in any transfer; in Germany, for example, depending on the federal state, obligatory curricula for working with children (usually between the ages of 3 and 6) exist for state-funded ECEC centers, but there is no formal instruction, which has been considered in our training.

A limitation of the results might be that the majority of the children changed their teachers before the follow-up test took place. The quality of interaction of the subsequent teachers was not controlled for in this study. However, any differences in the quality of interaction between the intervention groups that might emerge would tend to be random and not systematic. Instead, they would tend to underestimate the long-term effects of the intervention to some extent. However, the change of teacher may not only be a limitation but also a supplementary explanation for the result that we found specific group differences after the intervention between the math group and the general group, both of which diminished at follow-up testing for mathematical language and general language comprehension, while maintaining the differences with the control group. A further limitation of the study may be that, although we have used several implementation tools, like self-reflection and newsletter, we cannot be certain which implementation by the teachers has led to the different effects. However, a recently published meta-analysis found that informal math-specific interventions were more effective when they included caregiver training and a supportive implementation phase (e.g., short reminders) (Nelson et al., 2024). Further studies may address how different implementation strategies might relate to the development of numeracy skills and mathematical language. Another limitation might be that an above-average number of parents of the participating children held a college degree. Considering the influence of family background on the development of children's competencies, this may result in a bias. As we investigated the influence of PD training for teachers, the children were recruited exclusively via the participating teachers. This may have led to particularly motivated high-educated parents giving their consent for their children to participate.

In conclusion, our study indicates that the participation of teachers in a PD on math-specific interaction quality—combining sensitivity and cognitive-verbal stimulation as cornerstones with dialogic reading, playing board games and using mathematical language as powerful interaction settings and strategies—promotes the development of numeracy skills and mathematical language in two- to four-year-old children and thus forms a strong foundation for the children's later mathematical learning.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Grant 416593881.

CRediT authorship contribution statement

Nadine Besser: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis. **Anja Linberg:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Dorothea Dornheim:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Sabine Weinert:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Hans-Günther Roßbach:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Simone Lehrl:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization.

Declaration of interests

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecresq.2025.02.004.

Data availability

The data that has been used is confidential.

References

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development, 92*(2), 484–501. https://doi.org/10.1111/cdev.13508

Archambault, L., Leary, H., & Rice, K. (2022). Pillars of online pedagogy: A framework for teaching in online learning environments. *Educational Psychologist, 57*(3), 178–191. https://doi.org/10.1080/00461520.2022.2051513

Bücklein, C., Hoffer, R., & Strohmer, J. (2017). Measuring the quality of interactions in early childhood education and care settings for toddlers – an exploratory comparison of the observation tools GInA and CLASS Toddler. In H. Wadepohl, K. Mackowiak, K. Froehlich-Gildhoff, & D. Weltzien (Eds.), *Interaktionsgestaltung in Familie und Kindertagesbetreuung [Shaping interactions in families and childcare centers]* (pp. 83–114). Springer Fachmedien. https://doi.org/10.1007/978-3-658-10276-0_4.

Baron, F., Linberg, A., Dornheim, D., & Lehrl, S. (2025). Improving global and math-specific teacher-toddler interactions through an intervention for early childcare teachers: The role of activity settings. *Early Childhood Research Quarterly, 70*, 79–90. https://doi.org/10.1016/j.ecresq.2024.09.003

Beckh, K., & Becker-Stoll, F. (2016). Formations of attachment relationships towards teachers lead to conclusions for public child care. *International Journal of Developmental Science, 10*(3–4), 99–106. https://doi.org/10.3233/DEV-16197

Bornstein, M. H. (2017). The specificity principle in acculturation science. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 12*(1), 3–45. https://doi.org/10.1177/1745691616655997

Brunsek, A., Perlman, M., McMullen, E., Falenchuk, O., Fletcher, B., Nocita, G., Kamkar, N., & Shah, P. S. (2020). A meta-analysis and systematic review of the associations between professional development of early childhood educators and children's outcomes. *Early Childhood Research Quarterly, 53*, 217–248. https://doi.org/10.1016/j.ecresq.2020.03.003

Burchinal, M., Vernon-Feagans, L., Vitiello, V., & Greenberg, M. (2014). Thresholds in the association between child care quality and child outcomes in rural preschool children. *Early Childhood Research Quarterly, 29*(1), 41–51. https://doi.org/10.1016/j.ecresq.2013.09.004

Buysse, V., Winton, P. J., & Rous, B. (2009). Reaching consensus on a definition of professional development for the early childhood field. *Topics in Early Childhood Special Education, 28*(4), 235–243. https://doi.org/10.1177/0271121408328173

