Language in Good Company: Unraveling Peer Effects on Language Development in Early Education Using Integrative Data Analysis of Six German Studies

Franziska Hürlimann[1], Oliver Lüdtke[2, 3], and Daniel Schmerse[1]

[1] School of Education, University of Applied Sciences and Arts Northwestern Switzerland

[2] Department of Educational Measurement and Data Science, Leibniz Institute for Science and Mathematics Education (IPN), Kiel, Germany

[3] Centre for International Student Assessment (ZIB), Munich, Germany

Author Accepted Manuscript (AAM): This manuscript is not the copy of record and may not exactly replicate the final, authoritative document published in the APA journal. The final article is available, upon publication, at: 10.1037/edu0000977.

Supplemental materials available at: https://doi.org/10.1037/edu0000977.supp

Author Note

Franziska Hürlimann https://orcid.org/0009-0002-2567-8136

Oliver Lüdtke https://orcid.org/0000-0001-9744-3059

Daniel Schmerse https://orcid.org/0000-0001-5799-6724

This research was funded in whole, or in part, by the Swiss National Science Foundation (SNSF) (Grant 213572).

This study utilized secondary data provided by the Research Data Center at the Leibniz Institute for Educational Trajectories (FDZ-LIfBi) and the Research Data Center at the Institute for Educational Quality Improvement (FDZ at IQB) in Berlin.

Earlier versions of this research were presented at the 2024 conference of the German Society for Empirical Educational Research (GEBF), at the 2024 International Congress of the Study of Child Language (IASCL), and at the 2024 conference of the German Psychological Society (DGPs).

Correspondence concerning this article should be addressed to Franziska Hürlimann, School of Education, University of Applied Sciences and Arts Northwestern Switzerland, FHNW Campus, Bahnhofstrasse 6, 5210 Windisch, Switzerland. Email: franziska.huerlimann@fhnw.ch

Abstract

Peers play a central role in children's learning environments. Nevertheless, the empirical evidence regarding their influence on language development remains inconclusive. This study employed an integrative data analysis (IDA) approach to systematically investigate the effect of peers' average language abilities on individual language development across six German longitudinal studies. The analysis encompassed over 15,000 children (ages 3-10) from more than 2,500 linguistically and socioeconomically diverse preschool and primary school classrooms. In a consistent methodological approach, we estimated peer effects and moderation effects within each data set separately by specifying multilevel regression models, controlling for key individual background variables and peers' socioeconomic status. These effect sizes were then meta-analytically aggregated. The results demonstrated considerable variability across studies, with no significant aggregated main peer effect. However, a significant interaction between peers' language abilities and a child's prior skill level emerged. Children with lower initial language skills benefited, on average, more from being in classrooms with higher average peer language abilities, while higher skilled children's development was less affected by their peers. By employing IDA with meta-analytic aggregation, the study offers overarching conclusions that extend beyond the scope of individual studies, contributing to a more nuanced understanding of peer influences on language development across diverse educational settings. The findings underscore the crucial role of peers in shaping language development trajectories and suggest that considering peer effects in educational strategies and thoughtful classroom composition could enhance language development opportunities, especially for less skilled children.

Key words: peer effects; language development; integrative data analysis; early childhood education; classroom composition

Educational Impact And Implications Statement

Integrating results from six German studies, this research shows that children with weaker language skills tend to benefit more when their peers have stronger language abilities in preschool and early primary school. These findings highlight the important, yet nuanced role peers play in language development and suggest that thoughtful classroom composition may help reduce disparities in children's language skills. Teachers and policymakers might consider peer language ability levels when organizing groups to create more opportunities for children with weaker skills to learn from their peers.

Language in Good Company: Unraveling Peer Effects on Language Development in Early Education Using Integrative Data Analysis of Six German Studies

From the moment children enter preschool or primary school, they spend a substantial part of their day with their classroom peers. Peers significantly affect various developmental outcomes, including cognitive, academic, and social-emotional skills (DeLay, Zhang, et al., 2016; Lin et al., 2024; Rubin et al., 2015). Given this profound impact, classroom peer groups serve as pivotal developmental contexts, creating a variety of learning opportunities and experiences (Bandura, 1971; Bronfenbrenner & Morris, 2006), which may differ depending on the composition of these groups. Numerous studies focusing on school-aged children have reported that different aspects of classroom composition, such as social and ethnic-cultural backgrounds as well as skill levels in various domains, are related to students' learning and academic achievement (M. Becker et al., 2022; De Fraine et al., 2003; Rjosk et al., 2017; Sacerdote, 2011; Van Ewijk & Sleegers, 2010a, 2010b). This *compositional* or *peer effect* is operationalized as the effect of aggregated individual characteristics at the group level on an outcome, beyond the impact of individual characteristics themselves (Raudenbush & Bryk, 2002; Sacerdote, 2011). Importantly, this phenomenon is not limited to primary and secondary education. A growing body of research has begun to reveal similar peer effects within preschool contexts, demonstrating significant impacts on a range of developmental outcomes (Choi et al., 2018; Henry & Rickman, 2007; Mashburn et al., 2009; Ready & Reid, 2019; Schmerse & Hepach, 2021).

One particularly relevant domain of development in the preschool and early primary school years is language ability, which is foundational for further learning and overall academic success (Bleses et al., 2016; Schoon et al., 2021). Thus, a key question is to what extent classroom peers contribute to language development in these early educational contexts. Findings from studies addressing this question are mixed. Research has indicated

positive associations between peer characteristics and individual language development (Lin et al., 2023; Mashburn et al., 2009; Reid & Ready, 2013), with some studies suggesting stronger effects for certain groups of children (Justice et al., 2011; Ribeiro et al., 2017; Schmerse, 2021). Conversely, several studies did not find effects both in terms of a main peer effect and moderation effects (Choi et al., 2018; Kohl et al., 2022; Yang et al., 2023). Therefore, the existing evidence remains inconclusive, likely due to considerable variation in the methodological approaches adopted across these studies.

Addressing the inconsistencies, our study seeks to advance the existing peer-effect literature by conducting an integrative data analysis (IDA) of six German studies. Using a consistent analytical approach across different samples suited to produce robust evidence, we aim to explore the overall impact of peers' language skills on individual language development while controlling for the socioeconomic composition of the peer group and relevant individual variables. Additionally, we investigate whether peer effects are moderated by children's family backgrounds and initial language abilities. Thereby, we aim to contribute to a more comprehensive understanding of peer effects on language skills in early education settings.

Peer Effects in Language Development

The trajectories of children's language development are strongly shaped by contextual factors and interactions within their social environments (Hoff, 2006; Rowe & Weisleder, 2020; Tomasello, 2003). This influence is particularly well documented for family factors, such as socioeconomic status (SES) and the quality and quantity of parental language input (Hart & Risley, 1995; Hirsh-Pasek et al., 2015; Pace et al., 2017). In addition to the family environment, early childhood education (ECE) settings play a crucial role in language development. These settings are a common experience for the majority of children (OECD, 2022). Research on ECE settings' role in language development has primarily focused on

structural quality and teacher-child interactions (Cabell et al., 2015; Justice et al., 2018; Keys et al., 2013; Ulferts et al., 2019). However, there is a growing focus on the role of peers as a key component of the ECE environment (Chen et al., 2020; Gámez et al., 2019; Purtell et al., 2021; Washington-Nortey et al., 2022).

Several studies have examined compositional effects on language abilities, particularly looking at factors such as the proportion of children with migration backgrounds or those who speak different languages at home. Some studies have reported that a higher percentage of peers from these minority groups is associated with lower individual language skills (Atkins-Burnett et al., 2017; Garcia, 2018; Niklas et al., 2011; Reid & Ready, 2013), while others have found no significant effects (Bihler et al., 2018; Kohl et al., 2019; Willard et al., 2019). Moreover, research has indicated that the SES composition of the classroom is a salient predictor of language development (Bihler et al., 2018; Reid & Ready, 2013; Weiland & Yoshikawa, 2014).

In addition to class composition in terms of these demographic peer characteristics, studies have shown that the ability levels of peers significantly impact a child's language skills (Henry & Rickman, 2007). Specifically, the average language skills of classroom peers may shape the linguistic environment and influence individual language development. However, evidence on the impact of peers' average language level on individual language abilities is mixed. Some studies have supported the hypothesis that peers' average language level has a direct positive effect on individual language abilities (Atkins-Burnett et al., 2017; Foster et al., 2020; Garcia, 2021; Lin et al., 2023; Mashburn et al., 2009). For example, Mashburn et al. (2009) found that peers' fall expressive vocabulary positively predicted preschoolers' expressive and receptive vocabulary in spring, even when accounting for various background characteristics and classroom quality. Foster et al. (2020) identified a peer effect of mean receptive vocabulary level on children's individual receptive vocabulary

and auditory comprehension, after controlling for classroom quality, age composition, and peers' behavior problems. Conversely, other studies have reported no significant main peer effects of average language skills (Choi et al., 2018; Justice et al., 2011; Kohl et al., 2022; Ribeiro et al., 2017; Yang et al., 2023).

Yang et al. (2023) analyzed two large data sets from the United States but found little evidence of associations between average peer skills and children's language abilities. They pointed to the wide range of effect sizes reported in prior research, ranging from negligible to substantial impacts. These authors suggested that this disparity may be due to differences in sample characteristics and methodological differences and argued that such variations make it difficult to compare findings and draw robust conclusions about peer effects (Yang et al., 2023).

Taken together, there is mixed evidence regarding the effect that peers may have on children's early language development. An analysis of multiple studies under a unified analytical framework could help reconcile these disparities and provide a more robust and reliable understanding of peer effects on early language development.

The Role of the Socioeconomic Composition

In addition to the average skill levels, the SES of classroom peers is also a key feature of classroom composition. Research with older students provides evidence that a higher socioeconomic status of their school or classroom is associated with higher academic achievement across different subjects (for meta-analyses see Tan et al., 2023; Van Ewijk & Sleegers, 2010b). However, the effect of socioeconomic composition typically tends to diminish when measures of student's prior achievement or general cognitive abilities are taken into account (Gutiérrez, 2023; Tan et al., 2023). For example, two longitudinal studies with German primary school students found that when considering both socioeconomic and

achievement composition simultaneously, only the latter was associated with students' reading achievement (Bellin, 2009; Nikolova, 2011).

Whether these findings extend to the preschool context and other domains is less clear. A large body of research has consistently demonstrated the critical role of SES background on language development at the individual level. Children from low-SES families often start preschool or school with poorer language skills compared to their peers from higher-SES families (Pace et al., 2017). Targeted preschool programs for low-income children (e.g., Head Start, Weiland &Yoshikawa, 2014) and residential segregation frequently result in children attending preschools and primary schools that are predominantly composed of peers from similar socioeconomic backgrounds (B. Becker & Schober, 2017; Hogrebe et al., 2021). Although evidence from early education contexts is less extensive, it also suggests that the socioeconomic composition influences children's skill development (Coley et al., 2019; Dumont & Ready, 2020; Miller et al., 2017; Schechter & Bye, 2007). Specifically, for language skills, Reid and Ready (2013) showed that the average classroom SES is positively linked to both expressive and receptive vocabulary, independent of classroom quality. Weiland and Yoshikawa (2014) reported similar findings, indicating that a larger proportion of high-SES peers predicted better receptive vocabulary. Interestingly, in their study, SES composition remained a significant and stable predictor even when accounting for peer language skills and other peer and classroom characteristics. This contrasts with the findings in older students mentioned above where SES composition no longer shows an effect when considered simultaneously with achievement composition (Bellin, 2009; Nikolova, 2011). It also differs from an Australian study with preschoolers (Niklas & Tayler, 2018), which found that SES composition did not predict children's verbal abilities. However, Niklas and Tayler used cross-sectional data and did not account for peer

language skills, focusing instead on the average general cognitive ability level, which showed stronger associations with individual verbal abilities than average SES.

Given the evidence for the impact of school or class SES on child development, it is crucial to consider classroom peer SES simultaneously when investigating the effects of peer skill levels. Omitting either of these compositional measures can potentially lead to biased estimates (Gutiérrez, 2023). Moreover, the findings from preschool contexts are contradictory regarding the unique and combined effects of peer ability levels and socioeconomic backgrounds. Accounting for both compositional characteristics simultaneously, enables a more comprehensive conclusion about what remains of the peer language effect when SES composition is controlled for.

Are Peer Effects Moderated by Children's Background and Language Skills?

Diversity in home language environments, prior language skills, and family SES are closely intertwined and play significant roles in shaping children's language learning trajectories (Hoff, 2013; Pace et al., 2017; T. Schneider & Linberg, 2022). On average, these differences in linguistic and socioeconomic backgrounds can lead to varied proficiency levels in the societal majority language (Hoff, 2013; Pace et al., 2017). This raises the question of whether classroom peers might contribute differentially to the development of children from multilingual homes, those with different initial skill levels, and those from varied socioeconomic backgrounds.

Few studies have investigated the moderating role of these three aspects. Previous findings and theoretical considerations point to two contrasting patterns. On the one hand, *compensatory effects* suggest that disadvantaged children may benefit most from exposure to higher-skilled peer groups. Such groups may provide more opportunities to interact with and observe advanced language models and may help compensate for less enriching language environments at home or provide access to opportunities for learning the societal language in

the first place. If, at the same time, children with higher language skills were less susceptible to peer influence, this could potentially attenuate initial disparities in language abilities (Hanushek et al., 2003; Justice et al., 2011). Some studies support this assumption by demonstrating stronger effects for DLLs (Atkins-Burnett et al., 2017; Schmerse, 2021) or for children from lower-SES families (low maternal education: Ribeiro et al., 2017; low-income households: Yang et al., 2023). Justice et al. (2011) observed that peer effects were most pronounced for children with lower initial skills.

Conversely, the so-called *Matthew effect* suggests that competent peer groups would be most beneficial for children from more privileged backgrounds or those who already possess strong language skills, as they are more capable of leveraging enriched learning opportunities (Mashburn et al., 2009; Walberg & Tsai, 1983). In this scenario, children with lower skills may struggle to keep pace with the more advanced learning environment, which could exacerbate initial disparities. This perspective is supported by studies that identified stronger effects of peer language abilities for children with advanced initial language skills (Mashburn et al., 2009 for receptive vocabulary) or monolingual backgrounds (Foster et al., 2020 for auditory comprehension outcome). Evidence supporting the *Matthew effect* was also noted in a study on first graders' reading abilities (Kuzmina & Ivanova, 2018).

However, other research found no significant moderation of peer effect by DLL status (Kohl et al., 2022; Lin et al., 2023; Yang et al., 2023), prior skills (Choi et al., 2018; Kohl et al., 2022; Ribeiro et al., 2017; Yang et al., 2023), and family SES (Kohl et al., 2022; Yang et al., 2023 for maternal education).

