EARLY EDUCATION AND DEVELOPMENT
2026, VOL. 37, NO. 2, 492–517
https://doi.org/10.1080/10409289.2025.2571977

Instructional Talk in Early Science Education: How Teachers' Support and Children's Verbal Participation Shape Language Development and Learning About Materials

Ada Cecil Haen [a], Mirjam Steffensky [a], Anika Bürgermeister [b], Ilonca Hardy [c], Miriam Leuchter [d], Franziska Bednarski[b], Sina-Kristin Koschick [c], Lukas Lazzara [d], and Henrik Saalbach [b]

[a]Department of Social Sciences, Mathematics and Natural Sciences Education, University of Hamburg; [b]Department of Educational Psychology with a Focus on Teaching, Learning, and Development, University of Leipzig; [c]Institute of Early and Primary Education, Goethe-University Frankfurt; [d]Department of Educational Sciences, RPTU University of Kaiserslautern-Landau

ABSTRACT
The study investigates instructional support and children's verbal participation in video-recorded science activities. Using multilevel regression, a longitudinal analysis of data from 51 German preschool groups (188 children aged 4–6) was conducted to examine how Concept Development and Language Modeling (two dimensions of instructional support according to the Classroom Assessment Scoring System; Pianta et al. 2008), along with children's verbal participation, influenced language development and science knowledge related to the topic of materials. *Research Findings*: Prior language skills predicted growth in both domains; prior science knowledge influenced science outcomes. Unexpectedly, preschool teachers' Concept Development and Language Modeling did not predict outcomes, revealing limitations in instructional support practices. However, the length of children's verbal contributions positively affected their individual language development, while exploratory analyses indicated negative effects of turn frequency. A positive interaction effect between turn length and Concept Development on science outcomes was not confirmed in models using imputed data. *Practice or Policy*: Findings challenge assumptions about instructional support efficacy, pointing to the potential to integrate participation with scaffolding in early childhood education research. Given the low variance and limited instructional support quality, this research provides a foundation for refining pedagogical strategies and advancing the understanding of early childhood education.

Introduction

The foundation for educational success during formal schooling is laid in early childhood, a period that is highly predictive of children's long-term individual development (Camilli et al., 2010; Sylva et al., 2004). Language acquisition is central to this developmental process, as it serves both communicative functions in social interactions and as a cognitive tool for thinking and constructing complex, domain-specific knowledge structures (Kempert et al., 2019; Tomasello, 1999). This is also relevant for the development of children's science knowledge (e.g., Haug and Ødegaard 2014; Henrichs and Leseman 2014), which is the context of this study. Science education holds significant importance in

CONTACT Ada Cecil Haen ada.haen@uni-hamburg.de Faculty of Education, University of Hamburg, Von-Melle-Park 8, Hamburg, 20146, Germany.

This article has been corrected with minor changes. These changes do not impact the academic content of the article.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/10409289.2025.2571977.

early childhood, as children's everyday experiences with science phenomena lead them to develop a variety of their own explanations about the world (Eshach & Fried, 2005). Beyond these initial ideas, research shows that children can acquire science knowledge and skills from an early age (e.g., Gropen et al., 2017; Zimmerman & Klahr, 2018). Early science knowledge lays the foundation for understanding more complex scientific concepts in later educational stages and serves as a predictor for science achievement in primary school (Morgan et al., 2016).

Building upon the interplay between language and scientific knowledge development, the quality of early childhood education and care (ECEC) emerges as a key factor in supporting these developmental processes (Howard et al., 2024; Sylva et al., 2004). In this regard, the quality of teacher-child interaction plays a pivotal role, as developmental processes and learning are integrated into the social context and are driven by active participation and interactions (Callanan & Valle, 2008). The quality of social interactions is frequently conceptualized through three central domains (e.g., Pianta & Hamre, 2009): emotional and social dimensions, informed by attachment theory and self-determination theory; the organizational structures that underpin effective interactions; and cognitive aspects, rooted in co-constructivist learning theory (Vygotsky, 1978) and closely linked to the concept of *Scaffolding* (Wood et al., 1976). In this context, the child is not perceived as a passive recipient of the preschool teacher's instructions, but rather as an active participant whose engagement in interactions substantially influences learning outcomes. The multilevel supply-use model highlights the complex interplay of factors that influence learning (Brühwiler & Blatchford, 2011). One important aspect in this context is children's engagement with learning opportunities, which can be conceptualized in terms of emotional, cognitive, and behavioral dimensions (Fredricks et al., 2004). This engagement is regarded as an indicator of the various factors influencing how children take up and process learning opportunities. Verbal participation in discourse, for instance, is an observable indicator of behavioral engagement and has been shown to play an important role in individual learning and in relation to the effectiveness of scaffolding strategies (Bürgermeister et al., 2019; Herrmann et al., 2021; Mannel et al., 2016; Sedova et al., 2019).

While numerous studies have examined the impact of interaction quality on development in the context of literacy, numeracy, motor skills, and social-behavioral competencies (Suchodoletz et al., 2023), research on its influence on preschool children's science learning remains limited (Hu et al., 2017; Studhalter et al., 2021). Moreover, little is known about the interplay between children's participation in discourse and interaction quality, with a particular focus on teachers' instructional support (e.g., Bürgermeister et al., 2019). This study addresses this gap by investigating how instructional support (as defined by CLASS Pre-K, Pianta et al., 2008) and children's verbal participation influence the development of language skills and science knowledge. Employing a longitudinal study design and multilevel regression analyses, the research places particular emphasis on the interaction between these variables. By emphasizing the dynamic interplay between these factors, the study contributes to the ongoing discourse on the quality of ECEC in the context of early science education.

Instructional quality and its impact on children's development

Interaction quality in ECEC, as a key factor in supporting children's development and reducing educational disparities at school entry, has emerged as a critical focus in research (OECD, 2018).

Various conceptualizations of interaction quality agree on distinguishing three domains grounded in theoretical frameworks and empirical findings (Hofkens et al., 2023). 1) The emotional and social component of interactions is grounded in attachment theory. This perspective emphasizes that children need a familiar, secure environment and a trusting relationship with their parents (or a preschool teacher in institutional settings) to take risks and actively explore their environment (Bowlby, 1982). Additionally, self-determination theory highlights the fundamental needs for autonomy, competence, and relatedness as essential for intrinsic motivation to engage with a subject (Ryan & Deci, 2000). 2) The organizational domain of interaction quality relates to teachers' actions in structuring the learning environment. These include facilitating children's access to learning opportunities, promoting engagement, and implementing strategies to prevent or appropriately address disruptive behavior (Evertson & Weinstein, 2011). Empirical evidence shows a positive effect of structured learning environments on children's engagement and academic outcomes (Patall et al., 2024). 3) The cognitive domain, examined in relation to co-constructivist learning theory (Vygotsky, 1978), is closely connected to the concept of *Scaffolding*. Derived from the *Contingent Shift Framework* by Wood et al. (1976), scaffolding encompasses a range of strategies implemented dynamically and adaptively by a competent interaction partner to support children's learning. In this context, interaction quality is not solely influenced by scaffolding strategies; the extent of the child's active engagement with these learning opportunities is also a significant factor for individual development. From a science-specific perspective, scaffolding is particularly important for fostering children's understanding of science concepts as well as process-related knowledge (e.g., Schmerse et al., 2024), as it helps children to actively align new information with their intuitive concepts and gradually develop these toward scientifically appropriate concepts (Vosniadou, 2007). These cognitive scaffolding strategies are operationalized in the internationally well-established *Classroom Assessment Scoring System* (CLASS PreK, Pianta et al., 2008) through the Instructional Support domain, which comprises three dimensions that align with the concept of scaffolding (van de Pol et al., 2010): 1) *Concept Development* focuses on promoting children's higher-order thinking and understanding, both domain-specific and on a meta-level. Strategies are outlined as being pivotal for the acquisition of domain-specific knowledge: These include, for example, asking explanatory questions, encouraging comparisons and classifications, activating prior knowledge, linking concepts, and making references to children's everyday lives (Pianta et al., 2008). Effective use of these strategies requires alignment with a child's zone of proximal development van de Pol et al. (2010), 2) *Quality of Feedback* includes providing guidance when children struggle to understand concepts, encouraging back-and-forth exchanges, prompting reasoning, offering explanations, and fostering sustained interest and persistence (Pianta et al., 2008). 3) *Language Modeling* encompasses various language-related scaffolding strategies (Gibbons, 2015) that can enhance both language skills and domain-specific learning. Key techniques include open-ended questions, expanding on children's verbal contributions, self- and parallel-talk, and introducing new and varied vocabulary, which is crucial in acquiring domain-specific vocabulary (Pianta et al., 2008). In addition to Instructional Support, the CLASS framework also incorporates two additional domains that capture further aspects of interaction quality: The *Emotional Support* domain refers to fostering a positive social climate among peers, teachers and children, as well as demonstrating sensitive and responsive teacher behavior. *Classroom Organization* focuses on structuring interactions, managing behavior, efficiently using learning time, and supporting children's interests.

International studies indicate substantial variation in interaction quality across the three CLASS domains in ECEC. The quality ratings for Emotional Support and Classroom Organization are predominantly in the medium to high range, while ratings for Instructional Support are typically in the low to low-medium range (Hofkens et al., 2023). The relatively low quality ratings for Instructional Support align with findings on the limited use of specific scaffolding strategies within preschool settings, for example in the context of science (e.g., Leuchter & Saalbach, 2014; Studhalter et al., 2021). However, some studies have shown that the quality ratings for Classroom Organization and Instructional Support tend to be higher in structured, domain-specific activities – such as math, social studies, and science – compared to less structured activities like play, art, and everyday routines (Hu et al., 2016; Thorpe et al., 2020).

Findings on the impact of interaction quality across the three CLASS domains on children's development are inconsistent, with effect sizes tending to be small (Suchodoletz et al., 2023). Several factors lead to these small and inconsistent effects. Burchinal (2017) highlights methodological challenges, such as contextual differences in ECEC settings across countries and regions, variations in children's responses to learning opportunities, and intra-institutional variability in interaction quality. The choice of study design and the handling of covariates also influence outcomes. Longitudinal designs using CLASS have shown reduced sensitivity in predicting learning progress, potentially due to controlling for covariates like prior knowledge. Conversely, cross-sectional studies that do not control for covariates may overestimate these effects (Wang et al., 2022). Additionally, limitations in assessing interaction quality further obscure consistent findings. Moreover, maintaining a minimum level of quality over a longer period appears crucial for a positive impact on children's learning (Levickis et al., 2023). Another question is whether domain-specific criteria are necessary for assessing interaction quality in domain-specific contexts. High global interaction quality ratings do not necessarily reflect high quality in specific domains (Levickis et al., 2023). The CLASS instrument, while designed to assess global interaction quality across multiple domains and provide a broad overview of children's typical interactional experiences, may not effectively capture domain-specific nuances. A more focused use of the CLASS instrument, by examining specific dimensions such as Language Modeling rather than the broader domain of Instructional Support, represents a pragmatic strategy to address this limitation and to better capture domain-specific aspects of interaction quality. In doing so, it may provide clearer evidence of how teacher – child interaction affects children's developmental outcomes, such as language acquisition.