Ceulemans, A., Titeca, D., Loeys, T., Hoppenbrouwers, K., Rousseau, S., & Desoete, A. (2015). The sense of small number discrimination: The predictive value in infancy and toddlerhood for numerical competencies in kindergarten. *Learning and Individual Differences, 39*, 150–157. https://doi.org/10.1016/j.lindif.2015.03.009

Chen, H., Ning, X., Wang, L., & Yang, J. (2018). Acquiring new factual information: Effect of prior knowledge. *Frontiers in Psychology, 9*, 1734. https://doi.org/10.3389/fpsyg.2018.01734

Clements, D. H., & Samara, J. H. (2007). Effects of a preschool mathematics curriculum: summative research on the building blocks project. *Journal for Research in Mathematics Education, 38*(2), 136–163.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Costa, H. M., Outhwaite, L. A., & Van Herwegen, J. (2023). Early years practitioners' training, beliefs and practices concerning mathematics: implications for education and practice. *International Journal of Early Years Education, 0*(0), 1–18. https://doi.org/10.1080/09669760.2023.2254822

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review, 97*(2). https://doi.org/10.1257/aer.97.2.31. Article 2.

Dubowy, M., Ebert, S., Maurice, J., & Weinert, S. (2008). Sprachlich-kognitive kompetenzen beim eintritt in den kindergarten [Language and cognitive skills at the start of kindergarten]. *Zeitschrift Für Entwicklungspsychologie und Pädagogische Psychologie, 40*(3), 124–134. https://doi.org/10.1026/0049-8637.40.3.124

Egert, F., Fukkink, R., & Eckhardt, A. G. (2018). Impact of In-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of Educational Research, 88*(3), 401–433. https://doi.org/10.3102/0034654317751918

Fox-Boyer, A., Bäumer, T., Müller, M., & Merzbecher, S.. (2020). *TROG-D: Test zur Überprüfung des Grammatikverständnisses [Test to evaluate grammar comprehension]* (8. Auflage). Schulz-Kirchner Verlag.

Gasteiger, H., & Moeller, K. (2021). Fostering early numerical competencies by playing conventional board games. *Journal of Experimental Child Psychology, 204*, Article 105060. https://doi.org/10.1016/j.jecp.2020.105060

Grimm, H. (2016). SETK-2—Sprachentwicklungstest für zweijährige Kinder (2;0–2;11 Jahre). *Diagnose rezeptiver und produktiver Sprachverarbeitungsfähigkeiten (Language development test for two-year-old children (2;0-2;11 years). Diagnosis of receptive and productive language processing skills* (2nd ed.). Hogrefe.

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

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49–83). Paul H. Brookes Publishing Co.

Han, Y., & Ginsburg, H. P. (2001). Chinese and English mathematics language: The relation between linguistic clarity and mathematics performance. *Mathematical Thinking and Learning, 3*(2–3), 201–220. https://doi.org/10.1080/10986065.2001.9679973

Hornburg, C. B., Schmitt, S. A., & Purpura, D. J. (2018). Relations between preschoolers' mathematical language understanding and specific numeracy skills. *Journal of Experimental Child Psychology, 176*, 84–100. https://doi.org/10.1016/j.jecp.2018.07.005

Jenßen, L. (2021). A math-avoidant profession? Review of the current research about early childhood teachers' mathematics anxiety and empirical evidence. In S. Dunekacke, A. Jegodtka, T. Koinzer, K. Eilerts, & L. Jenßen (Eds.), *Early childhood teachers' professional competence in mathematics*. Routledge.

King, Y. A., & Purpura, D. J. (2021). Direct numeracy activities and early math skills: Math language as a mediator. *Early Childhood Research Quarterly, 54*, 252–259. https://doi.org/10.1016/j.ecresq.2020.09.012

Kluczniok, K., & Roßbach, H.-G. (2014). Conceptions of educational quality for kindergartens. *Zeitschrift für Erziehungswissenschaft, 17*(6), 145–158. https://doi.org/10.1007/s11618-014-0578-2. Suppl.

Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction, 19*(6), 513–526. https://doi.org/10.1016/j.learninstruc.2008.10.002

Kuger, S., Kluczniok, K., Kaplan, D., & Rossbach, H.-G. (2016). Stability and patterns of classroom quality in German early childhood education and care. *School Effectiveness and School Improvement, 27*(3), 418–440. https://doi.org/10.1080/09243453.2015.1112815

Kung, M., Schmitt, S. A., Zhang, C., Whiteman, S. D., Yang, F., & Purpura, D. J. (2019). The role of mathematical language in mathematics development in China and the US. *International Journal of Educational Research, 95*, 131–142. https://doi.org/10.1016/j.ijer.2019.02.008

La Paro, K. M., Hamre, B. K., & Pianta, R. C.. (2012). *Classroom assessment scoring system (CLASS), manual toddler*. Paul H. Brookes Publishing Co.