Overall, the available evidence provides inconsistent findings, with some suggesting potential differential effects; however, it remains unclear for which specific groups these effects occur. Notably, the existing studies vary considerably in their methodological approaches and samples. Particularly country-specific differences in the prevalence and

backgrounds of DLLs, the range and impact of socioeconomic disparities, and the structure and policies of ECE systems make it difficult to draw conclusive overarching inferences. Given that the present study used data from the German context, the following section briefly outlines the German ECE and primary school systems to provide the necessary context for our analysis.

ECE and Primary Schools in Germany

Children in Germany start compulsory education when they enter primary school at the age of 6 to 7 years. Prior to that, parents may voluntarily enroll their children in ECE provision, which is separate from the formal school system. Approximately 41% of children under 3 and over 90% of 3 to 6-year-olds attend ECE institutions[1] (Autor:innengruppe Bildungsberichterstattung, 2024).

The cost of ECE in Germany is income-related and publicly subsidized, making it widely accessible. Nevertheless, the age at which children enter ECE and the extent of their attendance varies based on parental needs and availability of spots. In general, parents are free to choose an ECE center, with proximity to their place of residence being an important criterion (B. Becker & Schober, 2017). However, although German ECE centers typically include a mix of socioeconomic, ethnic, and linguistic backgrounds, patterns of segregation exist due to residential segregation and also the possibility of choosing ECE centers. This results in children often attending facilities where their peers come from similar backgrounds. This phenomenon is particularly evident, for example, among children from

[1] In Germany, center-based ECE settings for 3 to 6-year-olds are predominantly referred to as *Kindergarten*. However, in terms of their objectives, these facilities are more akin to what is known as preschool, particularly in the USA. In this study, we primarily use the term preschool, but occasionally refer to Kindergarten when aligning with terminology used in some studies, treating the terms as interchangeable in the German context.

socioeconomically disadvantaged families and DLLs (B. Becker & Schober, 2017; Hogrebe et al., 2021; Riedel et al., 2010).

Pedagogical approaches in ECE emphasize child-initiated activities and free play rather than focusing on formal educational content, fostering an environment where children learn through exploration and social interaction (Kohl et al., 2022). The transition into compulsory schooling occurs at ages 6 to 7, with children typically being assigned to a primary school within their residential area. Primary school comprises 4 years in most German states. Upon completing primary school, students progress to different tracks of secondary education.

The Present Study

Previous studies have varied greatly in their methods and sample characteristics, possibly contributing to the inconsistencies of the empirical findings. In addition, most existing research relies on studies from the United States, with findings from European contexts being scarce (Kohl et al., 2022; Ribeiro et al., 2017; Schmerse, 2021). To provide more robust findings and extend the limited research in European contexts, the present study employs an IDA approach (Graham et al., 2022; Hofer & Piccinin, 2009) across six studies from the German ECE and early primary school context involving meta-analytic aggregation. Using a consistent methodological framework, the present study investigates the following research questions:

RQ1(a): Does the average language level of peers in preschool and primary school classrooms predict individual language abilities?

RQ1(b): To what extent does the peer effect on individual language abilities change when accounting for the group average SES?

RQ2: Are peer effects moderated by children's DLL status, prior language skills, and family socioeconomic background?

The current study expands upon prior research in several important ways: First, analyzing multiple data sets under a uniform approach allows us to evaluate the replicability and generalizability of peer effects while ensuring that observed differences do not stem from methodological discrepancies. Second, the included data sets enable us to examine peer effects in both preschool and early primary school grades, a still understudied context regarding peer effects on language development. Third, we account for the socioeconomic composition of peer groups, an important yet rarely considered factor in preschool and primary school contexts. Fourth, the meta-analytic synthesis provides an overarching summary across the different samples. Overall, with its integrative and systematic approach, this study aims to contribute to a more comprehensive and robust understanding of the independent and moderating impact of peer effects on language development.

Method

The present study utilized secondary data drawn from six German studies. Given that language abilities were the primary construct under investigation, each study included at least two measurement points of language and provided data on relevant child and family background variables. This section details the selected studies and data sets, the measures collected, and the analytical strategies consistently applied across all data sets.

Study Selection and Data Sets

Studies were selected based on predefined criteria to ensure their suitability for addressing our research questions: The studies (a) included children from preschools or primary schools; (b) contained information on class memberships; (c) provided sufficiently large samples at both the individual and group levels; (d) featured reliable and valid measures of language skills from at least two measurement points, enabling control of preceding language skills, and (e) included data on non-verbal cognitive abilities and key familial background characteristics, such as family language, SES, and maternal education.

Our search strategy prioritized relevant and accessible data sets that met these criteria. We focused on major German research repositories, particularly the Research Data Center at the Institute for Educational Quality Improvement (FDZ at IQB)[2] and the German Network of Educational Research Data[3], which systematically collects and provides access to research data from multiple research data centers and repositories.

As the availability of early education research data in Germany is still limited, our initial search targeted studies with preschool and primary school populations. From these, we selected data sets that met our inclusion criteria. To account for newly available studies, we conducted repeated searches between August 2020 and April 2024 in these repositories.

While we do not claim systematic completeness, our search covered the key sources of primary and pre-primary educational research data in Germany. We included all accessible data sets from these sources that met our predefined inclusion criteria. Through this process, we identified and included six studies. Table 1 provides an overview of the included studies, along with the corresponding references and sociodemographic characteristics.

Two of the finally included studies were longitudinal educational panel studies (BiKS-3-10 and NEPS) that followed children from preschool through their school years. In addition, two studies focused on preschool children (SEIKA-NRW LS and PHONO 1), and the remaining two studies included only primary school students (KuL and BiSpra 2). This allowed us to cover a broad age range, from 2.5 to 10 years.

For the BiKS and NEPS studies, we created separate data sets for the preschool and primary school samples to account for transitions into new educational contexts and the inclusion of additional participants, who were not part of the peer group at the previous measurement points. We also generated two separate data sets from the SEIKA study due to

[2] https://www.iqb.hu-berlin.de/fdz/studies/
[3] https://www.forschungsdaten-bildung.de/en/studies/search

the replenishment of the sample in Wave 2 with newly enrolled children. Further details on the design and characteristics of each study can be found in Supplement A. Overall, our analysis comprised 15,062 children across roughly 2,532 classes, with some overlap in the data sets. The proportion of DLLs varied across studies, ranging from 13.63% in the NEPS primary school data set to 55.3% in the BiSpra study, with representing a broad diversity of language backgrounds.

Measures

Language Skills

Language skills were assessed at multiple time points, serving both as an independent variable, measured at the individual level and aggregated at the group level, and as the outcome variable. Detailed information about the tests used across the studies, including information about mean and standard deviation and reliability, is provided in Table 2. From all studies except one (KuL), we employed a measure of receptive vocabulary. Most of the studies administered a German version of the Peabody Picture Vocabulary Test (PPVT). This standardized test requires the child to select one of four pictures that best represents the meaning of a spoken target word. For the primary school sample in the BiKS study, a similar picture selection task was used to assess verbal comprehension (subtest of KFT 1-3 [Cognitive ability test 1-3], Heller & Geisler, 1983). The BiSpra study used a test to assess academic vocabulary. Students completed a cloze task by selecting the best-fitting word from three options, presented audibly and visually (Heppt et al., 2020).

In the KuL study, the original vocabulary measure showed insufficient reliability. As an alternative measure, assessment of phonological awareness (PA) was used from this data set (subtest of FIPS [German version of Performance Indicators in Primary Schools [PIPS]], Bäuerlein et al., 2012). Students were asked to repeat pseudo-words, identify rhymes, and divide words into syllables.

While PA is primarily considered a metalinguistic skill associated with literacy development (Anthony & Francis, 2005), research also demonstrates strong concurrent and longitudinal associations with vocabulary and broader language competencies (e.g., Dickinson et al., 2003; Hipfner-Boucher et al., 2014; Lonigan et al., 2009). Furthermore, bidirectional relations between PA and vocabulary suggest they mutually influence each other during early language development (Gibbs, 2004).

We acknowledge that PA is not synonymous with vocabulary; however, given its established association with vocabulary and other language competencies, we consider PA a reasonable substitute in cases where direct vocabulary measures are unavailable or unsuitable.

Child and Family Background Variables

We included a number of key covariates related to language development to reduce potential selection bias and confounding effects. First, we included measures of non-verbal cognitive abilities to account for the child's cognitive processing skills independent of language competencies. All studies employed standardized tests, including matrix reasoning, figural analogies, pattern construction, or object assembly. For more details on the specific tests used see Table S2 in the Supplemental Materials. Second, we included family SES which is recognized as a critical factor associated with both language skills and selection into preschool and school contexts. Except for the SEIKA study, all studies provided a score of the International Socio-Economic Index of occupational status (ISEI, Ganzeboom et al., 1992). In our analysis, we utilized the highest value reported from either parent (HISEI). In the SEIKA study, information about the equivalized household income was available. Due to the skewed distribution, we applied a logarithmic transformation to the income variable across both SEIKA data sets. Third, we included the highest level of maternal education, measured as completed school degree in the PHONO study, levels of vocational qualification

(BiKS, KuL), and classified as levels according to the International Standard Classification of Education (ISCED; UNESCO, 2012) in the remaining studies. Fourth, for each child, we established DLL status. We operationalized a dichotomous variable for language use at home (i.e. languages used by parents in their interactions with the child) comparable across studies. Children exposed to at least one language other than German, whether in addition to German or exclusively, were classified as DLLs. Children who only experienced German were classified as single language learners (SLLs). For the PHONO data, only information about the child's native language was available (reported by parents). Consequently, children with a native language other than German or multiple native languages were classified as DLLs, while monolingual German-speaking children were classified as SLLs. Finally, the child's gender was coded as a dummy variable.

In addition to serving as covariates to account for individual differences, a child's language ability, DLL status, and SES are subsequently examined as moderators in separate analyses to assess their potential moderating effects.

For analyses, all non-dichotomous variables including the outcome were *z*-standardized to enable meaningful comparisons and facilitate interpretation of regression coefficients.

Group-Level Variables

For our primary predictor, the peer language skills, we calculated the average score of individual language abilities from a time point prior to the outcome measure from all available children within the same classroom. To operationalize the socioeconomic composition of the class, used as an additional control variable, we calculated the average of the individual HISEI scores of children sharing the group or classroom ID. In the SEIKA data, we computed classroom SES based on household income.

The average values derived in this way served as Level 2 predictors in our multilevel analyses. This approach to calculating peer skills, consistent with previous studies (e.g., Justice et al., 2011; Lin et al., 2023; Schmerse, 2021; Yeomans-Maldonado et al., 2019), allows for the estimation of the peer effect as a composition or contextual effect (Raudenbush & Bryk, 2002). Notably, the group averages were aggregated from *z*-standardized individual-level data after imputing missing values.

Analytical Strategy

Integrative Data Analysis

We addressed our research questions within an IDA framework, specifically employing the coordinated analysis approach. This method maintains each data set separately while analyzing them under a unified protocol (M. Becker et al., 2022; Graham et al., 2022; Hofer & Piccinin, 2009). This approach preserves measurement and sample heterogeneity, offering more flexibility than pooled analysis, which requires high uniformity across combined data sets (Graham et al., 2022). Previous studies on peer effects have predominantly relied on analyses of individual data sets and varied often in terms of included variables and analysis strategies. Through our IDA, we ensured consistency in covariates, uniformly addressed missing values, and applied the same multilevel regression models across all data sets. This uniform approach enabled us to conduct a thorough comparison and evaluate the generalizability and replicability of peer effects. In a subsequent step, we summarized the results meta-analytically. By employing this structured and comprehensive approach, we aimed to clarify existing ambiguities and contribute new evidence to the ongoing discussion in the field.

Multilevel Regression Models

To investigate the impact of peers' language skills on individual language development, a two-level multilevel lagged regression framework was applied, with children

at Level 1 nested within classrooms at Level 2 (Raudenbush & Bryk, 2002). Specifically, we regressed a subsequent language measure on the group-level predictor, which was aggregated from individual lagged language measures, while simultaneously controlling for each child's own earlier language measure. For the three data sets with more than two measurement points we also included analyses of non-adjacent measurements. We applied *z*-standardization, which normalizes data to a mean of zero and a standard deviation of one, thereby inherently including grand-mean centering. Consequently, the regression coefficient of the aggregated mean directly corresponds to the compositional effect.

All analyses were conducted using R, version 4.4.0 (R Core Team, 2024), in RStudio (RStudio Team, 2023) with the packages *lme4* (Bates et al., 2015) for multilevel regression models and *mitml* (Grund et al., 2023) for pooling the estimates across the imputed data sets. We started by testing an unconditional model (intercept-only) to calculate the intraclass correlation (ICC) for each data set, determining the proportion of variance attributable to differences between groups (Model 0). Subsequently, we specified random intercept models by progressively incorporating various covariates to investigate the main peer effect. Model 1 estimated the effect of peers' average language skills on a child's individual language skills, controlling for previous individual language skills without additional covariates. In Model 2, we included all Level 1 control variables, which encompass family background and non-verbal cognitive abilities. Model 3 further incorporated the Level 2 variable of average SES to examine whether the impact of socioeconomic composition of the class potentially explains or reduces the skill-based peer effect. Finally, to address our research questions concerning the moderating effects, we expanded Model 3 to include cross-level interactions (CLI). This involved examining how peer-average skills (Level 2) interact with different individual level predictors (DLL status, individual language skill at Time 1, and individual SES). Each CLI was tested in a separate model.