Nevertheless, numerous studies have examined the effects of interaction quality (as measured with CLASS) on literacy, numeracy, motor skills, and social-behavioral competencies (Suchodoletz et al., 2023). Meta-analyses on the effects of interaction quality in ECEC consistently reveal small but significant effects on children's developmental outcomes. Suchodoletz et al. (2023), analyzing 185 studies (primarily from the U.S., with over 80 using CLASS), found that higher interaction quality is significantly associated with better child learning outcomes in literacy and numeracy, although the effect sizes are small. Similarly, Wang et al. (2022) found weak effects on children's pre-literacy skills in their meta-analyses focusing on studies using CLASS in literacy contexts. Ulferts et al. (2019), based on data from 17 European longitudinal studies with over 16,000 children, identified small but significant positive effects of global (measured using tools like CLASS) and domain-specific interaction quality. Their findings show that global interaction quality has a stronger effect on literacy outcomes, whereas domain-specific quality more strongly influences numeracy outcomes.

Taken together, these findings suggest nuanced but inconsistent evidence, particularly with regard to the cognitive domain of interaction quality. Variability in observed effects on learning outcomes may reflect methodological differences (e.g., cross-sectional vs. longitudinal designs, use of covariates), as well as conceptual limitations of high-inference tools such as CLASS in capturing domain-specific instructional practice, especially in science contexts.

However, children's development in early education is influenced not only by the instructional support provided by preschool teachers, but also by the extent to which children actively engage in and make use of these learning opportunities. Nevertheless, the interplay between children's verbal participation and the instructional support provided by preschool teachers on individual learning, for example in the science domain, has received little attention in research. Addressing this gap, this study investigates the impact of Instructional Support, as measured by CLASS Pre-K, on the development of children's language skills and science knowledge within early science settings. Rather than assessing Instructional Support broadly, the study focuses on the subdimensions of Concept Development and Language Modeling, which are closely linked to domain-specific scaffolding strategies. Furthermore, children's individual verbal participation is integrated into the analyses to emphasize their active role in their own learning processes.

Children's verbal participation and its impact on their learning and development

As previously mentioned, the understanding of learning as a co-constructive process emphasizes engagement in social interactions and dialogic learning for cognitive and language development (Vygotsky, 1978). In this context, children's engagement is not simply a matter of attention or motivation but can be conceptualized as a multidimensional construct. It comprises *behavioral*, *emotional*, and *cognitive* components (Fredricks et al., 2004). Behavioral engagement refers to observable involvement in learning activities, such as effort, task persistence, rule compliance, and verbal participation in classroom discourse. Emotional engagement captures affective responses to peers, teachers, and learning activities, while cognitive engagement involves the mental investment and use of strategies to understand and master complex content (Fredricks et al., 2004).

Verbal participation in classroom discourse has long been a focus in educational research, particularly in studies examining teacher-student interaction (Mercer & Dawes, 2014). The present study builds on previous research and focuses specifically on quantifiable aspects of children's verbal participation (number and length of speaker turns) as indicators of children's behavioral engagement in early science activities. This quantitative approach facilitates systematic comparisons across children and settings. Moreover, previous research has highlighted that children's verbal participation is a crucial source of information for teachers, enabling them to adapt their scaffolding to the needs of individual children (Bürgermeister et al., 2019). This further justifies our focus on the extent of verbal participation. While qualitative dimensions of verbal participation, such as dialogic features, reciprocity, and the differentiation between on- and off-task contributions, are also important discourse markers, these aspects are not the focus of this study.

Co-constructivist concepts, such as *Sustained Shared Thinking*, also emphasize the involvement of both the child and the teacher in verbal interaction as a prerequisite for a cooperative problem-solving dialogue (Siraj-Blatchford et al., 2002). The perspective of the multilevel supply-use model, which describes the complex interplay of factors influencing a child's learning (Brühwiler & Blatchford, 2011), highlights active verbal participation as an observable indicator of their engagement with the learning opportunities provided by preschool teachers. This is supported by findings from a study by Sedláček and Šeďova (2020), which demonstrated a positive relationship between children's active verbal participation and their cognitive engagement. Preliminary evidence suggests that active verbal participation positively influences the effectiveness of teachers' scaffolding strategies for individual development: For example, in a study conducted by (Bürgermeister et al., 2019), an interaction effect was observed between the use of scaffolding and verbal participation on science learning. Similarly, in the primary school context, Herrmann et al. (2021)found an interaction effect between scaffolding and active participation on learning about the concept of evaporation, specifically in groups with a high proportion of multilingual children.

The positive effect of verbal participation on children's individual development is further supported by studies in the school context (Decristan et al., 2020; Hardy et al., 2019;Sedova et al., 2019) and further highlights the importance of fostering children's active verbal participation in classroom discourse as a fundamental task of teachers to promote development (Muhonen et al., 2016). However, there are also findings pointing to little impact of verbal participation. For instance, a study by Pauli and Lipowsky (2007) indicates that the participation of students in mathematical classroom discussions does not significantly influence posttest outcomes in mathematics. The evidence from the preschool context remain inconclusive, and analyses of children's verbal participation behavior in interactions are still limited (Bürgermeister et al., 2019; Tompkins et 2013). Descriptive analyses in this context, however, suggest that children's verbal participation in preschool interactions is relatively low. For example, preschool teachers have been observed to dominate speaking time during picture book readings (Simon & Sachse, 2011; Wirts & Cordes, 2021), science activities (Bürgermeister et al., 2019), and playtime (Tompkins et al., 2013; Wirts & Cordes, 2021).

Overall, the empirical findings on the developmental effects of verbal participation present a mixed picture, a pattern that may, at least partly, be attributed to the fact that most studies examine verbal engagement and other influencing factors, such as instructional support, in isolation. Yet, from a co-constructivist perspective and within the framework of the multilevel supply-use model, considering the interaction between these two dimensions may offer a more nuanced understanding of how children perceive, process, and benefit from learning opportunities. Empirical studies that explicitly address this interaction remain scarce and represent a critical gap in the current literature.

Moreover, inconsistencies in findings regarding the effect of verbal participation may also result from methodological differences, particularly in how verbal engagement is operationalized. Some studies measure the absolute number of speaker turns (e.g., Bürgermeister et al., 2019; Decristan et al., 2020) or the number of turns relative to the total number of turns in the interaction (e.g., Pauli & Lipowsky, 2007) or relative to the teacher (e.g., Bürgermeister et al., 2019). Other studies assess the duration of verbal contributions (e.g., Sedova et al., 2019; Tompkins et al., 2013; Wirts & Cordes, 2021), the number of words spoken (e.g., Herrmann et al., 2021; Wirts & Cordes, 2021), or the average number of words per speaker turn (e.g., Pauli & Lipowsky, 2007). Thus, two main approaches to assess verbal participation in classroom discourse can be identified: One focuses on the absolute number of speaker turns, without considering the length or duration of the turn, while the other emphasizes the length of verbal contributions, as indicated by either the number of words or the duration. The examination of individual children's verbal participation patterns – including both the number of speaker turns and the length of the turns – combined with an analysis of the instructional support provided by the preschool teacher, has the potential to demonstrate how interactions can facilitate engagement and lead to enhanced learning outcomes in early science activities.

Rationale and aims of the present study

Although a substantial body of research has examined the quality of interactions in ECEC and their influence on developmental outcomes in the context of literacy and numeracy, there is a notable lack of research focusing on instructional support in science settings and its impact on science knowledge development. In particular, there is limited research investigating the role of children's verbal participation behaviors, instructional support, and the interaction effects between individual participation and instructional support in these settings. This study aims to address this research gap by examining the impact of these factors on the development of children's language and science knowledge. Furthermore, this study analyzes the influence of prior knowledge.

As previously discussed, the CLASS Pre-K dimension of Concept Development includes relevant strategies that support science-specific learning, such as eliciting explanations, encouraging comparisons and classifications, activating prior knowledge, and linking concepts to everyday life. Concurrently, Language Modeling is closely linked to scaffolding language development. This study, therefore, focuses on the effects of Concept Development on the development of scientific knowledge and the effects of Language Modeling on language development, while controlling for prior language skills and/or science knowledge. By concentrating on the domain-specific application of these two dimensions, our study addresses critiques of CLASS ratings, which suggest that the three CLASS dimensions are overly broad and that high ratings for global interaction quality do not necessarily reflect high domain-specific quality (Levickis et al., 2023; Ulferts et al., 2019).

To this end, the following research questions are addressed in this study:

1. Do preschool teachers' Language Modeling and Concept Development predict the development of children's language skills and science knowledge, respectively?

Based on co-constructivist learning theory and the concept of scaffolding, higher levels of Language Modeling and Concept Development are expected to support children's language and science knowledge development.

2. Does children's individual verbal participation predict the development of their language skills and science knowledge?

It is expected that children's active verbal participation positively influences their development, with verbally engaged children in science activities demonstrating higher learning gains in language and science knowledge. This hypothesis is grounded in co-constructivist learning theory (Vygotsky, 1978), which highlights the importance of active verbal participation in teacher-child interactions for development.

The first two research questions serve as a foundation for the third, which takes a more integrative perspective by examining the interaction effect:

3. Does the interaction between the Language Modeling/Concept Development and children's individual verbal participation influence the development of children's language skills/science knowledge, respectively?

We hypothesize that the impact of teachers' instructional support will be particularly significant when children actively engage in verbal interactions. This hypothesis is based on two theoretical arguments. First, according to the multilevel supply-use model (Brühwiler & Blatchford, 2011), there is a dynamic interplay between teachers' instructional support and children's active engagement, as engaged children make greater use of the opportunities provided by teachers, thereby fostering their cognitive development. Second, the anticipated interaction effect can be attributed to teachers' ability to tailor their scaffolding more effectively to individual children's needs when they can analyze the verbal contributions children make during discourse (Bürgermeister et al., 2019). Without such contributions, this diagnostic adjustment would not be possible.

Methods

Sample

The data used for the analyses in this study were obtained from project ProfinK (*Professional learning communities for preschool teachers to support their ability to use diagnostic information for instructional decisions*), which involved preschool teachers from four medium to large German cities. Information about the study was communicated to preschool administrators and then to parents by participating teachers. Participation was voluntary, and informed consent was obtained from both preschool teachers and parents (on behalf of their children). The study was approved by the ethics committee of the RPTU University of Kaiserslautern-Landau (2022–001).

Study design

The project's objective was to provide professional development for preschool teachers with a focus on diagnostic and adaptive support in early science settings. The professional development focused on three central topics: science, as well as language and self-regulation, as key educational aspects of individual development. In the domain of science, preschool teachers were provided with ideas and background information on topics of materials, magnetism, floating and sinking, states of matter, and solving. In the domains of language and self-regulation, the professional development introduced specific scaffolding strategies designed to help preschool teachers address these developmental domains while implementing science activities. To support this implementation, a digital portfolio tool was available (Steffensky et al., 2025).

In addition to the initial 1.5-day professional development session, preschool teachers had the opportunity to participate in two-hour Professional Learning Community (PLC) sessions based on their interests and availability. These eight sessions were conducted over a period of approximately seven months. Within the PLCs, the implementation of science activities using the digital portfolio with various thematic focuses, such as diagnostic observation or language scaffolding, was discussed (Koschick et al., 2024).

Due to illness and other time constraints, not all preschool teachers who registered for the PLCs were able to attend regularly. As a result, the following classification was applied: To be included in the Intervention Group, a minimum attendance of 50%, corresponding to four to eight sessions, was required. This categorization was based on the assumption that meaningful engagement with the PLC format as a collaborative and participatory learning community requires attendance in at least 50% of the planned sessions (Steffensky et al., 2025). This categorization is supported by participants' perceptions of key features of PLCs, such as participation, shared values and discursive engagement (Darling-Hammond, 2017; Louis & Marks, 1998). Those who participated in at least four PLCs reported increasingly positive perceptions of these aspects over time (Steffensky et al., 2025). Preschool teachers who attended two to three sessions were assigned to the Professional Development Group, while those with lower attendance were designated as the Control Group (Koschick et al., 2024). Group membership was considered in the statistical analyses.