Lehrl, S., Kluczniok, K., & Rossbach, H.-G. (2016). Longer-term associations of preschool education: The predictive role of preschool quality for the development of mathematical skills through elementary school. *Early Childhood Research Quarterly, 36*, 475–488. https://doi.org/10.1016/j.ecresq.2016.01.013

Lehrl, S., Ebert, S., Blaurock, S., Rossbach, H.-G., & Weinert, S. (2020). Long-term and domain-specific relations between the early years home learning environment and students' academic outcomes in secondary school. *School Effectiveness and School Improvement, 31*(1), 102–124. https://doi.org/10.1080/09243453.2019.1618346

Lehrl, S., Linberg, A., Dornheim, D., Besser, N., Baron, F., Klemm, J., Bauer, N., Roßbach, H.-G., & Weinert, S. (2024). Frühe mathematische Entwicklung und die Bedeutung von Interaktionsqualität in Kindertageseinrichtungen: Ein Methodenbericht zur Konzeption und Umsetzung des EarlyMath Projekts. OSF. https://doi.org/10.31219/osf.io/wy9um.

Linberg, A., Lehrl, S., Dornheim, D., Weinert, S., & Roßbach, H.-G. (2020). *Mathematical Development and the Impact of Interactionquality in Early Childcare: Frühe mathematische Entwicklung und die Bedeutung von Interaktionsqualiät in Kindertageseinrichtungen*.

LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly, 22*(1), 3–17. https://doi.org/10.1016/j.ecresq.2006.05.001

Major, L., & Watson, S. (2018). Using video to support in-service teacher professional development: The state of the field, limitations and possibilities. *Technology, Pedagogy and Education, 27*(1), 49–68. https://doi.org/10.1080/1475939X.2017.1361469

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development, 79*(3), 732–749. https://doi.org/10.1111/j.1467-8624.2008.01154.x

NAEYC, & NCTM. (2010). Early childhood mathematics: Promoting good beginnings. *A joint position statement of the National Association for the Education of Young Children (NAEYC) and the National Council of Teachers of Mathematics (NCTM)*. https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/psmath.pdf.

Nelson, G., Carter, H., Boedeker, P., Knowles, E., Buckmiller, C., & Eames, J. (2024). A meta-analysis and quality review of mathematics interventions conducted in informal learning environments with caregivers and children. *Review of Educational Research, 94*(1), 112–152. https://doi.org/10.3102/00346543231156182

Nguyen, T., Watts, T. W., Duncan, G. J., Clements, D. H., Sarama, J. S., Wolfe, C., & Spitler, M. E. (2016). Which preschool mathematics competencies are most predictive of fifth grade achievement? *Early Childhood Research Quarterly, 36*, 550–560. https://doi.org/10.1016/j.ecresq.2016.02.003

Niklas, F., & Schneider, W. (2017). Home learning environment and development of child competencies from kindergarten until the end of elementary school. *Contemporary Educational Psychology, 49*, 263–274. https://doi.org/10.1016/j.cedpsych.2017.03.006

OECD. (2022). *Education at a glance 2022: OECD indicators*. OECD Publishing.

OECD. (2023). *PISA 2022 assessment and analytical framework*. OECD. https://doi.org/10.1787/dfe0bf9c-en

Payne, T. (2024). The math talk learning environment: Testing an early childhood math intervention. *Early Childhood Research Quarterly, 66*, 224–233. https://doi.org/10.1016/j.ecresq.2023.10.012

Peng, P., Lin, X., Ünal, Z. E., Lee, K., Namkung, J., Chow, J., & Sales, A. (2020). Examining the mutual relations between language and mathematics: A meta-analysis. *Psychological Bulletin, 146*(7), 595–634. https://doi.org/10.1037/bul0000231

Pillinger, C., & Vardy, E. J. (2022). The story so far: A systematic review of the dialogic reading literature. *Journal of Research in Reading, 45*(4), 533–548. https://doi.org/10.1111/1467-9817.12407

Purpura, D. J., & Logan, J. A. R. (2015). The nonlinear relations of the approximate number system and mathematical language to early mathematics development. *Developmental Psychology, 51*, 1717–1724. https://doi.org/10.1037/dev0000055

Purpura, D. J., & Lonigan, C. J. (2013). Informal numeracy skills: The structure and relations among numbering, relations, and arithmetic operations in preschool. *American Educational Research Journal, 50*(1), 178–209. https://doi.org/10.3102/0002831212465332

Purpura, D. J., & Reid, E. E. (2016). Mathematics and language: Individual and group differences in mathematical language skills in young children. *Early Childhood Research Quarterly, 36*, 259–268. https://doi.org/10.1016/j.ecresq.2015.12.020