Missing Data Treatment

The presence of missing data represents a significant challenge in statistical analysis. In the data sets used for our study, the proportion of missing values exhibits considerable variability, not only across different variables but also between the studies (see Table A1). Reasons for missing values can vary and may include attrition, missed measurement occasions, non-return of questionnaires, items left unanswered, and instances of missing by design. The BiSpra II data set, for example, incorporated planned missing values resulting from the test booklet design. Following other studies utilizing this data, we included only children who received a test booklet containing the vocabulary test in our analyses (Volodina et al., 2020). However, for approximately two-thirds of these children, the values for non-verbal cognitive abilities were missing by design. In the NEPS preschool data at Time 2, which was assessed after the transition to primary school, a significant number of language scores were missing. This was primarily due to children transferring to primary schools which did not participate in the study. It is unlikely that these and other missing values were missing completely at random. Consequently, traditional methods such as listwise deletion could lead to biased results (Enders, 2022). Given our robust set of covariates, it is reasonable to assume that the missing data might approximate the missing at random (MAR) condition. Accordingly, the utilization of modern imputation techniques, which leverage information from all correlated variables, can effectively mitigate potential biases and enhance the robustness of the statistical analyses. Multiple imputation (MI) is a highly recommended modern method for effectively handling missing data (Enders, 2022). Nevertheless, studies have demonstrated that results can be biased if the multilevel structure and nonlinear relations (e.g., cross-level interactions) are not adequately taken into account when imputing missing values (Grund et al., 2021). In recent years, significant methodological advances have been made to address this issue (Enders et al., 2020; Lüdtke et al., 2020). To ensure the reliability

and validity of our analyses, we implemented missing data treatment by using the R package *mdmb* (Robitzsch & Lüdtke, 2023), which applies substantive-model-compatible MI based on sequential modeling with a Bayesian estimation technique (Grund et al., 2021). Building on this method and in line with our coordinated analysis approach, we specified an imputation model that was as uniform as possible across all data sets and compatible with our analysis models. This model incorporated the multilevel framework with cross-level interactions and encompassed all relevant variables. Specifically, utilizing the R package *mdmb* enabled us to include the interaction term, aggregated group means during imputation, and to effectively handle the imputation of dichotomous variables. To facilitate convergence, all non-dichotomous independent variables were *z*-standardized prior to imputation and re-standardized across all imputed data sets after imputation. A total of 50 imputed data sets were generated for each original data set. Diagnostic plots and the $\hat{R}$ statistic were employed to assess the convergence of the algorithm, which was found to be adequate with all $\hat{R}$ values below 1.1. All imputed data sets were analyzed using the *testEstimate* function of the R package *mitml* (Grund et al., 2023), which implements Rubin's rules to pool estimates and standard errors.

Meta-Analytic Aggregation

We used meta-analytic techniques to integrate the results (i.e., coefficients from the multilevel analyses) across studies, employing the R package *metafor* (Viechtbauer, 2010). We specified random-effects models with restricted maximum likelihood estimation (REML). To account for dependencies among effect sizes from the same study, we implemented three-level meta-analytic models, nesting effect sizes within their respective studies. Furthermore, we applied robust variance estimation (RVE; clubSandwich, Pustejovsky, 2024) with small-sample correction to adjust standard errors for any remaining

within-study clustering. This approach guards against potential model misspecification and produces appropriately conservative standard errors (Pustejovsky & Tipton, 2022).

Additionally, some between-study dependencies exist due to overlapping participants in the preschool and primary school data sets from BiKS and NEPS, which were treated as separate studies within the hierarchical meta-analytic structure (for details see Supplement B). To address this potential issue, we conducted a sensitivity analysis by re-estimating effect sizes for the school samples after excluding overlapping participants.[4]

Robustness Checks

To ensure the robustness of our findings, we implemented two additional checks. First, we re-ran all MI and analyses excluding classrooms that contained fewer than three children. Second, we recalculated the group average for language skills by excluding the individual scores of the children whose outcomes were being predicted as recommended, for example, in the studies by Mashburn et al. (2009) or Choi et al. (2018). This provided each child with a unique peer score as a Level 1 predictor.

Transparency and Openness

The study utilized secondary data available as scientific-use files with a data use agreement. NEPS data can be obtained from the Research Data Center at the Leibniz Institute for Educational Trajectories (FDZ-LIfBi). All other data used in this study are available from the Research Data Center at the Institute for Educational Quality Improvement (FDZ at IQB) in Berlin. An overview of the data sets included in this study, with corresponding references, is presented in Table 1; full citations appear in the reference section. Data were analyzed using R, version 4.4.0 (R Core Team, 2024), in RStudio (RStudio Team, 2023). The analysis

[4] As our sensitivity analysis revealed that overlapping participants did not substantially alter effect estimates (see Table S1 in Supplement), we retained them in our primary analysis to maximize statistical power while avoiding unnecessary data exclusions.

code is available upon request from the corresponding author. This study and its analysis were not preregistered.

Results

All reported outcomes are based on the pooled results from the imputed data sets. As some studies included multiple measurement time points, we performed 17 separate analyses for each of our multilevel regression models. The ICC values for each data set and measurement time point (Table 3, Model 0), derived from the intercept-only model, indicate that between 7% and 37% of the variance in language skills is due to differences between groups. This range reflects notable variability, with most data sets suggesting that a moderate proportion of the variance is attributable to group-level differences.

Main Effect of Peer Language Abilities

Our first research question addressed the main effect of peers' language abilities on children's language skills. Table 3 presents the standardized regression coefficients and standard errors of peer-average language abilities across the three models with increasing inclusion of control variables for each data set and measurement time point, as well as the meta-analytically aggregated overall effects.

The first model tested the peer effect controlling only for individual prior language skills without additional control variables (Model 1). Across all data sets, the peer effect ranged from $\beta = -.05$ ($p = .560$) to $\beta = .41$ ($p < .001$). The meta-analytic summary revealed a significant positive association ($\beta = .13$, 95% CI [0.03, 0.24], $p = .022$), suggesting that, on average, children in peer groups with higher language abilities exhibited higher individual language skills at T2. However, this model did not control for potential confounding variables.

In the second model, we included Level 1 control variables to account for potential selection bias (Model 2). As expected, this inclusion resulted in a general attenuation of the

peer effects, with estimates ranging from β = - .07 (p = .374) to β = .28 (p = .004). This suggests that the peer effect observed in Model 1 may be partially attributable to individual-level covariates, although most studies showed only slight decreases. As illustrated in Table 3, three of the 17 estimates indicated a statistically significant positive effect of peer language abilities on individual language skills. After including Level 1 control variables, the overall effect was reduced to β =.09 (95% CI [-0.00, 0.17], p = .050).

To disentangle the potential influence of socioeconomic factors from the skill-based peer effect, Model 3 incorporated the average SES as a Level 2 control variable. The inclusion of SES led to varied adjustments in peer effect estimates (range: β = -.14, p = .141 to β = .35, p < .001). The changes in peer effects compared to Model 2 differed in magnitude and direction across the data sets (see Figure 1), with only two of the 17 estimates indicating a significant peer effect. The meta-analytically aggregated effect further decreased and did not attain statistical significance (β = .07, 95% CI [-0.04, 0.19], p = .180). Regarding the SES effect itself in this model, three out of 17 estimates were found to be significant, whereas the peer language skills were not. However, the aggregated SES effect remained small and statistically non-significant (β = .04, 95% CI [-0.05, 0.13], p = .315).

In summary, referring to our full model, the meta-analytic synthesis indicated that, on average, there was no significant main effect of peer language abilities on individual language skills when controlling for important variables. Nevertheless, the effects exhibited considerable variability across the different data sets.

Moderation of Peer Effects by Individual Characteristics

Our second research question examined whether peer effects vary depending on children's DLL status, prior language skills, and family SES. To investigate this, we added cross-level interaction (CLI) terms to Model 3 and estimated the moderation effects in separate models each. The results are also presented in Table 3.

The analyses for DLL status yielded consistently positive effects across all estimates (range β = .61, *p* = .003 to β = .01, *p* = .960), with three reaching statistical significance. The meta-analytic aggregation showed a significant moderation effect of β =.12 (95% CI [0.01, 0.23], *p* = .037). To probe this significant interaction between peer language skills and children's DLL status, we conducted simple slope analyses (Bauer & Curran, 2005; see Table S3 in the Supplemental Materials) and provided a graphical illustration of the effect. As depicted in Figure 2, the steeper slope for DLLs indicates that DLLs benefit more from being in higher-skilled peer groups compared to SLLs. Examination of the simple slopes revealed that the aggregated peer effect was statistically significant for DLLs (β = .16, 95% CI [0.02, 0.29], *p* = .027), while for SLLs, the effect did not significantly differ from zero (β = .03, 95% CI [-0.08, 0.14], *p* = .560).

The interaction between peer language abilities and children's prior language skills demonstrated significant negative estimates in seven out of 17 analyses. Across all analyses, the estimates ranged from β = -.23, *p* = .003 to β = .08, *p* = .185. There was a significant pooled moderation effect of β = -.06 (95% CI [-0.12, -0.00], *p* = .039). Figure 3 illustrates how the peer effect varies for children with low (-1 *SD*), average (*M*), and high (+1 *SD*) prior language skills. This figure shows that the peer effect is particularly pronounced for children with low prior skills; they benefit from being in groups with more skilled peers. Conversely, children with high prior skills were not impacted by the language abilities of their peers. The simple slope analysis corroborated these findings, indicating a significant aggregated peer effect only for children with low prior skills (β = .13, 95% CI [0.00, 0.25], *p* = .046; see Table S3 in the Supplemental Materials). Overall, the two significant moderation effects support the assumption of a compensatory effect.

Further analyses explored the interaction between peer language abilities and family SES (β = -.33, *p* < .001 to β = .13, *p* = .224). These revealed significant negative effects in

seven out of 17 analyses. The meta-analytically aggregated effect was not significant (β =-.06, 95% CI [-0.13, 0.01], p = .085), but the direction and magnitude of the effect were similar to that observed for prior language skills.

Robustness Checks

To ensure the robustness of our findings, we conducted all analyses with two modifications. Firstly, to assess the potential for bias due to averages drawn from small clusters, all groups comprising fewer than three children were excluded. This criterion did not affect all data sets; some had no groups smaller than three, while others had over 50% of such small groups. Most coefficients and standard errors exhibited marginal increases, however there was some variation across studies. The meta-analytically aggregated results indicated that the overall effects were slightly larger, but the trends remained consistent (see Table A2). The interaction with DLL status showed mostly minor changes compared to the full sample, except for two analyses that exhibited larger, opposite deviations in the estimates (NEPS preschool and BiKS school 1-3). The meta-analytic effect slightly decreased and no longer reached statistical significance. Results for models including interactions with prior language skills and family SES were generally consistent with the main analyses, replicating the significant overall effect for the CLI with prior skills (see Table A2).

Secondly, we recalculated the average peer skills excluding the target child (based on the data without small groups). In the main effect models (Models 1 – 3), this adjustment resulted in marginal changes in the estimates across data sets compared to the main analyses (see Table A3), with identical overall effects. In our analyses of moderation effects, we again found notable variations in the interaction of peer language skills with DLL status. The overall effect decreased and did not retain statistical significance. However, the significant interaction between peer language abilities and prior skills remained robust.

Exploratory Moderator Analysis of Age Group

To explore whether peer effects differ between preschool and primary school samples, we included study-level age category (preschool vs. primary school) as a moderator in our meta-analytic models. Specifically, this variable was included as a predictor in the meta-analytic models of the overall peer effect (Model 3) as well as in the models of the cross-level interaction effects. Results showed no significant effects across models. The age category predictor (preschool serves as the reference category) was not significant regarding peer effects of Model 3 (Est. = -0.04, SE = 0.10, p = .693) as well as for the effects of the cross-level interactions (DLL status: Est. = -0.08, SE = 0.08, p = .36; prior language skills: Est. = 0.02, SE = 0.05, p = .76; SES: Est. = 0.09, SE = 0.05, p = .09). These results suggest that peer effects in our models do not systematically differ between preschool and primary school samples.

Discussion

This study systematically investigated how peer language abilities affect individual development in preschool and primary school children using an IDA approach, addressing inconsistencies in existing research. We analyzed data from six German longitudinal studies, specifying separate but unified multilevel regression models to estimate peer effects and moderating effects. The estimates were then summarized meta-analytically, providing a more precise and reliable understanding of these effects that goes beyond the scope of individual studies. This meta-analytic aggregation yielded two primary findings. First, we found no evidence for a main effect of peer language abilities on individual language development. Second, a robust interaction effect with children's prior language skills was observed, supporting the compensatory effect hypothesis.

Main Peer Effects on Language Development

Although our primary focus was on the full model controlling for individual-level variables and classroom peer SES (Model 3), the initial analysis of Model 1 revealed a

significant pooled peer effect, which diminished once further controls like cognitive abilities and SES, were included in Model 2. This decrease underscores the importance of accounting for individual factors to avoid overestimating peer effects due to confounds that may affect both language abilities and selection into educational settings. The meta-analytic summary of the 17 estimates from Model 2 yielded a small effect with a p-value of .050. While this result is at the conventional threshold of statistical significance, it warrants cautious interpretation. Robustness checks yielded similar effect sizes but slightly larger p-values, underscoring the uncertainty of this finding. These results indicate a lack of strong evidence for a conclusive influence of peers' language abilities on individual language development. This conclusion was further reinforced when considering classroom peer SES as an additional control in Model 3. In this final model, the aggregated peer effect decreased even more in magnitude and was clearly non-significant. Adjusting for group SES allowed us to differentiate between effects attributable to peers' actual language abilities and confounded influences conditioned by the broader socioeconomic context. This is particularly important as SES composition is also intertwined with a wide array of factors that can influence the learning environment, including access to resources, quality of instruction, and parental engagement (Perry, 2012; Willms, 2010). The lack of a robust main peer effect suggests that the direct influence of peers' language abilities may be relatively small overall.

In summary, the meta-analytic synthesis did not support the assumption that being surrounded by peers with higher language skills influences individual language development for all children equally when controlling selection bias and peer group SES. This result is consistent with the findings of previous studies which also reported no significant peer effects when controlling for aspects of SES composition alongside other variables (Choi et al. 2018; Kohl et al.; 2022; Yang et al., 2023).

Overall, there was considerable variability in the effects across our individual studies, which reflects the broader inconsistencies in literature. Only a minority of our analyses indicated significant peer effects, suggesting that the presence and strength of peer effects might potentially be contingent on specific study or sample characteristics. While some data sets (e.g., PHONO or BiSpra) showed substantial positive effects even when controlling for SES composition, others revealed small or non-significant effects. Such variability may explain the discrepancies in previous research, where some studies report significant peer effects (e.g., Atkins-Burnett et al., 2017; Foster et al. 2020; Mashburn et al., 2009) while others do not (e.g., Kohl et al., 2022; Yang et al., 2023). Furthermore, we found that the pattern of results was largely robust to the exclusion of small groups from the relevant data sets and when calculating peer averages both including and excluding the target child. This suggests that the observed variability is more likely due to differences in study contexts or sample characteristics rather than methodological inconsistencies.