Table 1 presents the study design. A repeated-measures design with three measurement time points (MTP) was implemented between May 2022 and March 2024. The first MTP occurred prior to the

Table 1. Study design (based on Steffensky et al., 2025).

<table>
<tr><th colspan="2" rowspan="2">Participants</th><th colspan="2">1st MTP</th><th colspan="2">Intervention</th><th colspan="2">2nd MTP</th><th></th><th>3rd MTP</th></tr>
<tr><th colspan="2">May-Sep 2022</th><th>Sep 2022</th><th>Oct 2022 – Apr 2023</th><th colspan="2">May – Oct 2023</th><th></th><th>Nov 2023 - Mar 2024</th></tr>
<tr><td rowspan="3">Preschool teachers</td><td>Intervention Group (IG)</td><td rowspan="3">Questionnaire (biographical information, knowledge, attitudes/ beliefs)</td><td rowspan="4">Videography of a science activity</td><td rowspan="3">Kick-off event pro-fessional develop-ment, 1.5 days)</td><td>Participation in 4-8 PLCs (2 hours each)</td><td rowspan="3">Questionnaire</td><td rowspan="4">Videography of a science activity</td><td rowspan="5">Application phase (approx. 4 months)</td><td rowspan="3">Questionnaire</td></tr>
<tr><td>Professional Development Group (PDG)</td><td>Participation in 2-3 PLCs (2 hours each)</td></tr>
<tr><td>Control Group (CG)</td><td>Participation in 0-1 PLCs (2 hours each)</td></tr>
<tr><td colspan="2">Children</td><td>Standardized assessments in form of interviews (prior scientific knowledge, interest in sciences, non-verbal intelligence, self-regulation, language skills)</td><td colspan="2" rowspan="2"></td><td>Standardized assessments (scientific knowledge, interest in sciences, self-regulation, language skills)</td><td></td></tr>
<tr><td colspan="2">Parents</td><td colspan="2">Questionnaire (e.g. children's biographical information)</td><td>Questionnaire</td><td></td><td>Question-naire</td></tr>
</table>

Note. MTP: measurement time point; PLC: Professional Learning Community; the data and variables highlighted in gray are included in this analysis.

start of the professional development sessions for the preschool teachers. Children's development was assessed using a pre-post design.

Sample description

The present analyses focus on the subsample of children who were video-recorded in interaction with their preschool teacher during a science activity at MTP1. In total, 66 videos were recorded for research purposes, each documenting a science activity on the topic of "materials and their properties" ($M = 22.42$ min; $SD = 5.91$; min = 12; max = 45). No additional assistance or materials were provided to the teachers and the video recordings took place before the first intervention in the project.

The 66 recorded preschool teachers ($M = 40.43$ years, $SD = 10.87$; female: 74.2%) worked with small groups of 2 to 7 children ($M = 4.28$; $SD = 0.93$). Their professional backgrounds varied: 43.9% of the participants had completed training as preschool teachers, 18.2% had a degree in social/educational studies, 7.6% had other forms of educational or social training, 28.8% had completed other vocational training or for had missing information, and 1.5% had no vocational training. They had on average 12.13 years of work experience ($SD = 10.03$; min = 0, max = 35).

The children who were video-recorded during science activities ($N = 283$) had an average age of 4.45 years ($SD = 0.59$; min = 3, max = 6) at the first MTP. Of these children, 45.2% were female. Regarding their best spoken language, valid data were available for 159 children; among them, 88.1% primarily spoke German and 11.9% another language.

Analytic sample

The present longitudinal analyses include only those children who completed the assessment at both measurement points. Specifically, 183 children from 51 video-recorded groups were included in the models examining language development, and 186 children from the same 51 groups were included in the multilevel regression models on science knowledge development. Due to listwise exclusion of children with missing outcome data, these models should be interpreted as exploratory. To assess the robustness of the findings and to evaluate potential bias due to missing data, all analyses were re-run using multiply imputed data sets (see section *Multiple Imputation*).

Measures

In this study, the instructional support and the verbal participation of the children and preschool teachers were evaluated in video recordings at the first MTP. Biographical data on the children (gender, age) were collected, as well as their language skills and science knowledge at two MTPs. On average, the time interval between the two MTPs was $M_{language} = 323.88$ days ($SD = 55.25$; min = 181; max = 455) and $M_{science} = 310.58$ days ($SD = 50.47$; min = 146; max = 420).

Classroom Assessment Scoring System – CLASS Pre-K

To measure interaction quality, the CLASS Pre-K was used (Pianta et al., 2008). Due to the study design and available data, it was not possible to analyze four cycles of 20 min per preschool teacher, as recommended by CLASS. The video recordings were conducted by trained project assistants following a standardized protocol to ensure consistency across all settings. The recordings were divided into two cycles if their duration exceeded 31 min. The recorded videos varied in length ($M = 22.42$ min; $SD = 5.91$; min = 12; max = 45). Consequently, there were teachers for whom only one cycle could be analyzed and a few teachers whose ratings in the Instructional Support domain included three cycles. Following standard CLASS procedures, we calculated each preschool teacher's rating as the average of their completed observation cycles (Pianta et al., 2008). For all three CLASS domains, ratings are assigned on a scale of 1 to 7, with 1 and 2 representing low quality, 3, 4, and 5 indicating mid-range quality, and 6 and 7 high quality. All raters of the video data held a valid Teachstone® certificate. A total of 38% of the videos were double-rated, with the remainder rated by the first author. The reliability of ratings across all CLASS domains was excellent (Intraclass Correlation Coefficient [ICC] = .89, two-way random, single measure, absolute agreement). Interrater agreement was considered fair for the domain of Instructional Support (ICC = .56), fair for the dimension of Concept Development (ICC = .43), and good for Language Modeling (ICC = .64) (Cicchetti, 1994). In addition, weighted Cohen's kappa coefficients (squared weights) confirmed fair agreement for Concept Development ($\kappa = .42$, $p = .012$) and good agreement for Language Modeling ($\kappa = .63$, $p < .001$) (Cicchetti, 1994). Taken together, ICC and weighted kappa values indicate interrater reliability ranging from fair to good across the rated dimensions, with Concept Development yielding the lowest consistency. Final ratings were derived through critical discussions of discrepancies to create a master rating with the aim of ensuring analytical consistency.

To provide a more comprehensive view of rating reliability and to align with established reporting practices, some studies additionally report percentage agreement alongside ICC values, as specified in the CLASS manual (e.g., Suchodoletz et al., 2014). As outlined in the CLASS Pre-K manual, interrater agreement can be expressed as a percentage, with an acceptance deviation of ± 1 in the ratings. This margin is considered "to reflect an acceptable degree of accuracy in ratings" (Pianta et al., 2008, p. 96). In accordance with this procedure, there is a 95% agreement in the domain of Instructional Support, a 100% agreement in the dimension of Conceptual Development, and a 97% agreement in the dimension of Language Modeling.

Assessment of verbal participation of teachers and children

The verbal participation of the preschool teachers and children was quantified by speaker turns per minute (to allow comparisons between videos of varying lengths) and average number of words per speaker turn. This operationalization follows established approaches in the literature (e.g., Pauli & Lipowsky, 2007; Sedova et al., 2019) and captures two key aspects of verbal engagement: the frequency of contributions (speaker turns per minute) and their length (average words per turn), as outlined in the theoretical section on children's verbal participation. Transcripts based on standardized guidelines were analyzed using MAXQDA 2022 software (Release 22.8.0; VERBI Software, 2021). The software provided the total number of words spoken by each participant, as well as the number of speaker turns. Filler words were excluded, and parallel speech or interjections that did not interrupt a speech were not counted. All interactions from the science-related videos were included in this analysis.

Pre- and post-Assessment of children's language skills

Participating children's language skills were assessed using the *Wechsler Preschool and Primary Scale of Intelligence-IV* (WPPSI-IV; Wechsler, 2018). This instrument measures general and specific cognitive skills in children aged 2.6 to 7.7 years. The Vocabulary Acquisition Index subscale was utilized, which measures a child's active vocabulary (picture naming; 35 items) and receptive vocabulary (31 items). These language skills were assessed by presenting appropriate stimuli. Both subtests followed a cutoff criterion; they were terminated after the child answered three consecutive items incorrectly or not at all. For analysis, sum scores were used.

Pre- and post-Assessment of children's science knowledge

A pre-posttest procedure was employed to assess children's science knowledge. Items from previous studies were used to evaluate knowledge in the topic of "materials" (based on Steffensky et al., 2018). The scale assessed whether children could differentiate, categorize, and name relevant materials. For example, three items required children to match objects by material composition without any verbal response: "Look closely at this building block (spoon/thermos bottle). Which of these things is made of the same material as the building block (spoon/thermos bottle)?" In this task, children are required to sort the objects physically or by pointing, without speaking. In contrast, other items required verbal responses, such as: "Look, I have a hammer here. Can you tell me what this hammer is made of?" If the child correctly identifies one material, they receive one point. If they correctly name both materials, they get two points.

The scale comprises 12 items, with reliability values of Cronbach's alpha = 0.80 (pretest) and 0.76 (posttest), indicating fair to good reliability (Cicchetti, 1994). The science scale demonstrated moderate correlations with age ($r = .272$) and cognitive abilities ($r = .319$). A stronger correlation with language skills ($r = .674$) suggested that linguistic competencies influenced responses, likely due to the linguistic nature of the items. However, the correlations were not excessively high, supporting the construct validity of the scale.

Science knowledge was assessed via individual interviews (~30 min) conducted by trained research staff. The focus on "materials" in the scale reflects the science content emphasis of the ProfinK project. Knowledge about different materials serves as a foundation for understanding other science topics (e.g., recognizing that only certain metals are magnetic). Since all teachers initially explored this topic with the children, the scale serves as a reference for assessing science knowledge development, regardless of which specific phenomena were implemented. Additionally, it is expected that exploring other science topics will also promote children's learning about materials and their properties.

Assessment of biographic data

Children's ages were determined using parental consent forms for participation in the study and included as a control variable in the multilevel analyses.

Data analysis

Descriptive analyses were conducted using IBM SPSS version 29.0.1.0 (IBM Corp, 2023). To address the three research questions and given the hierarchical data structure (children nested in groups), multilevel regression analyses were conducted separately for the two domains language and science. This separation reflects both the distinct conceptual focus of the predictors, Language Modeling for language development and Concept Development for science knowledge, as well as the domain-specific orientation of the research questions (RQ1). Each model included individual-level predictors (prior knowledge, age, verbal participation) and group-level predictors (Language Modeling/Concept Development, participation in a professional development group within the project), as well as cross-level interactions (RQ3), to examine moderation effects. As verbal participation (RQ2) was operationalized using two distinct indicators, the average length (word count) of a child's speaker turn and children's speaker turns per minute, two separate models were estimated for each outcome domain. Multilevel regression analyses were conducted in R version 4.5.0 (R Core Team, 2025) using the *lme4* package (Bates et al., 2015).

All predictors were grand-mean centered (Enders & Tofighi, 2007). Teacher participation in professional development was considered using weighted effect coding to adjust for unequal group sizes (Control Group (CG): $N_{children} = 158$; Professional Development Group (PDG): $N_{children} = 31$; Intervention Group (IG): $N_{children} = 94$). The recoded values represent deviations from the weighted mean.