Purpura, D. J., Logan, J. A. R., Hassinger-Das, B., & Napoli, A. R. (2017). Why do early mathematics skills predict later reading? The role of mathematical language. *Developmental Psychology, 53*(9), 1633–1642. https://doi.org/10.1037/dev0000375

Purpura, D. J., Napoli, A. R., Wehrspann, E. A., & Gold, Z. S. (2017). Causal connections between mathematical language and mathematical knowledge: A dialogic reading intervention. *Journal of Research on Educational Effectiveness, 10*(1), 116–137. https://doi.org/10.1080/19345747.2016.1204639

Purpura, D. J., Litkowski, E. C., & Knopik, V. (2019). Mathematics and reading develop together in young children: Practical and policy considerations. *Policy Insights from the Behavioral and Brain Sciences, 6*(1), 12–20. https://doi.org/10.1177/2372732218815923

Purpura, D. J., O'Rear, C. D., Ellis, A., Logan, J. A. R., Westerberg, L., Ehrman, P., King, Y. A., Vander Tuin, M., Nordgren, I., Anderson, K., Cosso, J., Zippert, E., Napoli, A. R., Hornburg, C. B., Schmitt, S. A., & Dobbs-Oates, J. (2023). Unique and combined effects of quantitative mathematical language and numeracy instruction within a picture book intervention: A registered report. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000820

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375–394.

Ropohl, M., Nielsen, J. A., Olley, C., Rönnebeck, S., & Stables, K. (2018). The concept of competence and its relevance for science, technology and mathematics education. In J. Dolin, & R. Evans (Eds.), *Transforming assessment: through an interplay between practice, research and policy* (pp. 3–25). Springer International Publishing. https://doi.org/10.1007/978-3-319-63248-3_1.

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition, 108*(3), 662–674. https://doi.org/10.1016/j.cognition.2008.05.007

Seitz, M., & Weinert, S. (2022). Numeracy skills in young children as predictors of mathematical competence. *British Journal of Developmental Psychology, 40*(2), 224–241. https://doi.org/10.1111/bjdp.12408

Siegler, R., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*(3), 545–560. https://doi.org/10.1037/a0014239

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (.2012). *afex: Analysis of factorial experiments* (p. 1.3-1) [Dataset]. https://doi.org/10.32614/CRAN.package.afex.

Siraj, I., Melhuish, E., Howard, S. J., Neilsen-Hewett, C., Kingston, D., De Rosnay, M., Huang, R., Gardiner, J., & Luu, B. (2023). Improving quality of teaching and child development: A randomised controlled trial of the leadership for learning intervention in preschools. *Frontiers in Psychology, 13*. https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1092284.

Snyder, P., Hemmeter, M. L., Meeker, K. A., Kinder, K., Pasia, C., & McLaughlin, T. (2012). Characterizing key features of the early childhood professional development literature. *Infants & Young Children, 25*(3), 188. https://doi.org/10.1097/IYC.0b013e31825a1ebf

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2011). Pre-school quality and educational outcomes at age 11: Low quality has little benefit. *Journal of Early Childhood Research, 9*(2), 109–124. https://doi.org/10.1177/1476718X10387900

Toll, S. W. M., & van Luit, J. E. H. (2014). The developmental relationship between language and low early numeracy skills throughout kindergarten. *Exceptional Children, 81*(1), 64–78. https://doi.org/10.1177/0014402914532233

Towson, J. A., Macy, M., Abarca, D. L., Myers, K., & FitzPatrick, E. (2023). Examining teachers' use of dialogic reading strategies following a multiple component professional development intervention. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-023-01526-3

Trawick-Smith, J., Swaminathan, S., & Liu, X. (2016). The relationship of teacher–child play interactions to mathematics learning in preschool. *Early Child Development and Care, 186*(5), 716–733. https://doi.org/10.1080/03004430.2015.1054818

Turan, E., & De Smedt, B. (2022). Mathematical language and mathematical abilities in preschool: A systematic literature review. *Educational Research Review, 36*, Article 100457. https://doi.org/10.1016/j.edurev.2022.100457

Weinert, S., & Ebert, S. (2013). Spracherwerb im Vorschulalter: soziale disparitäten und einflussvariablen auf den grammatikerwerb [Language acquisition at preschool age: social disparities and influencing variables on grammar acquisition]. *Zeitschrift für Erziehungswissenschaft, 16*(2), 303–332. https://doi.org/10.1007/s11618-013-0354-8

Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30*(5), 679–689. https://doi.org/10.1037/0012-1649.30.5.679

Wijns, N., Purpura, D., & Torbeyns, J. (2023). Stimulating preschoolers' repeating patterning ability by means of dialogic picture book reading. *Journal of Educational Psychology, 115*(5), 732–746. https://doi.org/10.1037/edu0000756

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*(2), 155–193. https://doi.org/10.1016/0010-0277(90)90003-3

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P