Moderation of Peer Effects by Individual Characteristics

We examined potential moderating effects of DLL status, prior language skills, and family SES on the peer effect. Our analysis revealed a complex pattern of results. Initial analyses suggested an interaction effect between peer language skills and DLL status, indicating that DLLs might benefit on average more from their peers' language abilities compared to SLLs. This initial result corroborates the finding from Schmerse (2021), who used the BiKS preschool data and a growth curve modeling approach to examine peer effects on DLLs' language development. Our regression analysis of the BiKS data set replicated a significant interaction when examining non-adjacent time points. However, the effect was not consistent across all analyses and data sets. In the SEIKA 1-2 data, we found no significant interaction with DLL status, which aligns with the null findings reported by Kohl et al. (2022) using the same data. More importantly, our meta-analytically aggregated effect across

all data sets lacked robustness in subsequent checks when we excluded small groups and adjusted the mean of peer skills. The effect decreased in magnitude and became non-significant, consistent with other studies (Lin et al., 2023; Yang et al., 2023). The reduction in the DLL effect after our robustness checks suggests that it might be highly sensitive to specific analytical strategies and sample characteristics. Moreover, our dichotomization of learner status into DLL vs. SLL may mask the inherent heterogeneity within the group of DLLs, such as variations in language use at home and/or proficiency levels (Kim et al., 2018). This points to the idea that the level of individual majority language skills themselves, rather than DLL status per se, may be a more crucial factor in understanding peer effects on language development.

Supporting this notion, our analyses identified a robust and significant interaction effect with prior language skills. The aggregated effects of the simple slope analyses revealed that children with lower initial abilities benefit more from exposure to peers with higher language skills. This finding is particularly noteworthy and diverges from most previous studies in preschool contexts, which have reported null findings for this interaction (e.g., Choi et al., 2018; Kohl et al., 2022; Ribeiro et al., 2017; Yang et al., 2023). However, there are two exceptions that report significant effects, but in opposite directions. Mashburn et al. (2009) observed positive effects for children with higher initial language skills, suggesting a *Matthew effect*. Conversely, Justice et al. (2011) reported pronounced effects for children with lower prior language skills, specifically focusing on low SES backgrounds. Our meta-analytical aggregation provides robust support for this compensatory effect, extending its applicability beyond the low SES context.

The pooled interaction effect with individual SES backgrounds was not significant. However, several of our individual analyses showed significant interaction effects, and the magnitude of the pooled effect was similar to that of prior skills. Previous research has

reported mixed results using different SES measures. For instance, Yang et al. (2023) found a moderating effect for income, Ribeiro et al. (2017) for maternal education, while Kohl et al. (2022) reported null findings for both maternal education and household income.

In conclusion, our analysis demonstrates that prior language abilities represent the most reliable moderator of peer effects on language development. Although DLL status and SES background demonstrated significant effects in some individual studies, the meta-analytic aggregation did not reveal a robust effect. The non-significant findings for family SES and the unstable findings regarding DLL status as moderators suggest that these factors may not directly determine susceptibility to peer influences. Instead, they may operate indirectly through their association with a child's language skills. These findings underscore the complexity of peer effects in language development and highlight the need for caution when generalizing results from individual studies.

Why peers may matter

Peers may influence language development through various direct and indirect mechanisms (Purtell et al., 2021). Our finding that children with lower prior language skills benefit more in peer groups with higher language levels aligns with social learning theories and perspectives that emphasize the crucial role of interactions in language development (Bandura, 1971; Vygotsky, 1978). Higher average peer language levels may facilitate individual language skills directly by providing ample opportunities for diverse and complex language interactions through more skilled peers serving as language models (Bandura, 1971; Rydland et al., 2014). Moreover, the shared cultural references and age-appropriate topics that emerge in peer interactions can provide contextually relevant vocabulary and language structures, making the learning process more meaningful and memorable (Blum-Kulka & Snow, 2004; Cekaite et al., 2014). Unlike adult-child interactions, peer relationships are typically more symmetrical, potentially encouraging more active participation and dynamic

exchanges in language use. This symmetry particularly fosters the negotiation of different viewpoints, especially during discussions about rules and in role-play – patterns less common in the more asymmetrical dialogues with adults (Mammen et al., 2019). These factors combined create a rich, peer-driven language learning environment that complements and potentially enhances the more structured language instruction provided by adults (Blum-Kulka & Snow, 2004).

Additionally, it has been proposed that the average language level in a classroom may indirectly impact children's development, for example by prompting teachers to adjust their instruction, curriculum, and language use to better align with the overall classroom level (Harris, 2010; Pakarinen et al., 2011; Wilkinson & Fung, 2002). A generally supportive classroom climate and a high-quality environment may also foster enriching interactions among children (Mashburn et al., 2009). Nevertheless, our study relied on aggregated measures of peer language skills and could not capture the nuances of individual interactions or the indirect pathways through which these effects may occur. Some studies tested the mediating role of classroom quality in the preschool context but reported null findings (Kohl et al., 2022; Yeomans-Maldonado et al., 2019). Additionally, there is a growing body of research exploring the role of direct interactions and language input from peers in preschool classrooms (Chaparro-Moreno et al., 2019; Chen et al., 2020; DeLay, Hanish, et al., 2016; Gámez et al., 2019). These studies contribute to a better understanding of the significant role that peers play in language development, going beyond compositional features.

Limitations

A key strength of our study lies in its IDA approach with meta-analytically summarizing 17 distinct findings from individual studies. Notably, as Graham et al. (2020) highlighted in their IDA, each of these analyses could potentially have been treated as a study on its own, possibly contributing to the inconclusiveness in the field.

The IDA approach, combined with our robustness checks and the inclusion of two data sets previously used for peer effect analysis provided a unique opportunity for systematic comparison across multiple data sets. This enabled us to assess the robustness and generalizability of peer effects in ways not possible with single-study designs. By including the two data sets previously used, we demonstrated that prior findings are robust to differences in modeling approach, covariates, and handling of missing values. However, our broader analysis across multiple studies revealed that these results are not readily replicable and not generalizable across different kinds of samples without careful consideration.

Despite the strengths of our study, several limitations warrant consideration. First, every IDA approach faces trade-offs between complexity and comparability. Following recommendations by Weston et al. (2020), we addressed simpler questions and models across our multiple data sets rather than more complex analyses typically possible in targeted individual studies. Although this strategy permits broader comparisons, it may also restrict the depth of analysis for each data set (Graham et al., 2022; Weston et al., 2020). The use of shared variables across all data sets constrains both the selection of studies and the ability to analyze additional moderators or mediators. For example, measures of process quality frequently included in peer effect analyses in preschool contexts (e.g., Mashburn et al., 2009; Kohl et al., 2022; Schmerse, 2021; Yeomans-Maldonado et al., 2019) were unavailable in most of our data sets. These limitations illustrate the difficulty of achieving an appropriate balance between breadth and depth in multi-study analyses.

Second, due to the limited number of studies we were unable to conduct a thorough investigation of study-level factors that might explain variations in peer effects. Although we included data sets with both preschool and primary school children, our exploratory analysis testing age category (preschool vs. primary school) as a potential moderator revealed no significant effects. Theoretically, the age of peers could influence language development, as

older peers possess, on average, more advanced language skills, potentially enriching the language environment. Moreover, studies have suggested that older children tend to engage more positively with their peers (Vitiello et al., 2012), which could facilitate more frequent and higher-quality language interactions. Future research should further investigate potential developmental differences and other sources of variability in peer effects.

Third, our study's language measurements present certain limitations. We relied primarily on receptive vocabulary measures to assess both peer language skills and children's language outcomes. Receptive vocabulary is widely used in large-scale studies across diverse developmental stages, making it well-suited for multi-study designs such as IDA. Although receptive vocabulary alone does not capture the full breadth of language abilities, it correlates substantially with other linguistic domains in early childhood (Brinchmann et al., 2019; Dickinson et al., 2003; Lonigan et al., 2009). Nevertheless, peer interactions may influence language development in ways not fully captured by vocabulary tests. For example, peer discourse involving collaborative reasoning, joint decision-making, or perspective-taking may be particularly relevant for the acquisition of more complex constructions (Köymen & Tomasello, 2020).

Furthermore, one data set (KuL) used phonological awareness (PA) instead of vocabulary as a proxy for language competence. Although PA is associated with language development, it differs from the vocabulary measures used in the other data sets. To assess the impact of this difference, we conducted a sensitivity analysis by excluding the KuL study from our meta-analytic aggregation. This led to minor variations in effect sizes, with a slight increase in the estimated peer effects in Models 1 to 3, as well as in the cross-level interaction effect for DLL status (see Supplemental Materials, Table S4). The effect size for the cross-level interaction with prior language skill was marginally reduced, while the effect size for

the cross-level interaction with SES remained unchanged. Overall, the pattern of findings was largely consistent, though with reduced statistical power due to the smaller sample size.

Taken together, future research should examine the role of peer effects in shaping language development beyond receptive vocabulary.

Fourth, as mentioned above, our operationalization of DLL status as a dichotomous variable may have oversimplified a complex linguistic reality (Kim et al., 2018; Lin et al., 2023). For example, Lin et al. (2023) found that fluent and emergent bilinguals in classrooms contributed differently to language outcomes. Future studies could benefit from more detailed measures that allow for a more fine-grained analysis of how varying degrees of dual language experience might moderate peer effects on language development.

Fifth, our study relied on the average language abilities of peers, which presents certain limitations. As previously noted, this compositional measure could not capture the underlying mechanisms through which peer effects may operate, nor does it necessarily represent the true skill level of the peers with whom a child regularly interacts. Continued research is needed to explore the specific processes underlying compositional effects and to understand the role that direct interactions and peer networks may play in language development (Chen et al., 2020; DeLay, Hanish, et al., 2016). Additionally, focusing on average skill neglects other important aspects of peer group composition such as heterogeneity (Rjosk, 2020). Nevertheless, studies which examine heterogeneity of peer language skills on language development reported mostly null findings (Atkins-Burnett et al., 2017; Garcia, 2021; Yang et al., 2023).

Lastly, although our study employed longitudinal data with later measurements as outcomes, it remains correlational. Therefore, we are cautious about drawing causal conclusions from our findings. However, we incorporated a comprehensive set of control variables, including demographic factors, cognitive prerequisites, and SES composition, to

address potential selection bias. By including non-verbal cognitive abilities, we accounted for language-independent aspects of cognitive development, extending beyond typical controls in previous studies in preschool contexts. Furthermore, by incorporating SES as a group-level control, we aimed to capture broader contextual influences and more effectively disentangle the effects attributable to peers' actual language abilities from those related to socioeconomic factors. Nevertheless, the possibility of unmeasured confounders remains a central consideration. Further research could utilize alternative methodologies and models that may be more effective in elucidating causal scenarios.

Conclusion

Our study aimed to contribute to the understanding of the role of classroom peers in language development during early childhood education. The findings reveal that the effect of peer language abilities on language development is particularly pronounced for children with lower language skills. These children develop better language skills when surrounded by more competent peers than in groups with lower average language levels. This insight, supported by robust IDA methodology, has important implications for educational practice and policy. It suggests that thoughtful classroom composition and the promotion of peer interactions could enhance language development outcomes. This could involve creating environments and organizing activities where children with lower language skills are grouped with more competent peers and where language interactions and collaborative learning are encouraged. Importantly, our study found no evidence that children with average or strong language skills are negatively affected when placed in groups with lower average language levels, which might also be valuable information for parents.

While the observed effect sizes are modest, their potential to accumulate over time (Götz et al., 2022) underscores the importance of early educational environments in shaping children's language trajectories. Ultimately, this research highlights that in the realm of early

language development, alongside well-established factors like teacher-child interactions and other quality aspects, it is crucial to consider the role of peers. Our findings suggest that peer influences, if harnessed effectively, may help reduce disparities in language skills and potentially enhance opportunities for future academic success.

References

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, *14*(5), 255–259. https://doi.org/10.1111/j.0963-7214.2005.00376.x

Atkins-Burnett, S., Xue, Y., & Aikens, N. (2017). Peer effects on children's expressive vocabulary development using conceptual scoring in linguistically diverse preschools. *Early Education and Development*, *28*(7), 901–920. https://doi.org/10.1080/10409289.2017.1295585

Autor:innengruppe Bildungsberichterstattung (Ed.). (2024). *Bildung in Deutschland 2024: Ein indikatorengestützter Bericht mit einer Analyse zu beruflicher Bildung [Education in Germany 2024: An indicator-based report with an analysis of vocational education and training]*. wbv Media. https://doi.org/10.3278/6001820iw

Bandura, A. (1971). *Social learning theory*. General Learning Press.

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

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*, *40*(3), 373–400. https://doi.org/10.1207/s15327906mbr4003_5

Bäuerlein, K., Beinicke, A., Berger, N., Faust, G., Jost, M., & Schneider, W. (2012). *FIPS - Fähigkeitsindikatoren Primarschule [Performance Indicators in Primary Schools]*. Hogrefe.

Becker, B., & Schober, P. S. (2017). Not just any child care center? Social and ethnic disparities in the use of early education institutions with a beneficial learning

environment. *Early Education & Development*, *28*(8), 1011–1034. https://doi.org/10.1080/10409289.2017.1320900

Becker, M., Kocaj, A., Jansen, M., Dumont, H., & Lüdtke, O. (2022). Class-average achievement and individual achievement development: Testing achievement composition and peer spillover effects using five German longitudinal studies. *Journal of Educational Psychology*, *114*(1), 177–197. https://doi.org/10.1037/edu0000519

Bellin, N. (2009). *Klassenkomposition, Migrationshintergrund und Leistung: Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschülern [Classroom composition, migration background and achievement: Multilevel analyses of primary school students' language and reading comprehension]*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91359-9

Bihler, L.-M., Agache, A., Schneller, K., Willard, J. A., & Leyendecker, B. (2018). Expressive morphological skills of dual language learning and monolingual German children: Exploring links to duration of preschool attendance, classroom quality, and classroom composition. *Frontiers in Psychology*, *9*, Article 888. https://doi.org/10.3389/fpsyg.2018.00888

Bihler, L.-M., Agache, A., Willard, J., & Leyendecker, B. (2021). *Sprachbildung und -entwicklung im Kita-Alltag in Nordrhein-Westfalen—Längsschnitt (SEIKA-NRW LS) [Language education and development in kindergarten in North Rhine-Westphalia—Longitudinal (SEIKA-NRW LS)]* (Version 1) [Dataset]. Berlin: IQB - Institute for Educational Quality Improvement. https://doi.org/10.5159/IQB_SEIKA-NRW_LS_V1

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*(6), 1461–1476. https://doi.org/10.1017/S0142716416000060

Blossfeld, H.-P., & Roßbach, H.-G. (Eds.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS).* Springer VS. https://doi.org/10.1007/978-3-658-23162-0

Blum-Kulka, S., & Snow, C. E. (2004). Introduction: The potential of peer talk. *Discourse Studies*, *6*(3), 291–306. https://doi.org/10.1177/1461445604044290

Brinchmann, E. I., Braeken, J., & Lyster, S.-A. H. (2019). Is there a direct relation between the development of vocabulary and grammar? *Developmental Science*, *22*(1), e12709. https://doi.org/10.1111/desc.12709

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (6th ed., pp. 793–828). Wiley.