Exploratory analyses with complete-case samples

For the initial exploratory analyses, only children with complete data for both measurement points (MTP1 and MTP2) on the respective outcome variable were included. This resulted in a sample of N = 183 children (51 groups) for the language models and N = 186 children (51 groups) for the science models. While listwise exclusion entails a risk of bias in smaller samples, it was considered appropriate for the initial exploratory analyses in order to ensure transparency and interpretability of model estimation.

Correlation analyses identified significant predictors, followed by intercept-only models to calculate Intraclass Correlation Coefficient (ICC) for group effects. As the ICC exceeded .05 ($ICC_{language} = .238$; $ICC_{science} = .172$), multilevel modeling was considered appropriate (Thomas & Heck, 2001).

Table 2 provides an overview of the multilevel regression analysis plan. Model comparisons were performed using ANOVA, deviance, log-likelihood tests, and/or AIC and BIC. Predictors were added stepwise, beginning with level 1 variables, followed by level 2 variables, and finally cross-level interactions to test moderation effects (RQ3). Each variable was sequentially included as a fixed effect, random effect, and interaction term, with model improvements assessed at each step. Standardized regression coefficients were reported for all final models. An ANCOVA approach was used to analyze both metric and categorical variables where applicable, including in the final models.

Multiple imputation

To address potential bias due to missing data and to make full use of the available sample of 283 video-recorded children, we conducted multiple imputation using the *mice* package (v3.18.0; van Buuren & Groothuis-Oudshoorn, 2011). Since only scale-level child variables were imputed, we used the full dataset of all children participating in the ProfinK project (N = 383), not just the video-recorded subsample. Across this full sample, the proportion of missing values on the relevant outcome variables (language and science knowledge at MTP1 and MTP2) ranged from 4.5% to 34.6%. To assess the missing data mechanism, Little's MCAR test was conducted using the *MissMech* package (Jamshidian et al., 2014). Although multivariate normality assumptions were violated (Hawkins test, $p = .015$), the nonparametric component of the test yielded a non-significant result ($p = .053$), suggesting that the data were plausibly missing completely at random (MCAR; Little, 1988).

Table 2. Multilevel regression – data analysis plan (models 1–4).

Model	Included Variables
Intercept-only model	AV with random intercept
Random-intercept model/random-slopes model	Children's prior knowledge: Language Skills 1st MTP/Science knowledge 1st MTP Age Verbal participation: Average length (word count) of a children's speaker turn/ Children's speaker turns per minute
Intercept as outcome model	Teacher's professional development group Instructional Support Dimension: Language Modeling/Concept Development
Intercept and slopes as outcome model	Children's prior knowledge × Instructional Support Dimension: Language Skills 1st MTP × Language Modeling/ Science knowledge 1st MTP × Concept Development Verbal Participation × Instructional Support Dimension: Average length (word count) of a children's speaker turn × Language Modeling/ Concept Development/ Children's speaker turns per minute × Language Modeling/Concept Development

We used a two-level predictive mean matching method (*2 l.pmm*) and specified the grouping variable to reflect the nested data structure. Accounting for the hierarchical structure of the data is essential for producing valid estimates in subsequent multilevel analyses (Lüdtke et al., 2007). Two groups with only one or two children were excluded, as such small clusters do not allow for proper modeling of the hierarchical structure in multilevel imputation (van Buuren, 2018). The resulting imputation sample comprised $N = 378$ children. The predictor matrix was generated using the *make.predictorMatrix()* function and was refined to exclude identifier and item-level variables, retain only theoretically and empirically relevant predictors ($r > .10$), and minimize multicollinearity, particularly among related constructs in the science domain. A total of 50 imputations ($m = 50$) with 200 iterations each were generated (seed = 3103). Although some warnings regarding boundary (singular) fits were issued, which were likely caused by low variance within certain clusters, diagnostic checks (e.g., distributional comparisons with the raw data, correlations, and means) revealed no indications of implausible imputations. As an additional diagnostic, density plots of the imputed distributions for the four outcome variables (language and science knowledge at MTP1 and MTP2) are included in the supplemental material.

The final multilevel models were re-estimated using the multiply imputed datasets, with pooled estimates obtained using Rubin's rules (*pool()* function in *mice*). This dual approach, starting with exploratory complete-case analyses and followed by validation through multiply imputed models, ensures both transparency and robustness and enhances the interpretability of the findings. It also supports the use of exploratory models under listwise exclusion, given the partial uncertainty captured in key control variables. To evaluate the influence of missing data on the estimates, we computed the fraction of missing information (FMI) and relative increase in variance (RVI) for each pooled model. The FMI values ranged from 0.17 to 0.86 across the different parameters, indicating varying degrees of uncertainty due to imputation. Core predictors such as children's prior language skills (FMI = 0.86) and age (FMI = 0.75) exhibited higher FMIs, suggesting substantial missing data or between-imputation variance. Conversely, interaction terms (e.g. verbal participation × instructional support) and group contrasts (PDG/IG) exhibited lower FMI values (e.g. 0.20–0.48), suggesting greater stability. The corresponding RVI values followed this pattern, ranging from 0.21 to 5.96. A detailed overview is provided in Supplemental Tables 3 and 4. Following White et al. (2011), we calculated Monte Carlo Errors (MCE) to assess the stability of the pooled standard errors across imputations. An MCE value below 0.01 is typically considered acceptable. In the science knowledge models (Supplemental Table 4), all MCE values were ≤0.01, confirming high estimate precision. In the language models (Supplemental Table 3), MCE values were generally low. Slightly higher values occurred for control variables not central to our research questions (e.g., group contrasts), which also showed high standard errors. Importantly, all key predictors and interaction terms relevant to our hypotheses, such as children's verbal participation and its interaction with instructional support, showed MCE values < 0.1, indicating sufficient precision for interpretation.

Results

Descriptive statistics

CLASS PreK – Instructional Support

As expected, the ratings in the Instructional Support domain indicate low quality in the sample ($M = 2.90$, $SD = 0.72$, min = 1.33, max = 4.67). The variance is limited, and the upper range of the 7-point CLASS Pre-K scale is not fully utilized; no high-quality videos in terms of Instructional Support were observed in the sample. Considering the ratings of the particular dimensions, it can be seen that both the ratings of Concept Development ($M = 2.51$, $SD = .78$, min = 1.0, max = 5.0) and of Quality of Feedback ($M = 2.90$, $SD = .88$, min = 1.0, max = 5.0) are in the low quality range, while the quality rating of Language Modeling is in the lower medium range ($M = 3.28$, $SD = .91$, min = 1.0, max = 6.0).

Verbal participation

This section provides an overview of the central tendencies and distribution of verbal participation in the videos. For completeness, the number of words spoken per minute is also presented (Figure 1), in addition to the number of speaker turns per minute and the average length of speaker turns.

Regarding the absolute number of speaker turns, preschool teachers made an average of 5.67 turns per minute ($SD = 2.12$, min = 1.67, max = 11.40), the mean length of a turn was 17.92 words (SD = 9.63, min = 5.41, max = 63.16). The mean number of turns per minute for the children at the group level was 7.66 (SD = 3.45, min = 2.82, max = 17.33), with an average word count per turn of 5.09 ($SD = 2.53$, min = 1.91, max = 14.12). At the individual level, the mean number of turns per minute per child was 1.74 ($SD = 1.28$, min = 0.00, max = 6.76). The mean length of a turn varied among children, with a mean of 4.81 words ($SD = 3.12$, min = 0.00, max = 23.25).

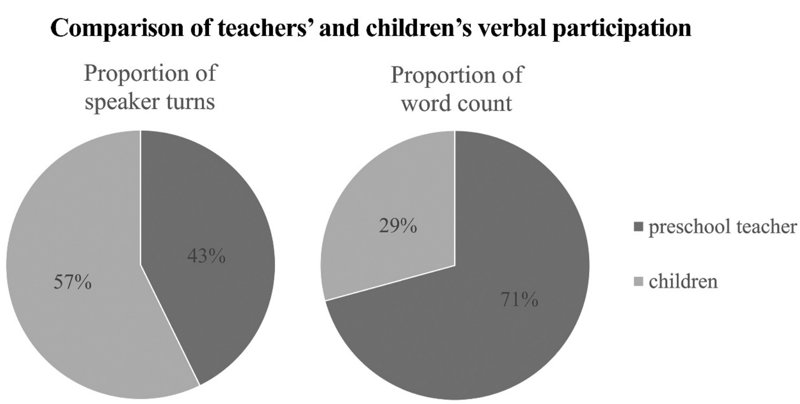

Figure 1. Comparison of teachers' and children's verbal participation. Proportion of speaker turns and word count in preschool science activities. Children take more turns (57%), while teachers dominate in terms of word count (71%). *Note*. Word counts per minute: overall ($M = 123.32$, $SD = 26.63$, min = 73.45, max = 242.44), preschool teacher ($M = 87.85$, $SD = 22.95$, min = 36.33, max = 139.66), children at group level ($M = 35.36$, $SD = 16.70$, min = 12.75, max = 102.78), individual child ($M = 8.05$, $SD =$ 7.08, min = 0.00, max = 50.63).

Table 3. Descriptive statistics of outcomes at MTP1 and MTP2 for video-recorded and complete-case samples.

		Sample of video-recorded children (*N* = 283)			Exploratory subsamples	
Outcome	MTP	*N*	Missing	*M* (*SD*); min; max	*N*	*M* (*SD*); min; max
Language skills	1	245	13,4%	31.93 (11.23); 4; 53	183	34.05 (10.43); 4; 53
	2	213	24,7%	38.48 (10.16); 2; 60	183	39.83 (9.26) 8; 60
Science knowledge	1	274	3,2%	.81 (0.45); 0; 1.92	186	0.87 (0.43) 0; 1.92
	2	212	25.1%	1.18 (0.40); 0.17; 2.00	186	1.21 (0.38); 0.17; 2.00

Note. MTP: measurement time point; language skills: the sum score corresponds to the number of correct items on the picture naming and receptive vocabulary subscales (WPPSI; Wechsler, 2018), with a possible maximum sum score of 66; science knowledge: the mean value ranges from 0 to 2.

Children`s language skills and science knowledge

Table 3 presents the descriptive statistics for children's language skills and science knowledge at the two MTPs.

Effects of instructional support, children's verbal participation, and their interaction on children's language and science knowledge development

The results of the exploratory multilevel regression analyses for language (models 1 and 2) and science knowledge (models 3 and 4) are shown in Table 4. Models 1 and 3 account for the average length (word count) of a child's speaker turn, while models 2 and 4 include the number of children's speaker turns per minute. This distinction reflects two established approaches to capturing verbal participation in classroom discourse: one emphasizing the frequency of turns, the other focusing on the length or elaboration of verbal contributions (see section *Children's verbal participation and its impact on their learning and development*). By incorporating both indicators separately, the analyses allow for a differentiated understanding of how distinct participation patterns in early science activities relate to children's development. Each model additionally includes individual-level predictors, such as prior knowledge and age, as well as group-level predictors including instructional support (Language Modeling or Concept Development) and teacher's participation in a professional learning community. Cross-level interactions between instructional support and the respective participation indicator were added to examine potential moderation effects.

For reasons of clarity, Table 4 reports only the summary of results; complete model outputs from both the exploratory and multiply imputed analyses are provided as supplemental materials. Significant effects that were replicated in the multilevel models based on multiply imputed data are highlighted in gray.

