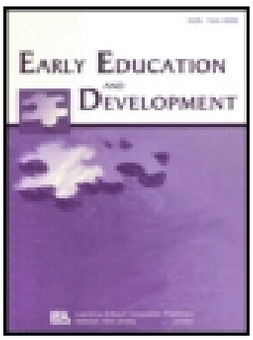

Early Education and Development

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/heed20

Refugee Children's Early Development during Attendance of Specialized Preschool Programs and Transition into First Grade in Germany

J. Busch, N. Cabrera, F. Ialuna, T. Buchmüller & B. Leyendecker

To cite this article: J. Busch, N. Cabrera, F. Ialuna, T. Buchmüller & B. Leyendecker (2021): Refugee Children's Early Development during Attendance of Specialized Preschool Programs and Transition into First Grade in Germany, Early Education and Development, DOI: 10.1080/10409289.2021.1970427

To link to this article: https://doi.org/10.1080/10409289.2021.1970427

Published online: 29 Aug 2021.

EARLY EDUCATION AND DEVELOPMENT
https://doi.org/10.1080/10409289.2021.1970427

Routledge
Taylor & Francis Group

Refugee Children's Early Development during Attendance of Specialized Preschool Programs and Transition into First Grade in Germany

J. Busch [a], N. Cabrera [b], F. Ialuna [a], T. Buchmüller [a], and B. Leyendecker [a]

[a]Child and Family Research, Faculty of Psychology, Ruhr-University Bochum; [b]Department of Human Development and Quantitative Methodology, University of Maryland

ABSTRACT

Research Findings: We assessed socio-emotional behavior, nonverbal reasoning, German receptive language, and motor skills of refugee children attending early childhood development [ECD] programs and of those who did not (N = 207, mean age = 69.4 months). Young refugee children overall demonstrated lower levels of development and more socio-emotional behavior problems. Attendance of preschool-based ECD programs was inconsistently linked to better outcomes. Only moderate improvements in German receptive language skills could be supported across different methodological approaches. Although socio-emotional problems of refugee children attending ECD programs persisted on high levels, those children showed overall fewer problems when compared to non-attenders at the transition to first grade, especially less hyperactivity/inattention and more prosocial behavior. *Practice or Policy*: Our study supports that refugee experiences during early childhood are linked to lower developmental learning foundations. Specialized ECD programs for refugees can compensate a general shortage in regular ECD services in times of increased demands. Such programs thus increase the chances of refugee children to keep pace academically with their non-refugee peers. However, as specialized programs for refugee children establish a non-inclusive route in the early education sector of Germany, they still have to empirically prove quality and promoting effects on the children's ECD.

In 2019, more than 31 million children below the age of 18 were internationally displaced from their countries of residency (United Nations International Children's Emergency Fund [UNICEF], 2019). In Germany, the only high-income country among the top 10 refugee-hosting countries worldwide, approximately one-third of the recently arrived 1.2 million refugees and asylees are below the age of 6 (German Federal Agency for Migration and Refugees, 2020). Beyond legal definitions, refugee migration constitutes forced displacement amid economic and social hardship (e.g., natural disasters, poverty), physical threats (e.g., persecution, civil war) and environmental chaos (breakdown of governments and societies). The interplay of stressors refugee children experience before and after resettlement are heterogeneous (Almqvist & Broberg, 1999; Montgomery, 2008). During displacement, refugee children are likely to be exposed to violence, separation, malnutrition, and a lack of age-adequate stimulation. After resettlement in high-income countries, children are more likely to find themselves in physically safer settings (Reed, Fazel, Jones, Panter-Brick, & Stein, 2012). In such contexts, however, refugee children still have to adapt to new socio-cultural systems and to acquire the relevant competences to benefit from host countries' education (Busch & Leyendecker, 2019).

CONTACT J. Busch julian.busch@rub.de Child and Family Research, Faculty of Psychology, Ruhr-University Bochum, Building/Room: GAFO 04/607, Universitätsstraße 150, Bochum 44801, Germany

Overall, experiences related to displacement increase children's risk of not acquiring the necessary developmental foundations required for positive learning trajectories (Bouchane et al., 2018). Among young refugee children, previous studies have demonstrated socio-emotional problems including attention difficulties, anxiety, conflicts with peers (Buchmüller et al., 2020; Kien et al., 2019), lower cognition-based performances and pre-academic skills (Aghajafari et al., 2020; Graham et al., 2016; Rousseau et al., 1996), and even higher rates of stunted growth (Walpole et al., 2018). This evidence is rather heterogeneous as it reflects situations of young children among different contexts and stages of forced migration. In the present study, we aim to distinctively understand young refugee children's readiness for elementary school enrollment during resettlement periods within a high-income country.