Cabell, S. Q., Justice, L. M., McGinty, A. S., DeCoster, J., & Forston, L. D. (2015). Teacher–child conversations in preschool classrooms: Contributions to children's vocabulary development. *Early Childhood Research Quarterly*, *30*, 80–92. https://doi.org/10.1016/j.ecresq.2014.09.004

Cekaite, A., Blum-Kulka, S., Grøver, V., & Teubal, E. (2014). Children's peer talk and learning: Uniting discursive, social, and cultural facets of peer interactions: Editors' introduction. In A. Cekaite, S. Blum-Kulka, V. Grøver, & E. Teubal (Eds.), *Children's peer talk. Learning from each other* (1st ed., pp. 3–20). Cambridge University Press. https://doi.org/10.1017/CBO9781139084536.003

Chaparro-Moreno, L. J., Justice, L. M., Logan, J. A. R., Purtell, K. M., & Lin, T.-J. (2019). The preschool classroom linguistic environment: Children's first-person experiences. *PloS One*, *14*(8), 0220227. https://doi.org/10.1371/journal.pone.0220227

Chen, J., Justice, L. M., Tambyraja, S. R., & Sawyer, B. (2020). Exploring the mechanism through which peer effects operate in preschool classrooms to influence language growth. *Early Childhood Research Quarterly*, *53*, 1–10. https://doi.org/10.1016/j.ecresq.2020.02.002

Choi, J. Y., Castle, S., Burchinal, M., Horm, D., Guss, S., & Bingham, G. E. (2018). Peer effects on low-income children's learning and development. *Journal of School Psychology*, *71*, 1–17. https://doi.org/10.1016/j.jsp.2018.10.002

Coley, R. L., Spielvogel, B., & Kull, M. (2019). Concentrated poverty in preschools and children's cognitive skills: The mediational role of peers and teachers. *Journal of School Psychology*, *76*, 1–16. https://doi.org/10.1016/j.jsp.2019.05.008

De Fraine, B., Van Damme, J., Van Landeghem, G., Opdenakker, M., & Onghena, P. (2003). The effect of schools and classes on language achievement. *British Educational Research Journal*, *29*(6), 841–859. https://doi.org/10.1080/0141192032000137330

DeLay, D., Hanish, L. D., Martin, C. L., & Fabes, R. A. (2016). Peer effects on Head Start children's preschool competency. *Developmental Psychology*, *52*(1), 58–70. https://doi.org/10.1037/dev0000066

DeLay, D., Zhang, L., Hanish, L. D., Miller, C. F., Fabes, R. A., Martin, C. L., Kochel, K. P., & Updegraff, K. A. (2016). Peer influence on academic performance: A social network analysis of social-emotional intervention effects. *Prevention Science*, *17*(8), 903–913. https://doi.org/10.1007/s11121-016-0678-8

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships

among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, *95*(3), 465–481. https://doi.org/10.1037/0022-0663.95.3.465

Dumont, H., & Ready, D. D. (2020). Do schools reduce or exacerbate inequality? How the associations between student achievement and achievement growth influence our understanding of the role of schooling. *American Educational Research Journal*, *57*(2), 728–774. https://doi.org/10.3102/0002831219868182

Enders, C. K. (2022). *Applied missing data analysis* (2nd ed.). The Guilford Press.

Enders, C. K., Du, H., & Keller, B. T. (2020). A model-based imputation procedure for multilevel regression models with random coefficients, interaction effects, and nonlinear terms. *Psychological Methods*, *25*(1), 88–112. https://doi.org/10.1037/met0000228

European Child Care and Education (ECCE) Study Group. (1999). *School-age assessment of child development: Long-term impact of pre-school experiences on school success, and familiy-school relationships* [Final report]. European Union.

Fischer, L., & Durda, T. (2020). *NEPS Technical report for receptive vocabulary: Scaling results of starting cohort 2 for kindergarten (Wave 1), grade 1 (Wave 3) and grade 3 (Wave 5)* (NEPS Survey Paper No. 65). Leibniz Institute for Educational Trajectories (LIfBi). http://doi.org/10.5157/NEPS:SP65:1.0

Foster, T. J., Burchinal, M., & Yazejian, N. (2020). The relation between classroom age composition and children's language and behavioral outcomes: Examining peer effects. *Child Development*, *91*(6), 2103–2122. https://doi.org/10.1111/cdev.13410

Gámez, P. B., Griskell, H. L., Sobrevilla, Y. N., & Vazquez, M. (2019). Dual language and English-Only learners' expressive and receptive language skills and exposure to

peers' language. *Child Development*, *90*(2), 471–479. https://doi.org/10.1111/cdev.13197

Ganzeboom, H., Graaf, P., & Treiman, D. (1992). A standard socio-economic index of occupational status. *Social Science Research*, *21*, 1–56. https://doi.org/10.1016/0049-089X(92)90017-B

Garcia, E. B. (2018). The classroom language context and English and Spanish vocabulary development among dual language learners attending Head Start. *Early Childhood Research Quarterly*, *42*, 148–157. https://doi.org/10.1016/j.ecresq.2017.09.005

Garcia, E. B. (2021). Peer effects on dual language learners' English and Spanish receptive vocabulary development. *Early Childhood Research Quarterly*, *56*, 180–189. https://doi.org/10.1016/j.ecresq.2021.03.015

Gentrup, S., & Rjosk, C. (2018). Pygmalion and the gender gap: Do teacher expectations contribute to differences in achievement between boys and girls at the beginning of schooling? *Educational Research and Evaluation*, *24*(3–5), 295–323. https://doi.org/10.1080/13803611.2018.1550840

Gibbs, S. (2004). Phonological awareness: An investigation into the developmental role of vocabulary and short-term memory. *Educational Psychology*, *24*(1), 13–25. https://doi.org/10.1080/0144341032000146412

Götz, F. M., Gosling, S. D., & Rentfrow, P. J. (2022). Small effects: The indispensable foundation for a cumulative psychological science. *Perspectives on Psychological Science*, *17*(1), 205–215. https://doi.org/10.1177/1745691620984483

Graham, E. K., Weston, S. J., Gerstorf, D., Yoneda, T. B., Booth, T., Beam, C. R., Petkus, A. J., Drewelies, J., Hall, A. N., Bastarache, E. D., Estabrook, R., Katz, M. J., Turiano, N. A., Lindenberger, U., Smith, J., Wagner, G. G., Pedersen, N. L., Allemand, M., Spiro, A., … Mroczek, D. K. (2020). Trajectories of Big Five personality traits: A

coordinated analysis of 16 longitudinal samples. *European Journal of Personality*, *34*(3), 301–321. https://doi.org/10.1002/per.2259

Graham, E. K., Willroth, E. C., Weston, S. J., Muniz-Terrera, G., Clouston, S. A. P., Hofer, S. M., Mroczek, D. K., & Piccinin, A. M. (2022). Coordinated data analysis: Knowledge accumulation in lifespan developmental psychology. *Psychology and Aging*, *37*(1), 125–135. https://doi.org/10.1037/pag0000612

Grund, S., Lüdtke, O., & Robitzsch, A. (2021). Multiple imputation of missing data in multilevel models with the R package mdmb: A flexible sequential modeling approach. *Behavior Research Methods*, *53*(6), 2631–2649. https://doi.org/10.3758/s13428-020-01530-0

Grund, S., Robitzsch, A., & Lüdtke, O. (2023). *Mitml: Tools for multiple imputation in multilevel modeling. R package version 0.4-5* [Computer software]. http://doi.org/10.32614/CRAN.package.mitml

Gutiérrez, G. (2023). Is it socioeconomic or academic? Disentangling sources of peer effects on student achievement. *British Journal of Sociology of Education*, *44*(1), 144–163. https://doi.org/10.1080/01425692.2022.2137465

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, *18*(5), 527–544. https://doi.org/10.1002/jae.741

Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, *112*(4), 961–1197. https://doi.org/10.1177/016146811011200404

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Paul H. Brookes Publishing Co.

Heller, K., & Geisler, H.-J. (1983). *Kognitiver Fähigkeitstest (Grundform), KFT 1-3 [Cognitive ability test, basic form]*. Beltz.

Henry, G. T., & Rickman, D. K. (2007). Do peers influence children's skill development in preschool? *Economics of Education Review*, *26*(1), 100–112. https://doi.org/10.1016/j.econedurev.2005.09.006

Heppt, B., Köhne-Fuetterer, J., Eglinsky, J., Volodina, A., Stanat, P., & Weinert, S. (2020). *BiSpra 2-4: Test zur Erfassung bildungssprachlicher Kompetenzen bei Grundschulkindern der Jahrgangsstufen 2 bis 4 [Test for assessing academic language competencies in primary school children in grades 2 to 4]*. Waxmann.

Heppt, B., Theisen, E., Weinert, Sabine, & Stanat, P. (2019). *Bildungssprachliche Kompetenzen (BiSpra II): Anforderungen, Sprachverarbeitung und Diagnostik. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente der Validierungsstudie [Academic language competencies – affordances, language processing, and diagnostics. Manual for the documentation of the instruments used in the validation study]*. Otto-Friedrich-Universität Bamberg.

Hipfner-Boucher, K., Milburn, T., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2014). Relationships between preschoolers' oral language and phonological awareness. *First Language*, *34*(2), 178–197. https://doi.org/10.1177/0142723714525945

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K. S., & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*(7), 1071–1083. https://doi.org/10.1177/0956797615581493

Hofer, S. M., & Piccinin, A. M. (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychological Methods*, *14*(2), 150–164. https://doi.org/10.1037/a0015566

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4–14. https://doi.org/10.1037/a0027238

Hogrebe, N., Pomykaj, A., & Schulder, S. (2021). Segregation in early childhood education and care in Germany: Insights on regional distribution patterns using national educational studies. *Diskurs Kindheits- Und Jugendforschung / Discourse. Journal of Childhood and Adolescence Research*, *1*, 36–56. https://doi.org/10.3224/diskurs.v16i1.04

Justice, L. M., Jiang, H., & Strasser, K. (2018). Linguistic environment of preschool classrooms: What dimensions support children's language growth? *Early Childhood Research Quarterly*, *42*, 79–92. https://doi.org/10.1016/j.ecresq.2017.09.003

Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer effects in preschool classrooms: Is children's language growth associated with their classmates' skills? *Child Development*, *82*(6), 1768–1777. https://doi.org/10.1111/j.1467-8624.2011.01665.x

Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., Ruzek, E. A., & Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development*, *84*(4), 1171–1190. https://doi.org/10.1111/cdev.12048

Kim, D.-H., Lambert, R. G., & Burts, D. C. (2018). Are young dual language learners homogeneous? Identifying subgroups using latent class analysis. *The Journal of Educational Research*, *111*(1), 43–57. https://doi.org/10.1080/00220671.2016.1190912

Kohl, K., Bihler, L.-M., Agache, A., Leyendecker, B., & Willard, J. A. (2022). Do peers matter? Peer effects on young children's vocabulary gains in German classrooms. *Journal of Educational Psychology*, *114*(1), 161–176. https://doi.org/10.1037/edu0000522

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), 1–16. https://doi.org/10.1177/2332858419832513

Köymen, B., & Tomasello, M. (2020). The early ontogeny of reason giving. *Child Development Perspectives*, *14*(4), 215–220. https://doi.org/10.1111/cdep.12384

Kristen, C., Stanat, P., Kogan, I., Lorenz, G., Gentrup, S., & Rahmann, S. (2018). *Kompetenzerwerb und Lernvoraussetzungen (KuL) [Competence Acquisition and Learning Preconditions (KuL)]* (Version 1) [Dataset]. Berlin: IQB - Institute for Educational Quality Improvement. https://doi.org/10.5159/IQB_KUL_V1

Kuzmina, Y., & Ivanova, A. (2018). The effects of academic class composition on academic progress in elementary school for students with different levels of initial academic abilities. *Learning and Individual Differences*, *64*, 43–53. https://doi.org/10.1016/j.lindif.2018.04.004

Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test Revision 4 (PPVT-4), deutsche Version*. Pearson Assessment.

Lin, T.-J., Chen, J., & Cheung, C. (2024). The roles of peers in students' cognitive and academic development. In P. A. Schutz & K. R. Muis, *Handbook of Educational Psychology* (4th ed., pp. 359–382). Routledge. https://doi.org/10.4324/9780429433726-19

Lin, T.-J., Chen, J., Lu, M., Sun, J., Purtell, K., Ansari, A., & Justice, L. (2023). The influence of classroom language contexts on dual language learners' language development. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000804

Lonigan, C. J., Anthony, J. L., Phillips, B. M., Purpura, D. J., Wilson, S. B., & McQueen, J. D. (2009). The nature of preschool phonological processing abilities and their relations to vocabulary, general cognitive abilities, and print knowledge. *Journal of Educational Psychology*, *101*(2), 345–358. https://doi.org/10.1037/a0013837

Lüdtke, O., Robitzsch, A., & West, S. G. (2020). Analysis of interactions and nonlinear effects with missing data: A factored regression modeling approach using maximum likelihood estimation. *Multivariate Behavioral Research*, *55*(3), 361–381. https://doi.org/10.1080/00273171.2019.1640104

Mammen, M., Köymen, B., & Tomasello, M. (2019). Children's reasoning with peers and parents about moral dilemmas. *Developmental Psychology*, *55*(11), 2324–2335. https://doi.org/10.1037/dev0000807

Mashburn, A., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language achievement during pre-kindergarten. *Child Development*, *80*(3), 686–702. https://doi.org/10.1111/j.1467-8624.2009.01291.x

Miller, P., Votruba-Drzal, E., McQuiggan, M., & Shaw, A. (2017). Pre-K classroom-economic composition and children's early academic development. *Journal of Educational Psychology*, *109*(2), 149–165. https://doi.org/10.1037/edu0000137

NEPS-Netzwerk. (2022). *Nationales Bildungspanel, Scientific Use File der Startkohorte Kindergarten (10.0.0) [National Educational Panel Study, Scientific Use File of Starting Cohort Kindergarten]* [Dataset]. Leibniz Institute for Educational Trajectories (LIfBi). https://doi.org/10.5157/NEPS:SC2:10.0.0

Niklas, F., Schmiedeler, S., Pröstler, N., & Schneider, W. (2011). Die Bedeutung des Migrationshintergrunds, des Kindergartenbesuchs sowie der Zusammensetzung der Kindergartengruppe für sprachliche Leistungen von Vorschulkindern [The importance of migration background, kindergarten attendance, and the composition of the kindergarten group for language skills of preschool children]. *Zeitschrift Für Pädagogische Psychologie*, *25*(2), 115–130. https://doi.org/10.1024/1010-0652/a000032

Niklas, F., & Tayler, C. (2018). Room quality and composition matters: Children's verbal and numeracy abilities in Australian early childhood settings. *Learning and Instruction*, *54*, 114–124. https://doi.org/10.1016/j.learninstruc.2017.08.006

Nikolova, R. (2011). *Grundschulen als differentielle Entwicklungsmilieus: Objektive und subjektive Kontextmerkmales der Schülerzusammensetzung und deren Auswirkung auf die Mathematik- und Leseleistungen [Primary schools as differential development environments: Objective and subjective contextual characteristics of student composition and their impact on mathematics and reading performance]*. Waxmann Verlag.