Development of children's language skills

The intercept values of the language development models indicate that a sum score of 39.87 (model 1)/39.88 (model 2) can be expected at the second MTP for children with average language skills at the first MTP. The analysis demonstrates that the children's initial language skills are significant predictors of language development. Across all groups, an increase of 0.65 (model 1; $p < .001$)/0.69 (model 2, $p < .001$) standard deviations can be expected from the first to the second MTP. The increase was consistent across all groups, as reflected in the R^2 values. Age of the children does not significantly influence language skills in our sample. To evaluate potential multicollinearity, correlations between the predictors (prior language skills and age) were examined. While prior knowledge and age exhibited a significant correlation ($r = .233$; $p = .002$), this moderate correlation did not indicate multicollinearity issues.

Contrary to hypothesis 1, Language Modeling showed no significant effect on children's language development, neither as a predictor nor in interaction with prior language skills. Similarly, the participation of preschool teachers in the Professional Development Group or Intervention Group was not found to be relevant. Regarding children's verbal participation, analyses showed that the mean length of children's speaker turns had a significant positive effect on language development (model 1, *Estimate* = .31, $p = .019$), consistent with hypothesis 2. This effect was replicated in the analysis with multiply imputed data (*Estimate* = .37, $p = .036$). In contrast, model 2 indicated a significant negative effect of the number of speaker turns

Table 4. Summary of multilevel regression models (listwise deletion) for the development of children's science language skills and knowledge.

Effect	Development of Children's Language Skill				Development of Children's Science Knowledge			
	Model 1		Model 2		Model 3		Model 4	
	Estimate [90%-CI]	*p*	Estimate [90%-CI]	*p*	Estimate [90%-CI]	*p*	Estimate [90%-CI]	*p*
Fixed Effects (Random Effects)								
Intercept	39.90 [39.18, 40.62]	**< .001**	39.92 [39.19, 40.64]	**< .001**	1.21 (0.10) [1.17 (0.03), 1.25 (0.12)]	**< .001**	1.21 (0.11) [1.17 (0.03), 1.25 (0.12)]	**< .001**
Individual-level variables								
Language Skills (1st MTP)	0.65 [0.58, 0.72]	**< .001**	0.68 [0.61, 0.79]	**< .001**	0.01 [< 0.01, 0.01]	**.004**	0.01 [< 0.01, 0.01]	**.004**
Science knowledge (1st MTP)	–	–	–	–	0.52 (0.20) [0.41 (0.07), 0.63 (0.25)]	**< .001**	0.52 (0.19) [0.42 (0.04), 0.62 (0.25)]	**< .001**
Age	0.04 [−0.08, 0.15]	.631	0.06 [−0.06, 0.17]	.422	< 0.01 [> −0.01, 0.01]	.531	< 0.01 [−0.01, 0.01]	.992
Average length (word count) of a children's speaker turn	0.30 [0.08, 0.51]	**.024**	–	–	> −0.01 [−0.01, 0.01]	.460	–	–
Children's speaker turns per minute	–	–	−0.83 [−1.45, −0.21]	**.027**	–	–	−0.01 [−0.04, 0.02]	.523
Group-level variables								
Teacher's professional development group								
PDG	−0.74 [−3.19, 1.71]	.618	−0.92 [−3.33, 1.48]	.529	0.01 [−0.10, 0.13]	.836	0.05 [−0.07, 0.16]	.526
IG	−0.04 [−0.91, 0.99]	.942	0.15 [−0.79, 1.09]	.790	0.02 [−0.03, 0.06]	.595	< 0.01 [−0.04, 0.05]	.903
Language Modeling	0.43 [−0.47, 1.33]	.430	0.50 [−0.38, 1.38]	.347	–	–	–	–
Concept Development	–	–	–	–	0.06 [< 0.01, 0.12]	.089	0.05 [> −0.01, 0.11]	.126
Interaction								
Language Skills (1st MTP) × Language Modeling	−0.03 [−0.12, 0.06]	.582	−0.05 [−0.14, 0.04]	.359	–	–	–	–
Average length (word count) of a children's speaker turn × Language Modeling	−0.11 [−0.48, 0.26]	.618	–	–	–	–	–	–
Children's speaker turns per minute × Language Modeling	–	–	0.21 [−0.47, 0.89]	.610	–	–	–	–
Science knowledge (1st MTP) × Concept Development	–	–	–	–	−0.06 [−0.18, 0.06]	.414	−0.03 [−0.15, 0.08]	.627
Average length (word count) of a children's speaker turn × Concept Development	–	–	–	–	0.02 [< 0.01, 0.03]	**.047**	–	–
Children's speaker turns per minute × Concept Development	–	–	–	–	–	–	0.05 [< 0.01, 0.09]	.059

Note. Model 1 and 2: $N_{children} = 183$, $N_{groups} = 51$; model 3 and 4: $N_{children} = 186$, $N_{groups} = 51$; *Estimate*: estimated effect size; *CI*: confidence interval (LL: lower limit, UL: upper limit); full model results including standard errors, AIC, BIC, marginal and conditional R^2 are provided in supplemental material; MTP: measurement time point; PDG: Deviation of the Professional Development Group (PDG) from the weighted mean value; IG: Deviation of the Intervention Group (IG) from the weighted mean value; language skills: the sum score corresponds to the number of correct items on the picture naming and receptive vocabulary subscales (WPPSI; Wechsler, 2018), with a possible maximum sum score of 66; science knowledge: the mean value ranges from 0 to 2. Significant effects highlighted in gray were replicated in the multilevel regression analyses using multiply imputed data (see Supplemental Tables 3 and 4).

per minute on language development ($p = .023$), indicating that a higher frequency of speaker turns per minute was associated with a decrease in language skills of −0.85 standard deviations. However, this effect could not be replicated in the imputed multilevel models (*Estimate* = −0.36, $p = .400$). The interaction between children's participation (level 1) and Language Modeling (level 2) showed no significant effects on children's language development, which contradicts hypothesis 3. In terms of model fit, model 2 exhibited a marginally superior fit according to the AIC and BIC values.

Development of children's science knowledge

The multilevel regression models indicate that a mean score of 1.21 can be expected at the second MTP for children with average science knowledge at the first MTP. The deviation of the group-specific learning development from the mean learning development across all groups is 0.1 standard deviations. The models further demonstrate that children's prior knowledge at the first MTP has a significant effect on the outcomes at the second MTP. Across all groups, a difference of 0.52 (model 3 and 4, $p < .001$) standard deviations can be predicted from the first to the second MTP. This effect was replicated in the multilevel regression models based on the multiply imputed dataset (*Estimate* = 0.38, $p < .001$). Group-specific learning development deviates by 0.21 standard deviations from the mean learning development across groups. This shows that there are differences in science knowledge development between the groups. The $R^2_{conditional}$ value of .6448 indicates the presence of random effects associated with the intercept and prior science knowledge. This suggests that both fixed and random effects together explain a significant portion of the variance in the model, as seen in Table 4.

Age is not a significant predictor in model 3 and 4, although science knowledge and age show a moderate, statistically significant correlation ($r = .360$, $p < .001$). Language skills at the first MTP also explain additional variance in the models. For children with average science knowledge, an increase of .01 standard deviations in the mean score of science outcomes at the second MTP can be expected (model 3 and 4, $p = .004$). This finding was also confirmed in the models with multiply imputed data (*Estimate* = 0.01, $p = .012$). The participation of preschool teachers in the Professional Development Group or Intervention Group was not a significant predictor. Contrary to hypotheses 1 and 2, neither Concept Development nor the number or length of speaker turns significantly influenced the development of science knowledge. However, in line with hypothesis 3, the exploratory analysis (model 3) revealed a significant positive interaction effect between average speaker turn length and Concept Development on children's science knowledge (*Estimate* = .02, $p = .047$). However, this result could not be replicated in the imputed multilevel models (*Estimate* = 0.01, $p = .206$). For model fit, model 4 exhibited a marginally superior fit based on the AIC and BIC values.

Discussion

Language is fundamental to children's (domain-specific) learning and development, serving as both a communication and cognitive tool for constructing knowledge through interaction (Tomasello, 1999). In ECEC, high-quality interactions between preschool teachers and children are crucial for supporting the development of children with diverse prerequisites (Suchodoletz et al., 2023). Emphasizing the active role of children in their own learning processes also highlights the importance of children's verbal participation (e.g., Sedova et al., 2019). This study examines the role of instructional support provided by preschool teachers and children's verbal participation in early science education settings, analyzing their effects on the longitudinal development of language and science knowledge. The discussion is structured as follows: First, the descriptive results are presented and contextualized, focusing on the instructional support and children's verbal participation. This is followed by a summary and interpretation of multilevel analysis results, examining the effects of these factors on children's language and science knowledge development.

Descriptive statistics of instructional support and children's verbal participation in early science settings

In alignment with prior research, we hypothesized low to medium quality ratings in the domain of Instructional Support according to CLASS Pre-K and a verbal dominance of preschool teachers in their interactions with children. As expected, the average quality rating of Instructional Support was low, with no teachers achieving high ratings and limited variance. The absence of high-quality interactions in the CLASS domain of Instructional Support is consistent with findings from other studies in ECEC contexts in Germany (e.g., Kohl et al., 2019; Suchodoletz et al., 2014) and other countries (Hofkens et al., 2023). These results underscore the challenging task for preschool teachers to provide high-quality instructional support. It requires appropriate content, pedagogical, and pedagogical content knowledge to implement support for higher-order thinking skills and promote deeper domain-specific understanding, as well as the maintenance of continuous participation through feedback and Language Modeling (Kunter et al., 2013; Peters et al., 2020).

Regarding verbal participation during the science activities, our findings revealed that preschool teachers dominated the verbal interaction, as reflected by their significantly longer speaker turns in terms of word count. This result aligns with a study by Bürgermeister et al. (2019), who found that Swiss preschool teachers spoke 72.5% of the words during science activities. Similarly, a study by Wirts and Cordes (2021) found a comparable distribution of verbal contributions during picture book activities and free play situations in preschool settings. In contrast to the average length of turns, the number of turns was relatively balanced between children and teachers. However, the range among the children was substantial. Some children did not speak at all, while others had few but relatively long turns (e.g., 'I know a game, how about we introduce the game called "feel what material it is" and it goes like this: everyone now feels an object and then feels what material it is.'). Meanwhile, others had many but very short (one-word) turns (e.g., yes/no answers or labeling a material like "metal" or "wood"). The observation of children who did not participate verbally at all is concerning, as language learning is particularly effective when children are linguistically active and develop their language skills in interactions (Wirts & Cordes, 2021), as also shown in the present study.

Effects of instructional support, children's verbal participation, and their interaction on children's language and science knowledge development

Contrary to hypothesis 1, Language Modeling had no significant effect on children's language development, neither as a predictor nor as a moderator. However, our analyses revealed that children's verbal participation, particularly the length of speaker turns, had a significant positive effect on their individual language development. This result supports the findings of previous studies (e.g., Sedova et al., 2019) and was also replicated in the analyses using multiply imputed data, indicating robust evidence across analytical strategies. In contrast, the negative effect of the number of speaker turns per minute on language development, which was found in the exploratory models, could not be replicated in the models based on imputed data. Therefore, this result should be interpreted with caution and regarded as an exploratory finding that requires further investigation. Nevertheless, it may point to the possibility that a high frequency of short or fragmented child contributions, reflected in a higher number of speaker turns per minute, does not necessarily support language development. A high number of turns may indicate short answers or frequent interruptions in the conversation, which could reduce engagement in cooperative and problem-solving dialogue and limit the opportunity to produce extended responses or utilize complex language structures. One potential explanation may be found in the interaction patterns observed in preschool settings, for example with regard to preschool teachers' questioning. Research has shown that preschool settings are characterized by a predominance of closed-ended questions, leading to fragmented conversational patterns and a restriction of children's verbal engagement (e.g., Hormann & Skowronek, 2019). Conversely, open-ended questions have been shown to elicit more complex and extensive linguistic responses from children (e.g., Lee & Kinzie, 2012). However, questions that are linguistically and cognitively stimulating remain infrequent in everyday ECEC practice (e.g., Hamel et al., 2021; Lee & Kinzie, 2012).