Readiness for school refers to an interdependent set of developmental domains that are necessary for academic learning of young (refugee) children within elementary school settings. According to the American Academy of Pediatrics, such domains include motor, social and emotional, language and cognitive development (High, 2008). Previous studies with non-refugee children described mechanisms of how such domains are relevant for academic learning and substantiate links to later academic achievement.

Basic knowledge of the majority language is essential to build relationships, to navigate social situations, and to benefit from verbally mediated instruction. Evidence from second-language learners shows that competence in the majority language was critical to keep pace with math and reading achievements throughout the elementary school years (Halle et al., 2012). Oral academic language including listening comprehension and vocabulary have been argued to present the language domains most critical for second language learners' academic achievement (Bailey & Butler, 2007). Beyond language, children's strong cognitive abilities promote academic learning as they represent core competencies to solve educational tasks. Specifically, strong cognitive reasoning skills, covering how well-novel problems can be mastered, were found to predict better math and language skills from third to fourth grade of elementary school (Lakin, 2012). A longer duration of enrollment in early education was linked to better cognitive skills (Burrage et al., 2008). Further, motor skills have demonstrated effects on academic achievement (e.g., Carlson et al., 2013; Murrah, 2010). They can be differentiated into fine, gross, and visual-motor skills and comprise physical movements but also their neurocognitive underpinnings. While within education contexts gross motor skills are essential for developing physical well-being and social competence (Wilson et al., 2013), fine motor skills encompass the coordination and precise movements needed for many learning activities (e.g., oral language, paper-pencil tasks). Finally, building socio-emotional skills during early childhood facilitates academic achievement during the elementary school years (Ladd et al., 1999). Better socio-emotional skills were specifically found to support positive relations to peers and teachers, foster belongingness and adjustment and also children's engagement in education. Overall, the different contributors to academic learning also seem to be interrelated with a strong focus on cognitive abilities (Sabol & Pianta, 2012).

Early childhood development [ECD] programs have the overall goal to facilitate the unfolding of children's individual potential across developmental domains relevant for academic learning (e.g., Howes et al., 2008). Therefore, ECD programs can include all child- and caregiver-centered efforts for preventive (i.e., children at-risk for developmental disturbances) and compensative (i.e., children already affected by hampering circumstances) initiatives below school-entry level. A substantial body of studies has investigated variations of effectiveness among ECD programs. Outcomes of this research indicate that (1) process (e.g., stimulating teacher–child interactions) and structural characteristics (e.g., child-suited premises and availability of educative material) of ECD programs show relations with child development, and (2) longer program attendance may be particularly beneficial for children from immigrant families and those living in poverty (Ramey et al., 2000; Weiland & Yoshikawa, 2013). Especially for at-risk populations, ECD programs can enrich learning opportunities by providing nurturing environments and stimulating interactions. Evidence, however, on refugee children attending ECD programs, the programs' impact and the determinants thereof is still limited (Murphy et al.,

2018). A recent report reviewed national policies on ECD services and their implications for supporting refugee children among nine high-income countries (Park et al., 2018). This report suggested that countries' responses do not meet the legal requirements for this specific group (e.g., supporting access to national ECD services) and described a general lack of guiding evidence on how to effectively serve the ECD needs of refugee children. Achieving a better understanding of young refugee children's development therefore provides a basis for further action, thus creating targeted ECD programs that specifically address children's developmental needs, facilitating their learning and enrollment into elementary schools after resettlement.

Some previous studies on refugees and related underserved populations have focused on ECD programs initiated in conflicted and deprived settings by non-governmental institutions (see Murphy et al., 2018). For mother-child dyads affected by the Yugoslav wars, a 5-month-long caregiver-centered intervention in combination with medical checkups increased the quality of maternal caregiving behavior, improved children's cognitive development, and reduced their socio-emotional problem behavior (Dybdahl, 2001). In a rural region of Pakistan, a community-based intervention providing psychosocial stimulation and nutrition promoted physical, social-emotional and cognitive development of deprived children (Yousafzai et al., 2014). In Uganda, refugee children's attendance of high-quality playgroups for 3 months fostered their well-being and overall child development (Metzler et al., 2019). Evidence suggests that specialized ECD programs for refugee children are effective in their specific contexts of implementation. Specifically, programs in low-resource contexts demonstrated effectiveness when they combined development-stimulating activities for children or caregiver support with the provision of basic needs, such as nutrition, health, and protection. Notably, ECD programs in low-resource contexts were less likely organized by governmental stakeholders and not linked to the national frameworks of ECD policies. In non-conflicted, high-resource contexts of resettlement, however, the basic needs are more likely secured. The implementation of specific ECD programs can focus on education and healthcare services, for example, with the goal of preparing refugee children for school entry. Still, empirical studies on specialized ECD programs in high-resource resettlement contexts are very limited. A study by Rousseau et al. (2009) found promising effects of creative expression workshops for refugee children attending preschools in Canada. The study documented that impact of a low-intensity program can support children's socio-emotional development. Such evidence suggests that some ECD domains of refugee children could be especially susceptible to change through program attendance.