OECD. (2022). *Education at a Glance 2022: OECD Indicators*. OECD. https://doi.org/10.1787/19991487

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, *3*(1), 285–308. https://doi.org/10.1146/annurev-linguistics-011516-034226

Pakarinen, E., Lerkkanen, M., Poikkeus, A., Siekkinen, M., & Nurmi, J. (2011). Kindergarten teachers adjust their teaching practices in accordance with children's academic pre-skills. *Educational Psychology*, *31*(1), 37–53. https://doi.org/10.1080/01443410.2010.517906

Perry, L. B. (2012). What do we know about the causes and effects of school socio-economic composition? A review of the literature. *Education and Society*, *30*(1), 19–35. https://doi.org/10.7459/es/30.1.03

Purtell, K. M., Ansari, A., Yang, Q., & Bartholomew, C. P. (2021). The role of preschool peers in children's language development. *Seminars in Speech and Language*, *42*(2), 88–100. https://doi.org/10.1055/s-0041-1723838

Pustejovsky J. E. (2024). *clubSandwich: Cluster-Robust (Sandwich) variance estimators with small-sample corrections*. R package version 0.5.11.9999 [Computer software]. http://jepusto.github.io/clubSandwich/

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science*, *23*(3), 425–438. https://doi.org/10.1007/s11121-021-01246-3

R Core Team. (2024). *R: A language and environment for statistical computing.* (Version 4.4.0) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed., Vol. 1). SAGE Publications.

Ready, D. D., & Reid, J. L. (2019). Children's executive function development and school socio-economic and racial/ethnic composition. *Early Childhood Research Quarterly*, *47*, 457–471. https://doi.org/10.1016/j.ecresq.2018.08.002

Reid, J. L., & Ready, D. D. (2013). High-Quality preschool: The socioeconomic composition of preschool classrooms and children's learning. *Early Education and Development*, *24*(8), 1082–1111. https://doi.org/10.1080/10409289.2012.757519

Ribeiro, L. A., Zachrisson, H. D., & Dearing, E. (2017). Peer effects on the development of language skills in Norwegian childcare centers. *Early Childhood Research Quarterly*, *41*, 1–12. https://doi.org/10.1016/j.ecresq.2017.05.003

Riedel, A., Schneider, K., Schuchart, C., & Weishaupt, H. (2010). School choice in German primary schools. How binding are school districts? *Journal for Educational Research Online*, *2*(1), 94–120. https://doi.org/10.25656/01:4569

Rjosk, C. (2020). Dispersion of student achievement and classroom composition. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (pp. 1–33). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_47-1

Rjosk, C., Richter, D., Lüdtke, O., & Eccles, J. S. (2017). Ethnic composition and heterogeneity in the classroom: Their measurement and relationship with student outcomes. *Journal of Educational Psychology*, *109*(8), 1188–1204. https://doi.org/10.1037/edu0000185

Robitzsch, A., & Lüdtke, O. (2023). *Mdmb: Model based treatment of missing data. R package version 1.8-7* [Computer software]. http://doi.org/10.32614/CRAN.package.mdmb

Roßbach, H.-G., Tietze, W., & Weinert, Sabine. (2005). *Peabody Picture Vocabulary Test – Revised. Deutsche Forschungsversion des Tests von L. M. Dunn & L. M. Dunn von 1981. [German research version of the PPVT]*. Universität Bamberg/FU-Berlin.

Rowe, M. L., & Weisleder, A. (2020). Language development in context. *Annual Review of Developmental Psychology*, *2*(1), 201–223. https://doi.org/10.1146/annurev-devpsych-042220-121816

RStudio Team. (2023). *RStudio: Integrated development for R.* [Computer software]. RStudio, PBC. http://www.rstudio.com

Rubin, K. H., Bukowski, W. M., & Bowker, J. (2015). Children in peer groups. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (7. ed., pp. 1–48). Wiley. https://doi.org/10.1002/9781118963418.childpsy405

Rydland, V., Grøver, V., & Lawrence, J. (2014). The potentials and challenges of learning words from peers in preschool: A longitudinal study of second language learners in Norway. In A. Cekaite, S. Blum-Kulka, V. Grøver, & E. Teubal (Eds.), *Children's peer talk. Learning from each other* (pp. 214–234). Cambridge University Press.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 3, pp. 249–277). Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00004-1

Schechter, C., & Bye, B. (2007). Preliminary evidence for the impact of mixed-income preschools on low-income children's language growth. *Early Childhood Research Quarterly*, *22*(1), 137–146. https://doi.org/10.1016/j.ecresq.2006.11.005

Schmerse, D. (2021). Peer effects on early language development in dual language learners. *Child Development*, *92*(5), 2153–2169. https://doi.org/10.1111/cdev.13588

Schmerse, D., & Hepach, R. (2021). How socialization goals and peer social climate predict young children's concern for others: Evidence for a development shift between 2 and 4 years of age. *Social Development*, *30*(1), 239–257. https://doi.org/10.1111/sode.12478

Schneider, T., & Linberg, T. (2022). Development of socio-economic gaps in children's language skills in Germany. *Longitudinal and Life Course Studies*, *13*(1), 87–120. https://doi.org/10.1332/175795921X16233448663756

Schneider, W., Stanat, P., Artelt, C., Blatter, K., Faust, V., Jäger, D., & Schöppe, D. (2017). *Kurz- und langfristige Effekte eines Trainings zur phonologischen Bewusstheit bei Kindergartenkindern deutscher und nichtdeutscher Herkunftssprache (PHONO I) [Short- and Long-term Effects of a Phonological Awareness Training Program for Kindergarteners With and Without German Home Language (PHONO I)]* (Version 1) [Dataset]. Berlin: IQB - Institute for Educational Quality Improvement. https://doi.org/10.5159/IQB_Phono_1_v1

Schoon, I., Nasim, B., & Cook, R. (2021). Social inequalities in early childhood competences, and the relative role of social and emotional versus cognitive skills in predicting adult outcomes. *British Educational Research Journal*, *47*(5), 1259–1280. https://doi.org/10.1002/berj.3724

Tan, C. Y., Hong, X., Gao, L., & Song, Q. (2023). Meta-analytical insights on school SES effects. *Educational Review*, 1–29. https://doi.org/10.1080/00131911.2023.2184329

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

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

UNESCO (Ed.). (2012). *International Standard Classification of Education: ISCED 2011*. UNESCO. http://uis.unesco.org/en/topic/international-standard-classification-education-isced

Van Ewijk, R., & Sleegers, P. (2010a). Peer ethnicity and achievement: A meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, *21*(3), 237–265. https://doi.org/10.1080/09243451003612671

Van Ewijk, R., & Sleegers, P. (2010b). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*(2), 134–150. https://doi.org/10.1016/j.edurev.2010.02.001

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Vitiello, V. E., Booren, L. M., Downer, J. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly*, *27*(2), 210–220. https://doi.org/10.1016/j.ecresq.2011.08.005

Volodina, A., Weinert, S., & Mursin, K. (2020). Development of academic vocabulary across primary school age: Differential growth and influential factors for German monolinguals and language minority learners. *Developmental Psychology*, *56*(5), 922–936. https://doi.org/10.1037/dev0000910

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.; Nachdr.). Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Walberg, H., & Tsai, S.-L. (1983). Matthew effects in education. *American Educational Research Journal*, *20*(3). https://doi.org/10.2307/1162605

Washington-Nortey, P.-M., Zhang, F., Xu, Y., Ruiz, A. B., Chen, C.-C., & Spence, C. (2022). The impact of peer interactions on language development among preschool English Language Learners: A systematic review. *Early Childhood Education Journal*, *50*(1), 49–59. https://doi.org/10.1007/s10643-020-01126-5

Weiland, C., & Yoshikawa, H. (2014). Does higher peer socio-economic status predict children's language and executive function skills gains in prekindergarten? *Journal of Applied Developmental Psychology*, *35*(5), 422–432. https://doi.org/10.1016/j.appdev.2014.07.001

Weinert, S., Roßbach, H.-G., Faust, G., Blossfeld, H.-P., & Artelt, C. (2013). *Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS-3-10) [Educational Processes, Competence Development and Selection Decisions in Preschool and School Age (BiKS-3-10)]* (Version 6) [Dataset]. Berlin: IQB - Institute for Educational Quality Improvement. https://doi.org/10.5159/IQB_BIKS_3_10_V6

Weinert, S., Stanat, P., Heppt, B., & Schuth, E. (2017). *Bildungssprachliche Kompetenzen (BiSpra II): Anforderungen, Sprachverarbeitung und Diagnostik (Verbundvorhaben; 2. Förderphase). Schlussbericht* [Final Report]. Otto-Friedrich-Universität Bamberg. https://doi.org/10.2314/GBV:896641961

Weinert, S., Stanat, P., Heppt, B., Theisen, E., & Köhne-Fuetterer. (2019). *Bildungssprachliche Kompetenzen (BiSpra II)—Anforderungen, Sprachverarbeitung und Diagnostik [Academic Language Competencies (BiSpra II): Affordances, Language Processing, and Diagnostics]* (Version 1) [Dataset]. Berlin: IQB - Institute for Educational Quality Improvement. https://doi.org/10.5159/IQB_BiSpra_II_v1

Weston, S. J., Graham, E. K., & Piccinin, A. M. (2020). Coordinated data analysis: A new method for the study of personality and health. In P. L. Hill & M. Allemand (Eds.), *Personality and healthy aging in adulthood: New directions and techniques* (pp. 75–92). Springer. https://doi.org/10.1007/978-3-030-32053-9_6

Wilkinson, I. A. G., & Fung, I. Y. Y. (2002). Small-group composition and peer effects. *International Journal of Educational Research*, *37*(5), 425–447. https://doi.org/10.1016/S0883-0355(03)00014-4

Willard, J. A., Agache, A., & Leyendecker, B. (2019). Classroom composition, classroom quality and German skills of very young dual language learners and German-only learners. *Early Childhood Research Quarterly*, *49*, 269–281. https://doi.org/10.1016/j.ecresq.2019.07.004

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record*, *112*(4), 1008–1037. https://doi.org/10.1177/016146811011200408

Yang, Q., Ansari, A., Purtell, K. M., Pianta, R. C., Whittaker, J. V., & Vitiello, V. E. (2023). Classroom skill compositions and preschoolers' early academic and executive function outcomes. *Early Childhood Research Quarterly*, *64*, 290–301. https://doi.org/10.1016/j.ecresq.2023.04.003

Yeomans-Maldonado, G., Justice, L. M., & Logan, J. A. R. (2019). The mediating role of classroom quality on peer effects and language gain in pre-kindergarten ECSE classrooms. *Applied Developmental Science*, *23*(1), 90–103. https://doi.org/10.1080/10888691.2017.1321484

Appendix

Additional Tables

Table A1

Percentage of Missing Data for All Included Variables

Data set	N	Language skill	Cognitive ability	DLL-status	SES	Maternal education	Gender
SEIKA-NRW LS Data set 1	626	T1: 7.5 T2: 11.7 T3: 15.2	13.9	0	15.3	10.9	0
SEIKA-NRW LS Data set 2	950	T1: 11.9 T2: 10.9	13.8	2.0	24.4	17.7	0
PHONO 1	572	T1: 9.1 T2: 4.4	2.8	3.8	20.3	20.8	0
BiKS-3-10 preschool	547	T1: 3.1 T2: 8.6 T3: 21.4 T4: 25.8	T1: 2.7 T2: 6.0 T3: 17.6	0	0.4	3.1	0
BiKS-3-10 primary school	898	T1: 8.5 T2: 24.7 T3: 59.1	T1: 9.5 T2: 24.5	4.0	9.7	41.4	0
NEPS SC2 preschool	2975	T1: 4.5 T2: 81.9	11.0	16.8	21.6	16.7	14.8
NEPS SC2 primary school	6912	T1: 6.5 T2: 19.0	11.8	23.1	12.0	14.6	10.8
KuL	1065	T1: 5.1 T2: 6.1	4.9	21.5	15.0	26.3	0.2
BiSPra II	517	T1: 10.3 T2: 9.3	66.0 [a]	2.5	43.5	43.3	0

Note. SEIKA-NRW LS = Language Education and Development in Kindergarten in North Rhine-Westphalia - Longitudinal; PHONO = Short- and Longterm Effects of a Phonological Awareness Training Program for Kindergardeners With and Without German Home Language; BiKS = Educational Processes, Competence Development and Selection Decision in Preschool and School Age; NEPS SC2= National Educational Panel Study, Starting Cohort 2; KuL = Competence Acquisition and Learning Preconditions; BiSpra = Academic Language Competencies – Affordances, Language Processing, and Diagnostics; DLL = dual language learner; SES = socioeconomic status.

[a] The main part of missing values is due to the test booklet design, as not all students were given the corresponding test.