The expected positive interaction effect between Language Modeling and children's participation was not observed, likely due to the sample's moderate quality ratings and the absence of high-quality interactions in this dimension. It can be assumed that maintaining a medium to high level of interaction quality over an extended period is necessary to positively impact the development of a heterogeneous learning group (Levickis et al., 2023).

In examining the impact of instructional quality on children's science knowledge development, we focused on Concept Development and explored whether children's participation influenced this domain-specific learning outcome. Contrary to hypothesis 1 and 2, neither Concept Development nor children's verbal participation, when considered independently, showed significant effects on science learning. The absence of notable effects of Concept Development on science knowledge may be attributed to the low average quality ratings observed in the sample (Levickis et al., 2023). Additionally, the high-inference nature of the CLASS instrument might lack sufficient specificity for domain-specific settings (Wang et al., 2022). Although the analysis focused explicitly on the dimension of Concept Development, which is closely related to domain-specific support and higher-order thinking, this focus might not have fully addressed this limitation. As Ulferts et al. (2019) noted, interaction quality, as measured by high-inference tools, tends to have stronger effects on language learning than on domain-specific outcomes such as math. Nevertheless, CLASS ratings provide valuable insights into how learning activities can be supported, which has implications for the professional development of preschool teachers (Ulferts et al., 2019).

Furthermore, the study revealed an absence of evidence supporting the hypothesis 2, namely that children's participation as predictors had a positive effect on their learning outcomes in the science knowledge domain. While the exploratory analyses identified a significant interaction effect between Concept Development and the average length of children's speaker turns (hypothesis 3), this effect was not replicated in the multiply imputed models. Given this discrepancy, the interaction effect should be interpreted with caution and regarded as an initial indication rather than a robust finding requiring validation in future studies. Nonetheless, this exploratory result may offer a theoretically plausible indication of how instructional support and children's active participation (associated with the utilization of learning opportunities) can interact to foster science knowledge. This interpretation is supported by previous research (Herrmann et al., 2021), including a study in preschool context by Bürgermeister et al. (2019), that examined the interplay between children's speaker turns at the group level and the use of scaffolding strategies by the preschool teacher to activate prior knowledge. Teachers' scaffolding strategies can be more effectively adapted to children's individual needs when children actively engage through verbal contributions. Such contributions provide diagnostic insights into children's domain-specific knowledge and language proficiency (Bürgermeister et al., 2019). Thus, active verbal participation by children can be seen as a prerequisite for the effectiveness of scaffolding strategies.

Although the interaction effect was not replicated across models, it can be considered within the framework of a co-constructivist process of learning (Vygotsky, 1978), in which learning is understood as a collaborative process shaped through social interaction. In this context, verbal participation enables preschool teachers to identify children's zone of proximal development and to provide contingent scaffolding. Furthermore, within the framework of the multilevel supply-use model (Brühwiler & Blatchford, 2011), verbal participation may serve as an indicator of the child's engagement with learning opportunities. However, the efficacy of co-constructive processes for domain-specific learning is likely influenced by additional factors, such as the composition and size of the learning group, group dynamics and peer relations, and the conceptual and linguistic quality of children's verbal contributions, as well as the extent to which these contributions relate to the domain-specific content.

Our study highlights the critical role of prior knowledge in predicting both language and science learning outcomes. These findings are consistent with prior research on science knowledge development (e.g., Saçkes et al., 2011), which demonstrates that children with higher levels of prior knowledge tend to achieve better learning outcomes in both the short and long term. A notable result of this study is the finding that children's prior language knowledge significantly predicts their science knowledge outcomes, aligning with a study by Henschel et al. (2023), who identified a reciprocal relationship between science vocabulary and science knowledge. This underscores the importance of considering prior knowledge as a central factor in development and providing children varied opportunities to engage with domain-specific content from an early age, thus laying the foundation for future learning (National Research Council, 2012).

Overall, our study provides valuable exploratory insights into the interaction effects between preschool teachers' instructional support and children's verbal participation, emphasizing the dynamic interplay between these variables and their potential influence on children's learning and development. Although no long-term effects of the CLASS dimensions were found, the exploratory finding of an interaction effect between teachers' Concept Development and the length of children's verbal contributions may point to the complex nature of interactions in ECEC. For language development, the length of verbal contributions appears to be a crucial factor, emphasizing the importance of fostering longer, meaningful verbal exchanges rather than brief contributions. This effect was replicated across both analytical approaches. When children are given chances to participate in extended discourse and make active use of the provided learning opportunities, their language development benefits significantly.

Although this assumption is only supported by exploratory findings and was not consistently confirmed in the present analyses, theoretical frameworks and earlier studies (Bürgermeister et al., 2019, p. 3) suggest that the efficacy of scaffolding strategies, as conceptualized through the CLASS dimensions Concept Development and Language Modeling, may be enhanced when combined with meaningful opportunities for children's verbal participation. Further empirical validation is needed to substantiate this interaction. Taken together, these considerations point to the importance of creating learning environments that provide not only high-quality instructional support but also opportunities for children's active and extended verbal participation. For instance, open-ended questions have been shown to stimulate language production by eliciting extended responses from children (Haen et al., 2025; Lee & Kinzie, 2012). The results of this study highlight the necessity of considering the interplay between interaction quality and children's active engagement with learning opportunities.

Limitations and further research

The findings of this study must be interpreted considering certain limitations that may have implications for future research. The relatively small sample size, particularly at the group level, and the high proportion of participating preschool teachers with an academic degree (nearly 20%), which exceeds the German ECEC sector average by more than double (Autorengruppe Bildungsberichterstattung, 2020), limit the generalizability of the results. Future studies should include larger, more diverse samples and consider additional covariates, such as teachers' professional background and children's socio-economic and language background. The latter could not be included in the present analyses due to a high amount of missing data in the parental questionnaires. Including these variables may help identify whether specific groups of children, such as dual-language learners who speak little German at home, especially benefit from instructional support in early science settings.

Importantly, the limited replication of some effects in the multiply imputed dataset may in part reflect methodological constraints of the imputation process. In particular, the small cluster sizes at level 2 (ranging from 3 to 8 children per group) may have reduced the precision of imputations, despite acceptable convergence diagnostics and distributional checks (van Buuren, 2018). This limitation underscores the need for cautious interpretation of effects that were only significant in the exploratory complete-case models and not replicated in the imputed analyses. Consequently, the findings should be regarded as preliminary, interpreted in conjunction with prior research and relevant theoretical frameworks, and examined through further empirical investigation.

Instructional Support ratings were based on a single video recording of a science activity, and in some cases on a single CLASS cycle, which may not fully reflect typical classroom interactions. Future research should assess longer and more varied (science) activities to provide a more comprehensive evaluation of instructional support. Although the CLASS tool is validated in the German context (Stuck et al., 2016; Suchodoletz et al., 2014), the timing of observed interactions within the daily routine can influence results (Praetorius et al., 2014; Thorpe et al., 2020). In addition, the quality of interactions between preschool teachers and individual children within a group may vary (Burchinal, 2017), which raises questions about the generalizability of CLASS ratings for drawing conclusions about interaction quality in a specific group. The quality of instructional support observed in the recorded science activity likely does not reflect teachers' behavior in everyday routines or larger group settings. This raises the question of the ecological validity of the results. Moreover, factors such as the proximity to and the extent of contact with the teacher, as well as the quantity of high-quality interactions, may also significantly influence children's development Anders and Oppermann (2024).

Furthermore, the relatively low interrater reliability for the Concept Development dimension indicates that ratings in this domain should be interpreted with caution. Although high percentage agreement was found within the accepted deviation range (Pianta et al., 2008), exact agreement, ICC and weighted kappa values indicated only fair consistency. Such measurement error may have attenuated associations with children's outcomes, limiting the strength of inferences for Concept Development. In order to address this concern, sensitivity analyses were conducted in which Concept Development was excluded from the science models or substituted by Language Modeling, and vice versa for the language models. These analyses indicated that the overall pattern of significant and nonsignificant effects remained stable across model specifications and data handling approaches (listwise deletion vs. multiple imputation). In the exploratory science models, however, the significant interaction between Concept Development and children's average turn length was no longer evident when Concept Development was replaced by Language Modeling.

The limited variance and absence of high CLASS ratings may explain why the effects of interaction quality on the development of language and science knowledge were not statistically significant. Restricted variability in a predictor can substantially reduce the statistical power of a test, as it constrains the range of possible correlations and thus lowers the likelihood of detecting a true effect. When a measure is confined to a narrow range, the test's sensitivity to subtle associations is markedly diminished (Cohen, 2022).

Future research could consider using additional observation tools that explicitly focus on science-specific contexts. For example, including indicators of factual content accuracy or the appropriate use of science-specific vocabulary could be implemented. Regarding the analysis of children's participation, this study focused on the quantity of verbal contributions but did not assess their quality. This limitation concerns important aspects such as the content-related accuracy of children's verbal contributions and the use of domain-specific vocabulary, which have been shown to play a central role in the development of science knowledge (Henschel et al., 2023). Although the videos were recorded during science activities, the inclusion of all verbal contributions without re-coding to differentiate science-related from off-task contributions, may have limited the conclusiveness of the findings regarding the effects of children's participation on science learning. A more differentiated qualitative analysis of the children's verbal contributions, such as categorizing them as on-task or off-task, or examining reasoning, questioning, and elaboration, could also provide deeper insights into how participation influences the development of science knowledge. In addition to verbal participation, the video recordings offer the potential to evaluate children's non-verbal communication, which could provide further insight into the individual children's engagement in the dialogs.

Concluding implications: fostering verbal participation to support learning

This study examined the role of instructional support, focusing on the CLASS Pre-K dimensions of Concept Development and Language Modeling, as well as children's verbal participation in the development of language and science knowledge in early childhood education. By analyzing 51 video-recorded science activities, the study highlights the interplay between the quality of teacher-child interactions and individual participation, using multilevel regression analyses. The results demonstrated that the length of children's turns had a significant positive effect on language outcomes. In contrast, the quality of instructional support did not show a significant effect on learning in either domain. However, exploratory analyses suggested a positive interaction effect between the length of children's turns and the quality of Concept Development (CLASS Pre-K, Pianta et al., 2008) on children's science knowledge development, although this finding requires further validation with more robust data. Prior knowledge emerged as a significant predictor for both language and science knowledge development, with prior language skills also shown to predict science knowledge outcomes.

These findings provide a basis for practice-oriented recommendations on how preschool teachers can support children's language and science learning through high-quality interactions and verbal participation: First, science-related activities provide a meaningful and authentic context for fostering both language and science learning, as they naturally invite children to observe, describe, compare, and explain (Peterson & French, 2008). To realize this potential, preschool teachers should use strategies that elicit extended verbal contributions. For instance, asking open-ended questions (Lee & Kinzie, 2012) and waiting for a few seconds for a response can increase the length of children's answers. Second, particular attention should be given to children who rarely or never participate verbally, as limited verbal engagement places them at greater risk of missing out on both language and domain-specific learning opportunities. Teachers should therefore monitor not only the overall balance between teacher and child talk, but also the balance of participation among individual children. Third, targeted professional development is essential to support preschool teachers in implementing these practices. Such programs should raise awareness of the importance of scaffolding and verbal participation and provide concrete strategies, such as those mentioned above, for use in everyday routines and domain-specific activities (e.g., science), in order to provide more inclusive learning opportunities in preschool settings.

Acknowledgments

We want to thank the preschool teachers, their institutions, the children and their parents who participated in the project. We would also like to thank the student assistants who contributed to the completion of the study. Special thanks go to Magdalene Jesche, who helped us with the statistical analyses.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Federal Ministry of Education and Research (BMBF) under Grant [01NV2110].

ORCID

Ada Cecil Haen http://orcid.org/0009-0009-2796-9288
Mirjam Steffensky http://orcid.org/0000-0002-7157-6088
Anika Bürgermeister http://orcid.org/0000-0002-1528-8540
Ilonca Hardy http://orcid.org/0000-0002-3195-1373
Miriam Leuchter http://orcid.org/0000-0002-7962-6561

Sina-Kristin Koschick http://orcid.org/0009-0003-4075-8131
Lukas Lazzara http://orcid.org/0000-0002-9570-067X
Henrik Saalbach http://orcid.org/0000-0003-4387-9214

Author contributions

Ada Cecil Haen: Writing – original draft, Conceptualization, Data curation, Formal analysis, Methodology. Mirjam Steffensky: Writing – review and editing, Conceptualization, Methodology, Funding acquisition, Project administration. Anika Bürgermeister: Writing – review and editing, Conceptualization, Data curation, Methodology, Funding acquisition, Project administration. Ilonca Hardy: Writing – review and editing, Conceptualization, Methodology, Funding acquisition, Project administration. Miriam Leuchter: Writing – review and editing, Conceptualization, Methodology, Funding acquisition, Project administration. Franziska Bednarski: Data curation, Formal analysis. Sina-Kristin Koschick: Data curation. Lukas Lazzara: Data curation. Henrik Saalbach: Writing – review and editing, Conceptualization, Methodology, Funding acquisition, Project administration.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Anders, Y., & Oppermann, E. (2024). Frühpädagogische Qualität in Kindertageseinrichtungen: Eine Erweiterung des Struktur-Prozess-Modells [Quality in early childhood education and care: A theoretical research synthesis and an integral framework of pedagogical quality]. *Zeitschrift für Erziehungswissenschaft*. Advance online publication. *27*(2), 551–577. https://doi.org/10.1007/s11618-024-01218-7

Autorengruppe Bildungsberichterstattung. (2020). *Bildung in Deutschland, 2020: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt [Education in Germany, 2020: An indicator-based report with an analysis of education in a digitalised world]*. https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2020/pdf-dateien-2020/bildungsbericht-2020-barrierefrei.pdf10.3278/6001820gw

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bowlby, J. (1982). *Attachment and loss* (2 ed. Vol. 1). Basic Books.

Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning & Instruction*, *21*(1), 95–108. https://doi.org/10.1016/j.learninstruc.2009.11.004

Burchinal, M. (2017). Measuring early care and education quality. *Child Development Perspectives*, *12*(1), 3–9. https://doi.org/10.1111/cdep.12260

Bürgermeister, A., Große, G., Leuchter, M., Studhalter, U., & Saalbach, H. (2019). Interaktion von pädagogischen Fachkräften und Kindern in naturwissenschaftlichen Lerngelegenheiten im Kindergarten [Teacher-child interaction in preschool science learning: Focussing on quantitative aspects of children's talk]. *Frühe Bildung*, *8*(1), 13–21. https://doi.org/10.1026/2191-9186/a000406

Callanan, M. A., & Valle, A. (2008). Co-constructing conceptual domains through family conversations and activities. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 49, pp. 147–165). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(08)00004-2

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record the Voice of Scholarship in Education*, *112*(3), 579–620. https://doi.org/10.1177/016146811011200303

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cohen, J. (2022). *Applied multiple regression/correlation analysis for the behavioral sciences* (Third ed.). Routledge. https://doi.org/10.4324/9780203774441

Darling-Hammond, L. (2017). *Empowered educators: How high-performing systems shape teaching quality around the world*. Wiley.

Decristan, J., Fauth, B., Heide, E. L., Locher, F. M., Troll, B., Kurucz, C., & Kunter, M. (2020). Individuelle Beteiligung am Unterrichtsgespräch in Grundschulklassen: Wer ist (nicht) beteiligt und welche Konsequenzen hat das für den Lernerfolg? [Students' differential participation in classroom discourse in primary schools: Who participates (not), and what are the consequences for student learning?]. *Zeitschrift Für Pädagogische Psychologie*, *34*(3–4), 171–186. https://doi.org/10.1024/1010-0652/a000251

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Eshach, H., & Fried, M. N. (2005). Should science be taught in early childhood? *Journal of Science Education and Technology*, *14*(3), 315–336. https://doi.org/10.1007/s10956-005-7198-9

Evertson, C. M., & Weinstein, C. S. (Eds.). (2011). *Handbook of classroom management: Research, practice, and contemporary issues*. Routledge.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential for the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Gibbons, P. (2015). *Scaffolding language, scaffolding learning: Teaching English language learners in the mainstream classroom* (2nd ed.). Heinemann.

Gropen, J., Kook, J. F., Hoisington, C., & Clark-Chiarelli, N. (2017). Foundations of science literacy: Efficacy of a preschool professional development program in science on classroom instruction, teachers' pedagogical content knowledge, and children's observations and predictions. *Early Education & Development*, *28*(5), 607–631. https://doi.org/10.1080/10409289.2017.1279527

Haen, A. C., Hardy, I., Saalbach, H., Leuchter, M., & Steffensky, M. (2025). Preschool Teachers' Language Scaffolding in Early Science Education in Linguistically Heterogeneous Groups. In J. Ziehm-Eicher, M. Hasselhorn, & H.-G. Roßbach (Eds.), *Kinder mit erhöhtem Risiko für Bildungsmisserfolg* Edition ZfE (Vol. 20, pp. 129–162). Springer VS; Springer Fachmedien Wiesbaden GmbH. https://doi.org/10.1007/978-3-658-48674-7_5

Hamel, E., Joo, Y., Hong, S. -Y., & Burton, A. (2021). Teacher questioning practices in early childhood science activities. *Early Childhood Education Journal*, *49*(3), 375–384. https://doi.org/10.1007/s10643-020-01075-z

Hardy, I., Sauer, S., & Saalbach, H. (2019). Empirische Arbeit: Frühe sprachliche Bildung im Kontext Naturwissenschaften: Effekte einer Intervention im Kindergarten. *Psychologie in Erziehung und Unterricht*, *66*(3).

Haug, B. S., & Ødegaard, M. (2014). From words to concepts: Focusing on word knowledge when teaching for conceptual understanding within an inquiry-based science setting. *Research in Science Education*, *44*(5), 777–800. https://doi.org/10.1007/s11165-014-9402-5

Henrichs, L. F., & Leseman, P. P. (2014). Early science instruction and academic language development can go hand in hand. The promising effects of a low-intensity teacher-focused intervention. *International Journal of Science Education*, *36*(17), 2978–2995. https://doi.org/10.1080/09500693.2014.948944

Henschel, S., Heppt, B., & Hardy, I. (2023). Structure and associations of science vocabulary, general academic vocabulary, and science knowledge. *The Journal of Educational Research*, *116*(6), 356–370. https://doi.org/10.1080/00220671.2023.2269545

Herrmann, A., Bürgermeister, A., Lange-Schubert, K., & Saalbach, H. (2021). Die Bedeutung von Partizipation und Scaffolding für die Leistung im naturwissenschaftlichen Sachunterricht in Klassen mit hohem und niedrigem Anteil mehrsprachiger Schüler*innen [The role of participation and scaffolding for science learning in primary school classes with high and low proportion of multilingual pupils]. *Zeitschrift Für Grundschulforschung*, *14*(2), 305–323. https://doi.org/10.1007/s42278-021-00112-z

Hofkens, T., Pianta, R. C., & Hamre, B. K. (2023). Teacher-student interactions: Theory, measurement, and evidence for universal properties that support students' learning across countries and cultures. In R. Maulana, M. Helms-Lorenz, & R. M. Klassen (Eds.), *Effective teaching around the world* (pp. 399–422). Springer International Publishing. https://doi.org/10.1007/978-3-031-31678-4_18

Hormann, O., & Skowronek, M. (2019). Wie adaptiv ist der Einsatz von Sprachlehrstrategien in Kitas? Ergebnisse einer Videoanalyse. *Frühe Bildung*, *8*(4), 194–199. https://doi.org/10.1026/2191-9186/a000444

Howard, S. J., Lewis, K. L., Walter, E., Verenikina, I., & Kervin, L. K. (2024). Measuring the quality of adult–child interactions in the context of ECEC: A systematic review on the relationship with developmental and educational outcomes. *Educational Psychology Review*, *36*(1). https://doi.org/10.1007/s10648-023-09832-3

Hu, B. Y., Dieker, L., Yang, Y., & Yang, N. (2016). The quality of classroom experiences in Chinese kindergarten classrooms across settings and learning activities: Implications for teacher preparation. *Teaching & Teacher Education*, *57*, 39–50. https://doi.org/10.1016/j.tate.2016.03.001

Hu, B. Y., Fan, X., Wu, Z., LoCasale-Crouch, J., Yang, N., & Zhang, J. (2017). Teacher-child interactions and children's cognitive and social skills in Chinese preschool classrooms. *Children & Youth Services Review*, *79*, 78–86. https://doi.org/10.1016/j.childyouth.2017.05.028

IBM Corp. (2023). *IBM SPSS statistics for windows* (version 29.0.1.0.) [Computer software].

Jamshidian, M., Jalal, S., & Jansen, C. (2014). MissMech: An R package for testing homoscedasticity, multivariate normality, and missing completely at random (MCAR). *Journal of Statistical Software*, *56*(6). https://doi.org/10.18637/jss.v056.i06

Kempert, S., Schalk, L., & Saalbach, H. (2019). Übersichtsartikel: Sprache als Werkzeug des Lernens: Ein Überblick zu den kommunikativen und kognitiven Funktionen der Sprache und deren Bedeutung für den fachlichen Wissenserwerb [Language as tool for learning: A review of the communicative and cognitive functions of language and its relevance for academic learning]. *Psychologie in Erziehung Und Unterricht*, *66*(3). https://doi.org/10.2378/PEU2018.art19d

Kohl, K., Willard, J. A., Agache, A., Bihler, L. -M., & Leyendecker, B. (2019). Classroom quality, classroom composition, and age at entry: Experiences in early childhood education and care and single and dual language learners' German vocabulary. *AERA Open*, *5*(1), 233285841983251. https://doi.org/10.1177/2332858419832513

Koschick, S.-K., Hardy, I., Bürgermeister, A., & Venitz, L. (2024). Partizipation und diskursive Orientierung in Professionellen Lerngemeinschaften des Elementarbereichs [Participation and discursive orientation in professional learning communities in early childhood education]. *Erziehung und Unterricht*, *174*(3-4), 324–333.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805–820. https://doi.org/10.1037/a0032583

Lee, Y., & Kinzie, M. B. (2012). Teacher question and student response with regard to cognition and language use. *Instructional Science*, *40*(6), 857–874. https://doi.org/10.1007/s11251-011-9193-2

Leuchter, M., & Saalbach, H. (2014). Verbale Unterstützungsmaßnahmen im Rahmen eines naturwissenschaftlichen Lernangebots in Kindergarten und Grundschule [Teachers' tutoring activities in the context of science education in kindergarten and primary school]. *Unterrichtswissenschaft*, *42*(2), 117–131.