(Appendix continues)

Table A2

Peer Effects of Average Language Abilities and Cross-Level Interaction Effects (Cluster with Fewer than Three Children Excluded)

Study (measurement interval)	Model 0	Model 1		Model 2		Model 3		CLI: peer average language abilities with DLL status		CLI: peer average language abilities with prior language skill		CLI: peer average language abilities with family SES	
	ICC	β	*SE*	β	*SE*	β	*SE*	β	*SE*	β	*SE*	β	*SE*
SEIKA preschool 1–2	0.066	-0.05	0.08	-0.07	0.07	-0.06	0.07	0.03	0.13	0.08	0.06	**-0.18**	0.07
SEIKA preschool 1–3	0.071	-0.03	0.08	-0.04	0.08	-0.03	0.08	0.09	0.15	0.04	0.07	0.02	0.08
SEIKA preschool 2–3	0.120	0.06	0.05	0.04	0.05	0.03	0.05	**0.22**	0.09	-0.03	0.03	0.10	0.05
PHONO 1 preschool	0.370	**0.32**	0.06	**0.27**	0.07	**0.35**	0.08	0.14	0.10	**-0.14**	0.04	-0.01	0.04
NEPS preschool	0.255	0.09	0.09	0.07	0.09	0.08	0.10	-0.18	0.11	-0.01	0.06	-0.09	0.06
BiKS preschool 1–2	0.184	0.08	0.07	0.03	0.07	-0.04	0.07	0.11	0.13	-0.04	0.05	-0.06	0.06
BiKS preschool 1–3	0.166	0.00	0.08	-0.06	0.08	-0.09	0.08	**0.32**	0.14	**-0.12**	0.05	**-0.16**	0.06
BiKS preschool 1–4	0.222	0.11	0.10	0.07	0.09	-0.03	0.10	**0.58**	0.16	**-0.21**	0.06	**-0.26**	0.08
BiKS preschool 2–3	0.171	0.12	0.08	0.01	0.08	0.03	0.09	0.28	0.15	**-0.17**	0.06	**-0.17**	0.07
BiKS preschool 2–4	0.245	**0.25**	0.10	0.17	0.10	0.09	0.12	**0.61**	0.20	**-0.23**	0.08	**-0.33**	0.08
BiKS preschool 3–4	0.205	0.20	0.11	0.20	0.11	0.11	0.12	0.31	0.20	-0.08	0.07	**-0.24**	0.08
BiKS school 1–2	0.146	**0.29**	0.09	**0.24**	0.09	0.16	0.10	0.18	0.18	-0.13	0.07	-0.05	0.08
BiKS school 1–3	0.270	**0.42**	0.17	**0.38**	0.16	0.26	0.17	0.40	0.30	-0.11	0.13	0.09	0.14
BiKS school 2–3	0.236	0.23	0.14	0.16	0.15	0.10	0.17	0.38	0.24	-0.01	0.11	0.17	0.12
NEPS school 1– 3	0.200	**0.07**	0.02	0.04	0.02	-0.01	0.03	0.03	0.04	-0.03	0.01	-0.00	0.02
BiSPra II	0.327	**0.41**	0.10	**0.29**	0.10	**0.28**	0.11	0.07	0.13	0.03	0.06	0.01	0.07
KuL	0.139	-0.03	0.09	-0.05	0.08	-0.14	0.10	0.01	0.11	**-0.16**	0.05	-0.03	0.06
Overall		**0.15**	0.05	0.10	0.04	0.08	0.05	0.11	0.06	**-0.07**	0.03	-0.05	0.04

Note. Cluster with fewer than three children were excluded prior to imputation. Data sets with excluded clusters are presented in italicized text.
CLI = cross-level interaction; DLL = dual language learner; SES = socioeconomic status; ICC = intraclass correlation coefficient. Model 0: intercept-only model; Model 1: peer effect without control variables; Model 2: peer effect controlled for the following Level 1 covariates: gender, DLL status (0 = single language learner, 1 = dual language learner), nonverbal cognitive skills, SES, maternal education (all non-dichotomous Level 1 variables were *z*-standardized); Model 3: additionally controlled for Level 2 SES; CLI-models: based on Model 3. Statistically significant effects ($p < .05$) are printed in bold.

(Appendix continues)

Table A3

Peer Effects of Average Language Abilities and Interaction Effects Based on the Adjusted Average of Peer Abilities

Study (measurement interval)	Model 0	Model 1		Model 2		Model 3		Interaction: peer average language abilities with					
								DLL status		prior language skill		family SES	
	ICC	β	*SE*	β	*SE*	β	*SE*	β	*SE*	β	*SE*	β	*SE*
SEIKA preschool 1–2	0.066	-0.04	0.07	-0.06	0.07	-0.06	0.07	-0.01	0.13	0.06	0.06	**-0.16**	0.07
SEIKA preschool 1–3	0.071	-0.02	0.08	-0.04	0.08	-0.03	0.08	0.06	0.14	0.02	0.06	0.02	0.08
SEIKA preschool 2–3	0.120	0.04	0.04	0.03	0.04	0.02	0.04	**0.15**	0.08	-0.02	0.04	0.07	0.05
PHONO 1 preschool	0.370	**0.30**	0.06	**0.25**	0.06	**0.32**	0.07	0.10	0.10	**-0.14**	0.04	-0.00	0.04
NEPS preschool	0.255	0.07	0.07	0.05	0.07	0.06	0.07	-0.18	0.11	-0.05	0.06	-0.09	0.05
BiKS preschool 1–2	0.184	0.06	0.06	0.03	0.05	-0.04	0.06	-0.03	0.12	-0.01	0.05	-0.03	0.05
BiKS preschool 1–3	0.166	0.01	0.07	-0.04	0.06	-0.06	0.07	0.20	0.13	-0.10	0.06	-0.09	0.06
BiKS preschool 1–4	0.222	0.09	0.08	0.06	0.08	-0.03	0.08	**0.38**	0.15	**-0.21**	0.07	**-0.21**	0.07
BiKS preschool 2–3	0.171	0.10	0.07	0.01	0.07	0.02	0.08	0.17	0.14	**-0.13**	0.06	-0.09	0.07
BiKS preschool 2–4	0.245	**0.21**	0.08	0.14	0.08	0.09	0.09	**0.39**	0.18	**-0.25**	0.08	**-0.25**	0.07
BiKS preschool 3–4	0.205	0.16	0.09	0.15	0.09	0.08	0.10	0.16	0.20	-0.07	0.07	**-0.21**	0.08
BiKS school 1–2	0.146	**0.23**	0.08	0.19	0.08	0.12	0.09	0.10	0.16	-0.14	0.07	-0.03	0.08
BiKS school 1–3	0.270	**0.36**	0.14	**0.32**	0.14	0.23	0.15	0.30	0.27	-0.11	0.13	-0.07	0.13
BiKS school 2–3	0.236	0.20	0.12	0.14	0.13	0.07	0.15	0.36	0.23	-0.01	0.10	0.16	0.11
NEPS school 1– 3	0.200	**0.06**	0.02	0.03	0.02	-0.00	0.02	0.03	0.04	**-0.05**	0.02	-0.02	0.02
BiSPra II	0.327	**0.36**	0.09	**0.26**	0.09	**0.25**	0.10	0.05	0.13	0.02	0.07	-0.00	0.07
KuL	0.139	-0.03	0.08	-0.05	0.08	-0.13	0.09	-0.01	0.11	**-0.12**	0.05	0.01	0.06
Overall		**0.13**	0.05	0.09	0.04	0.07	0.05	0.07	0.04	**-0.07**	0.02	-0.04	0.03

Note. Cluster with fewer than three children were excluded prior to imputation. The peer averages were calculated excluding the target child in each imputed data set. DLL = dual language learner; SES = socioeconomic status; ICC = intraclass correlation coefficient. Model 0: intercept-only model; Model 1: peer effect without control variables; Model 2: peer effect controlled for the following Level 1 covariates: gender, DLL status (0 = single language learner, 1 = dual language learner), nonverbal cognitive skills, SES, maternal education (all non-dichotomous Level 1 variables were *z*-standardized); Model 3: additionally controlled for Level 2 SES; Interaction-Models: based on Model 3. Statistically significant effects ($p < .05$) are printed in bold.

Table 1

Study Characteristics and Sample Description

Study	Year of data collection	Sample	Reference	Analysis sample: Included time points	*n* Children	*n* Classes	Average cluster size	SES [a] Median /Mean (SD)	% female	% DLL	Maternal education (% University degree)
SEIKA-NRW LS [b]	2016-2019	Kindergarten children in North Rhine-Westphalia aged 2;6 to 3;11 years at first assessment	Data: Bihler et al. (2021)	3	Data set 1: 626	Data set 1: 162	3.86	1489.03	49.52	32.75	22.32
				2	Data set 2: 950	Data set 2: 192	4.95	1487.86	49.26	36.32	22.00
PHONO 1 [b]	2009-2011	Children in their last year of kindergarten to grade 1 in three German states (Baden-Wuerttemberg, Bavaria, Berlin)	Data: W. Schneider et al. (2017)	2	572	45	12.7	50.67 (18.04)	49.48	33.92	- [c]
BiKS-3-10 [b]	2005-2012	Children in their first year of kindergarten (who were supposed to enter primary school three years later) to grade 4 of primary school in two German states (Bavaria, Hesse)	Data: Weinert et al. (2013)	4	Preschool: 547	Preschool: 97	5.64	52.46 (16.28)	48.10	19.56	22.12
				3	Primary school: 898	Primary school: 188	5.10	53.26 (15.79)	53.12	17.48	15.37
NEPS SC2 [d]	2011-2021	Children in their penultimate year of kindergarten (Wave 1) from a nationwide, representative sample of German kindergarten and students at regular schools from grade 1 to 7 (from Wave 3 onwards)	Blossfeld & Roßbach (2019) Data set: NEPS-Netzwerk (2022)	2	Preschool: 2975	Preschool: 720	4.13	54.73 (20.47)	42.00	21.38	17.31
				2	Primary school: 6912	Primary school: 1024	6.75	59.79 (19.35)	45.46	13.63	20.80
KuL [b]	2013, 2014	Students in grade 1 in North Rhine-Westphalia	Data: Kristen et al. (2018)	2	1065	66	16.14	53.20 (19.50)	48.08	23.94	17.56
BiSpra II [b]	2014, 2015	Students in grades 2 and 3 (at Time 1) in four German states (Baden-Wuerttemberg, Brandenburg, Hamburg, North Rhine-Westphalia)	Data [e]: Weinert et al. (2019) Final report: Weinert et al. (2017) Codebook: Heppt et al. (2019)	2	517 [f]	38	13.61	40.89 (18.78)	49.52	55.32	8.51

Note. All descriptive information is based on the raw data.

SEIKA-NRW LS = Language Education and Development in Kindergarten in North Rhine-Westphalia - Longitudinal; PHONO = Short- and Longterm Effects of a Phonological Awareness Training Program for Kindergardeners With and Without German Home Language; BiKS = Educational Processes, Competence Development and Selection Decision in Preschool and School Age; NEPS SC2 = National Educational Panel

Study, Starting Cohort 2; KuL = Competence Acquisition and Learning Preconditions; BiSpra = Academic Language Competencies – Affordances, Language Processing, and Diagnostics; SES = socioeconomic status; DLL = dual language learner.

[a] For SEIKA data sets monthly net equivalized household income in Euros; all others Highest International Socio-Economic Index of Occupational Status (HISEI, Ganzeboom et al., 1992).

[b] The data were made available by the Research Data Centre at the Institute for Educational Quality Improvement (FDZ at IQB). [c] In the PHONO 1 study, only information on the highest completed school degree was available. The proportion of mothers with a university entrance qualification was 21.5%. [d] This paper uses data from the National Educational Panel Study (NEPS; see Blossfeld & Roßbach, 2019). The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network. [e] The data were collected as part of the BiSpra II project (Head: S. Weinert, P. Stanat; funded by BMBF [Federal Ministry of Education and Research]). [f] Subsample of the total sample (N = 1179) comprising only students, who were presented with a test booklet including the academic vocabulary subtest.

Table 2

Overview of Language Tests Used in the Included Studies and Descriptive Information

Data set	Time point	Mean age (years)	Details on language test	*M* (*SD*) [a]	ICC
SEIKA-NRW LS Data set 1	T1	3.4	PPVT 4 [b]; Items: 228; α = .81	52.39 (26.46)	.11 [c]
	T2	4.3		81.11 (24.91)	.05
	T3	5.2		102.24 (24.38)	.05
SEIKA-NRW LS Data set 2	T1	4.1	PPVT 4 [b]; Items: 228;	71.82 (28.93)	.12
	T2	5.0		92.76 (28.55)	.11
PHONO 1			PPVT-R [d]; Items: 175;		
	T1	5.5	WLE reliability = 0.94;	65.33 (24.22)	.39
	T2	- [e]	WLE reliability = 0.93	77.99 (24.36)	.38
BiKS-3-10 preschool	T1	3.7	PPVT-R [d]; Items: 175; Split-half reliability = .82 - .96	34.2 (18.98)	.27
	T2	4.7		56.11 (21.71)	.19
	T3	5.6		77.80 (22.10)	.20
	T4	7.1		103.33 (17.25)	.16
BiKS-3-10 primary school	T1	7.4	KFT 1-3 Subtest Verbal Comprehension [f]; Items: 15	7.09 (2.51)	.11
	T2	8.2		8.66 (2.62)	.12
	T3	9.3		10.11 (2.37)	.15
NEPS SC2 preschool			Adapted PPVT-R [g];		
	T1	5.0	Items: 77; WLE reliability = .89	-0.05 (1.12) (WLE - score)	.26
	T2	7.0	Items: 66; WLE reliability = .87	1.56 (0.77) (WLE - score)	.20
NEPS SC2 primary school			Adapted PPVT-R [g]		
	T1	7.0	Items: 66; WLE reliability = .87	1.44 (0.84) (WLE - score)	.22
	T2	8.7	Items: 72; WLE reliability = .84	2.51 (0.94) (WLE - score)	.20
KuL	T1	6.0	FIPS Subtest Phonological Awareness [h]; Items: 26; α = .82	13.28 (4.63)	.16
	T2	7.0		20.83 (4.62)	.14
BiSpra II			BiSpra 2-4 Subtest Academic Vocabulary [i]		
	T1	8.4	Items: 24 [j]; EAP Reliability .53-.67	0.01 (0.89) (WLE - score)	.24
	T2	9.4	Items: 29 [j]; EAP Reliability .69-.75	0.01 (1.14) (WLE - score)	.33

Note. All descriptive information is based on the raw data.

PPVT = Peabody Picture Vocabulary Test; ICC = intraclass correlation coefficient; EAP = expected a posteriori; WLE = weighted maximum likelihood estimate; α = Cronbach's alpha.

[a] Refers to the mean of the raw sum scores unless otherwise stated. [b] German version (Lenhard et al., 2015). [c] The ICC of T1 refers to the sample including the peers from the cross-sectional part (see Supplement A for details). [d] German research version, (Roßbach et al., 2005); previous research using the PPVT-R version reported split-half reliabilities for 4- and 8-years-old, respectively (European Child Care and Education (ECCE) Study Group, 1999). [e] T2 occurred approximately 20 weeks after T1. [f] KFT 1-3 = Kognitiver Fähigkeitstest 1-3 (Cognitive ability test), Heller & Geisler, 1983; α (overall KFT with 4 subtests) = .80. [g] Adapted version of PPVT-R, reliabilities reported by Fischer and Durda (2020). [h] FIPS = Fähigkeitsindikatoren Primarschule (German version of Performance Indicators in Primary Schools [PIPS]), Bäuerlein et al., 2012; reliabilitiy reported by Gentrup and Rjosk (2018). [i] BiSpra 2-4 = Test zur Erfassung bildungssprachlicher Kompetenzen bei Grundschulkindern der Jahrgangsstufen 2 bis 4 [Test for Assessing Academic Language Competencies in Primary School Students from Grades 2 to 4],

Heppt et al., 2020; reliabilities reported in the codebook (Heppt et al., 2019). [j] This represent the total number of items, but depending on grade and test booklet a selection of these items were presented.