Levickis, P., Cloney, D., Roy-Vallières, M., & Eadie, P. (2023). Associations of specific indicators of adult-child interaction quality and child language outcomes: What teaching practices influence language? *Early Education & Development*, *35*(4), 647–666. https://doi.org/10.1080/10409289.2023.2193857

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, *106*(4), 532–575. https://doi.org/10.1086/444197

Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, *58*(2), 103–117. https://doi.org/10.1026/0033-3042.58.2.103

Mannel, S., Hardy, I., Sauer, S., & Saalbach, H. (2016). Sprachliches Scaffolding zur Unterstützung naturwissenschaftlichen Lernens im Kindergartenalter [Language scaffolding to foster science learning in kindergarten-aged children]. In E. Tschirner, O. Bärenfänger, & J. Möhring (Eds.), *Sammelband zum Symposium"Kompetenzprofile Deutsch als fremde Bildungssprache"* (pp. 97–113). Stauffenburg Verlag.

Mercer, N., & Dawes, L. (2014). The study of talk between teachers and students, from the, 1970s, until the 2010s. *Oxford Review of Education*, *40*(4), 430–445. https://doi.org/10.1080/03054985.2014.934087

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher*, *45*(1), 18–35. https://doi.org/10.3102/0013189X16633182

Muhonen, H., Rasku-Puttonen, H., Pakarinen, E., Poikkeus, A. -M., & Lerkkanen, M. -K. (2016). Scaffolding through dialogic teaching in early school classrooms. *Teaching & Teacher Education*, *55*, 143–154. https://doi.org/10.1016/j.tate.2016.01.007

National Research Council. (2012). *A framework for K-12 science Education: A framework for K-12 science Education: Practices, crosscutting concepts, and Core ideas*. National Academies Press. https://doi.org/10.17226/13165

OECD. (2018). *Engaging young children: Lessons from research about quality in early childhood education and care*. Starting Strong. OECD Publishing. https://doi.org/10.1787/9789264085145-en

Patall, E. A., Yates, N., Lee, J., Chen, M., Bhat, B. H., Lee, K., Beretvas, S. N., Lin, S., Man Yang, S., Jacobson, N. G., Harris, E., & Hanson, D. J. (2024). A meta-analysis of teachers' provision of structure in the classroom and students' academic competence beliefs, engagement, and achievement. *Educational Psychologist*, *59*(1), 42–70. https://doi.org/10.1080/00461520.2023.2274104

Pauli, C., & Lipowsky, F. (2007). Mitmachen oder zuhören? Mündliche Schülerinnen- und Schülerbeteiligung im Mathematikunterricht [Student Participation in Whole-Class Discussions and Teacher-Student Interactions in Mathematics Classrooms]. *Unterrichtswissenschaft*, *35*(2). https://doi.org/10.25656/01:5488

Peters, S., Wolstein, K., Mischo, C., & Ehm, J. -H. (2020). Wissen, Wissensorientierung und die Qualität von Fachkraft-Kind-Interaktionen: Eine Analyse ihrer Zusammenhänge [Knowledge, orientation to knowledge and quality of interaction of early childhood teachers and children: An analysis of their relationships]. *Zeitschrift für Erziehungswissenschaft*, *23*(6), 1227–1250. https://doi.org/10.1007/s11618-020-00979-1

Peterson, S. M., & French, L. (2008). Supporting young children's explanations through inquiry science in preschool. *Early Childhood Research Quarterly*, *23*(3), 395–408. https://doi.org/10.1016/j.ecresq.2008.01.003

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109–119. https://doi.org/10.3102/0013189X09332374

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system manual Pre-K*. Teachstone.

Praetorius, A. -K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning & Instruction*, *31*, 2–12. https://doi.org/10.1016/j.learninstruc.2013.12.002

R Core Team. (2025). *R: A language and environment for statistical computing* (version 4.5.0) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68
Saçkes, M., Trundle, K. C., Bell, R. L., & O'Connell, A. A. (2011). The influence of early science experience in kindergarten on children's immediate and later science achievement: Evidence from the early childhood longitudinal study. *Journal of Research in Science Teaching*, *48*(2), 217–235. https://doi.org/10.1002/tea.20395
Schmerse, D., Dominke, H., Mohr, J., & Steffensky, M. (2024). Children's understanding of scientific inquiry: The role of instructional support and comparison making. *Journal of Educational Psychology*, *116*(2), 233–255. https://doi.org/10.1037/edu0000836
Sedláček, M., & Šeďova, K. (2020). Are student engagement and peer relationships connected to student participation in classroom talk? *Learning, Culture & Social Interaction*, *26*, 100411. https://doi.org/10.1016/j.lcsi.2020.100411
Sedova, K., Sedlacek, M., Svaricek, R., Majcik, M., Navratilova, J., Drexlerova, A., Kychler, J., & Salamounova, Z. (2019). Do those who talk more learn more? The relationship between student classroom talk and student achievement. *Learning & Instruction*, *63*, 101217. https://doi.org/10.1016/j.learninstruc.2019.101217
Simon, S., & Sachse, S. (2011). Sprachförderung in der Kindertagesstätte – verbessert ein Interaktionstraining das sprachförderliche Verhalten von Erzieherinnen? [Promoting language skills in day care - can interaction training improve childhood educators' language-promoting behaviour?]. *Empirische Pädagogik*, *25*(4), 462–480.
Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R., & Bell, D. (2002). *Researching effective pedagogy in the early years*. University of London, Institute of Education, Department of Educational Studies.
Steffensky, M., Anders, Y., Barenthien, J., Hardy, I., Leuchter, M., Oppermann, E., Taskinen, P., & Ziegler, T. (2018). Early Steps into Science – EASI Science. Wirkungen früher naturwissenschaftlicher Bildungsangebote auf die naturwissenschaftlichen Kompetenzen von Fachkräften und Kindern [Early Steps into Science – EASI Science: Effects of early childhood science education programs on teachers' and children's science competencies. In Y. Anders, J. Barenthien, I. Hardy, A. Hartinger, R. Kästner, M. Leuchter, E. Oppermann, S. Pauen, A. Rank, M. Steffensky, P. Taskinen, S. Tietze, A. Wildemann, & T. Ziegler (Eds.), *Wirkungen naturwissenschaftlicher Bildungsangebote auf pädagogische Fachkräfte und Kinder* (Vol. 10, pp. 50–137). https://doi.org/10.2307/j.ctvmx3jn8.8.
Steffensky, M., Haen, A. C., Lazzara, L., Koschick, S.-K., Bürgermeister, A., Venitz, L., Saalbach, H., Leuchter, M., & Hardy, I. (2025). Diagnostik und adaptive Förderung in der frühen naturwissenschaftlichen Bildung: Ein App-basierter Ansatz zur Professionalisierung von pädagogischen Fachkräften [Diagnostics and adaptive support in early childhood science education: An app-based professional development approach for preschool teachers]. In K. Beck, R. A. Ferdigg, D. Katzenbach, J. Kett-Hauser, S. Laux, & M. Urban (Eds.), *Förderbezogene Diagnostik in der inklusiven Bildung: Kompetenzbereiche – Fachdidaktik* (pp. 73–93). Waxmann.
Stuck, A., Kammermeyer, G., & Roux, S. (2016). The reliability and structure of the Classroom Assessment Scoring System in German pre-schools. *European Early Childhood Education Research Journal*, *24*(6), 873–894. https://doi.org/10.1080/1350293X.2016.1239324
Studhalter, U. T., Leuchter, M., Tettenborn, A., Elmer, A., Edelsbrunner, P. A., & Saalbach, H. (2021). Early science learning: The effects of teacher talk. *Learning and Instruction*, *71*, 101371. https://doi.org/10.1016/j.learninstruc.2020.101371
Suchodoletz, A. V., Fäsche, A., Gunzenhauser, C., & Hamre, B. K. (2014). A typical morning in preschool: Observations of teacher-child interactions in German preschools. *Early Childhood Research Quarterly*, *29*(4), 509–519. https://doi.org/10.1016/j.ecresq.2014.05.010
Suchodoletz, A. V., Lee, D. S., Henry, J., Tamang, S., Premachandra, B., & Yoshikawa, H. (2023). Early childhood education and care quality and associations with child outcomes: A meta-analysis. *PLOS ONE*, *18*(5), e0285985. https://doi.org/10.1371/journal.pone.0285985
Sylva, K., Melhuish, E., Sammons, P., & Siraj-Blatchford, I. (2004). The effective provision of pre-school education [EPPE] project: Technical paper 12 the final report: Effective pre-school education.
Thomas, S. L., & Heck, R. H. (2001). Analysis of large-scale secondary data in higher education research: Potential perils associated with complex sampling designs. *Research in Higher Education*, *42*(5), 517–540. https://doi.org/10.1023/A:1011098109834
Thorpe, K., Rankin, P., Beatton, T., Houen, S., Sandi, M., Siraj, I., & Staton, S. (2020). The when and what of measuring ECE quality: Analysis of variation in the classroom assessment scoring system (CLASS) across the ECE day. *Early Childhood Research Quarterly*, *53*, 274–286. https://doi.org/10.1016/j.ecresq.2020.05.003
Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.
Tompkins, V., Zucker, T. A., Justice, L. M., & Binici, S. (2013). Inferential talk during teacher-child interactions in small-group play. *Early Childhood Research Quarterly*, *28*(2), 424–436. https://doi.org/10.1016/j.ecresq.2012.11.001
Ulferts, H., Wolf, K. M., & Anders, Y. (2019). Impact of process quality in early childhood education and care on academic outcomes: Longitudinal meta-analysis. *Child Development*, *90*(5), 1474–1489. https://doi.org/10.1111/cdev.13296
van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Chapman & Hall/CRC Interdisciplinary statistics series. CRC Press Taylor & Francis Group.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, *22*(3), 271–296. https://doi.org/10.1007/s10648-010-9127-6

VERBI Software. (2021). *Maxqda 2022* [Computer software]. maxqda.com

Vosniadou, S. (2007). Conceptual change and education. *Human Development*, *50*(1), 47–54. https://doi.org/10.1159/000097684

Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes*. Harvard University Press.

Wang, S., Wang, X., Liang, L., & Xie, H. (2022). Teacher-child interaction domains measured by the CLASS and Children's Pre-literacy skills: A systematic Review and meta-analysis. *Early Education & Development*, *35*(3), 1–23. https://doi.org/10.1080/10409289.2022.2159259

Wechsler, D. (2018). *Wechsler preschool and primary scale of intelligence – fourth edition: Deutsche Bearbeitung herausgegeben von Franz Petermann und Monika Daseking* [German adaptation by Franz Petermann and Monika Daseking]. Pearson.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. https://doi.org/10.1002/sim.4067

Wirts, C., & Cordes, A. -K. (2021). Wer spricht denn da? Redeanteile und das sprachanregende Potenzial verschiedener Fragetypen [Look who's talking!? Talk ratios and the encouraging potential of different types of questions]. *Forschung Sprache*, *9*(2), 151–159.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Zimmerman, C., & Klahr, D. (2018). Development of scientific thinking. In J. T. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (pp. 1–25). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119170174.epcn407