Table 3

Peer Effects of Average Language Abilities and Cross-Level Interaction Effects on Children's Language Skills

Study (measurement interval)	Model 0	Model 1		Model 2		Model 3		CLI peer average language abilities with					
								DLL status		prior language skill		family SES	
	ICC	β	*SE*	β	*SE*	β	*SE*	β	*SE*	β	*SE*	β	*SE*
SEIKA preschool 1–2	0.07	-0.05	0.08	-0.07	0.07	-0.06	0.07	0.03	0.13	0.08	0.06	**-0.18**	0.07
SEIKA preschool 1–3	0.07	-0.03	0.08	-0.04	0.08	-0.03	0.08	0.09	0.15	0.04	0.07	0.02	0.08
SEIKA preschool 2–3	0.13	0.07	0.05	0.05	0.05	0.06	0.05	0.13	0.08	-0.01	0.03	-0.01	0.06
PHONO 1 preschool	0.37	**0.32**	0.06	**0.27**	0.07	**0.35**	0.08	0.14	0.10	**-0.14**	0.04	-0.01	0.04
NEPS preschool	0.26	0.09	0.07	0.06	0.08	0.08	0.08	0.00	0.11	-0.03	0.05	**-0.11**	0.05
BiKS preschool 1–2	0.18	0.08	0.07	0.03	0.07	-0.04	0.07	0.11	0.13	-0.04	0.05	-0.06	0.06
BiKS preschool 1–3	0.17	0.00	0.08	-0.06	0.08	-0.09	0.08	**0.32**	0.14	**-0.12**	0.05	**-0.16**	0.06
BiKS preschool 1–4	0.22	0.11	0.10	0.07	0.09	-0.03	0.10	**0.58**	0.16	**-0.21**	0.06	**-0.26**	0.08
BiKS preschool 2–3	0.17	0.12	0.08	0.01	0.08	0.03	0.09	0.28	0.15	**-0.17**	0.06	**-0.17**	0.07
BiKS preschool 2–4	0.25	**0.25**	0.10	0.17	0.10	0.09	0.12	**0.61**	0.20	**-0.23**	0.08	**-0.33**	0.08
BiKS preschool 3–4	0.21	0.20	0.11	0.20	0.11	0.11	0.12	0.31	0.20	-0.08	0.07	**-0.24**	0.08
BiKS school 1–2	0.13	**0.16**	0.08	0.14	0.08	0.13	0.08	0.24	0.14	-0.09	0.05	-0.07	0.06
BiKS school 1–3	0.19	0.14	0.13	0.12	0.13	0.10	0.14	0.05	0.24	-0.07	0.08	-0.05	0.11
BiKS school 2–3	0.19	0.16	0.12	0.12	0.12	0.10	0.12	0.28	0.23	-0.02	0.08	0.13	0.11
NEPS school 1– 3	0.20	**0.07**	0.02	**0.04**	0.02	0.00	0.02	0.05	0.04	**-0.03**	0.01	-0.01	0.02
BiSPra II	0.33	**0.41**	0.10	**0.28**	0.10	**0.25**	0.11	0.06	0.12	0.03	0.06	0.03	0.07
KuL	0.14	-0.03	0.09	-0.05	0.08	-0.14	0.10	0.01	0.11	**-0.16**	0.05	-0.03	0.06
Overall		**0.13**	0.05	0.09	0.04	0.07	0.05	**0.12**	0.04	**-0.06**	0.02	-0.06	0.03

Note. CLI = cross-level interaction; DLL = dual language learner; SES = socioeconomic status; ICC = intraclass correlation coefficient.
Model 0: intercept-only model; Model 1: peer effect without control variables; Model 2: peer effect controlled for the following Level 1 covariates: gender, DLL status (0 = single language learner, 1 = dual language learner), nonverbal cognitive skills, SES, maternal education (all non-dichotomous Level 1 variables were *z*-standardized); Model 3: additionally controlled for Level 2 SES; CLI-models: based on Model 3. Statistically significant effects ($p < .05$) are printed in bold.

Figure 1

Peer Effects of Average Language Abilities on Language Skills at T2 in Model 2 and Model 3.

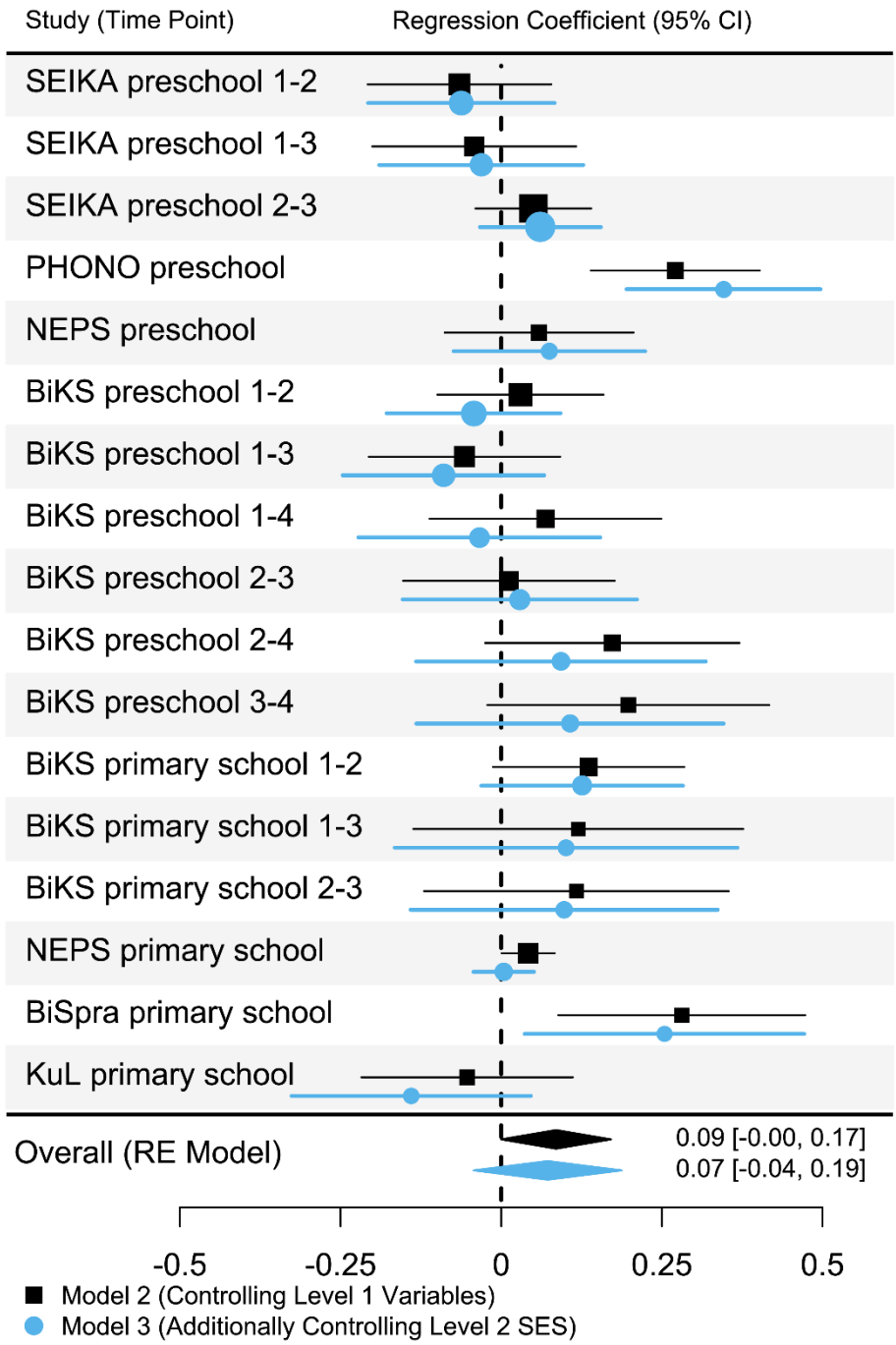
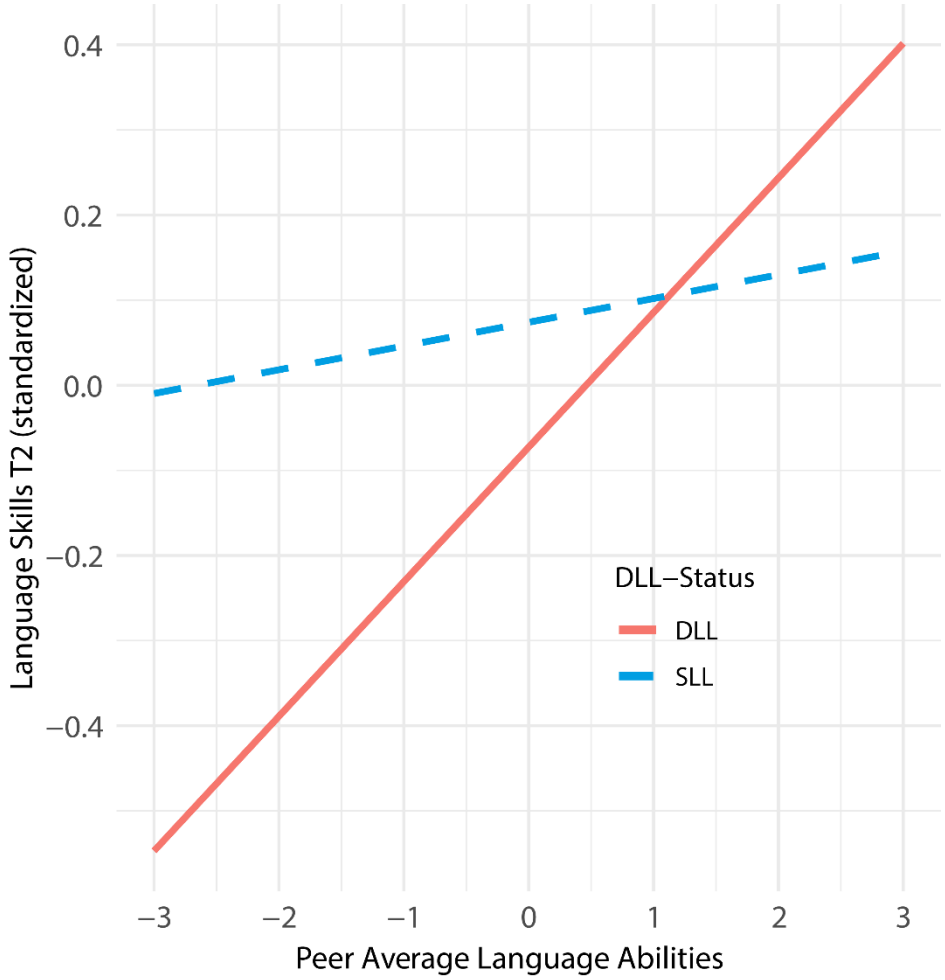


Figure 2

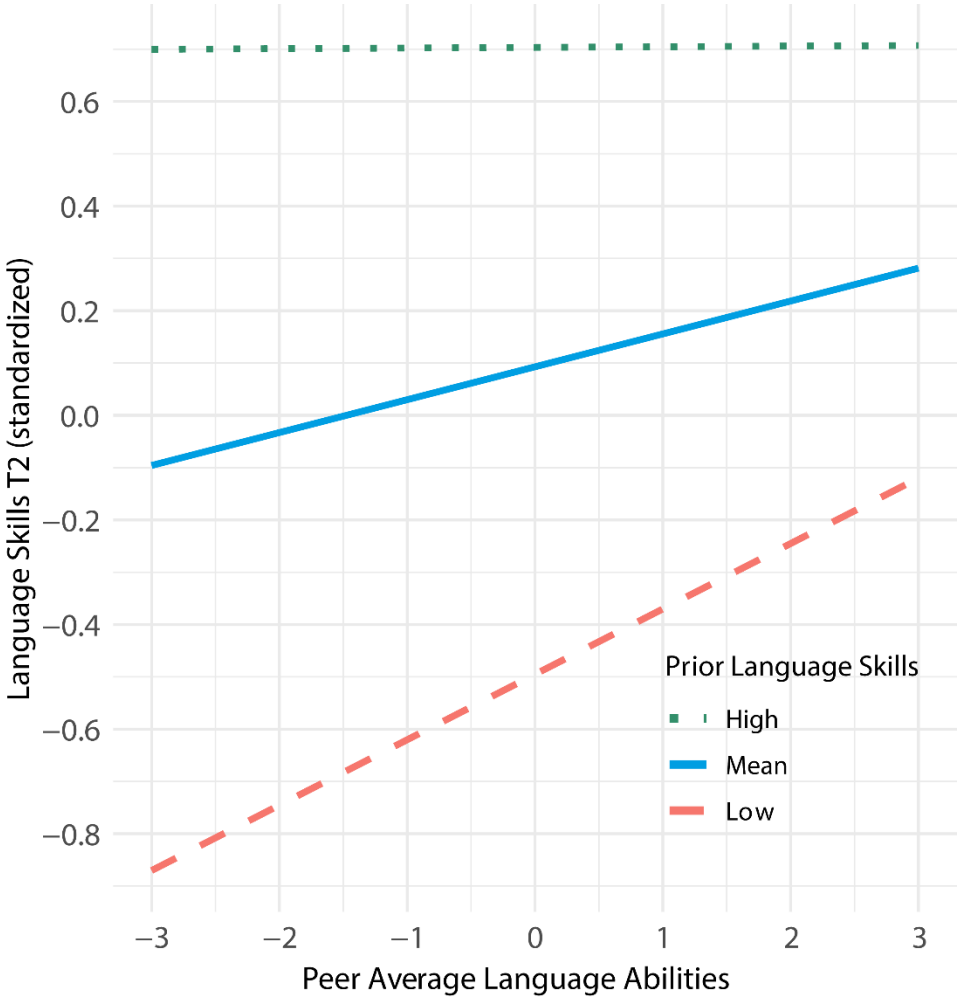
Illustration of the Cross-Level Interaction Between Peer Language Abilities and DLL Status Based on Pooled Coefficients.



Note. DLL = dual language learner; SLL = single language learner. The slope for DLL significantly differs from zero.

Figure 3

Illustration of the Cross-Level Interaction Between Peer Language Abilities and Prior Language Skills Based on Pooled Coefficients.



Note. *Low*, *mean*, and *high* prior language skills correspond to –1, 0, and +1 SD from the mean. The slope at low prior language skills significantly differs from zero.