Beyond the specialized interventions, regular ECD services in high-resource contexts are likely to be available to refugee children as well. Accessing these services can, however, be difficult when the demand from the native-born population is high and large numbers of young refugee children arrive within a short-time period (Gambaro et al., 2017; Morantz et al., 2013). In consequence, many arriving refugee children are at risk of entering first grade without prior attendance of ECD programs (Park et al., 2018). In Germany in 2015, targeted ECD programs for recently arrived refugee children were funded based on a new policy, so-called "Bridging Projects" [BPs]. The policies' overall goal was to serve the developmental needs of young refugee children. Stakeholders such as the Communal Youth Welfare offices or private childcare agencies were therefore granted flexibility in implementing BPs to suit the local circumstances and diverse demands of refugee families with young children. There were few requirements for BPs, but they included a minimum of one staff member with an early education certification and a teacher–child ratio of 1 adult for 5 children, or lower. BP attendance is subsidized by the state and providers receive a rate of 30€ per 5 children per hour. Close to one thousand BPs have been funded annually since 2015. They differ in format (e.g., parent–child groups, child-only groups), settings (e.g., public spaces, refugee accommodations, facilities for education), group sizes (M = 8.60, SD = 4.05), and frequency (M = 10.41, SD = 8.27 weekly hours; all own calculations based on data provided by state authorities; see also Busch et al., 2021).

In our study, we focus on BPs providing preschool-based programs to prepare refugee children for the transition into first grade. These BPs intend to facilitate children's development for subsequent academic learning through quality early education. They are exclusively set up in daycare centers or

elementary schools in small groups, are run by better trained staff when compared to other BPs, and they offer frequent attendance. Preschool-based BPs aim to achieve high structural and process quality by providing fixed curricula. Those curricula combine play-based and instructional learning activities to facilitate different domains of development and also behavioral adjustment. For example, the programs offer arts and crafts (promoting fine motor skills), apply repeating timetables for session structuring (adaptation to temporal routines), combine group-based, individual, and peer activities (regulation of impulses, behavior, and emotions in peer and teacher interactions), and explore outside-world phenomena (stimulate cognitive development). Additionally, the preschool-based BPs emphasize German language acquisition through a number of language learning activities (shared reading, storytelling, singing, facilitating, and scaffolding conversation).

Current Study

Investigating the preschool-based BPs, the present study contributes to our understanding how we can foster the ECD of young refugee children. While young refugee children appear at-risk for reaching school readiness, we have only limited knowledge about their specific developmental needs during resettlement periods in a high-income country. We moreover still lack guiding evidence on ECD programs for young refugee children as previous studies primarily report on non-governmentally funded ECD programs from low-resource, conflicted, or deprived contexts. Therefore, our goals were to investigate child development of recently arrived refugee children attending preschool-based BPs during the resettlement period in Germany. We (1) expected that young refugee children would show lower levels compared to non-refugee peers but would demonstrate positive changes over time of BP attendance. Based on prior research mainly with non-refugee children or in low-income countries, we (2) expected that refugee children who attend BPs would demonstrate better development at the transition to first grade when compared to refugee children who did not attend any ECD programs.

Methods

Participants

We overall included n_1 = 152 refugee children attending preschool-based BPs and n_2 = 55 attending first grade in our study. As the primary study group, we assessed refugee children from 10 preschool-based BPs that were all part of the same regional ECD initiative to support refugee children within certain districts of the largest metropolitan area in Germany (Ruhr-Area). Those districts are multi-ethnic, have a population with low socio-economic status and, as traditional arrival quarters for immigrants, they are hosting disproportionally large numbers of recently arrived refugee families. Children from preschool-based BPs eligible for participation (study sample T_1) were between 3 and 7 years of age, had arrived in Germany within the three years prior to data collection, and attended BPs regularly for at least two months before assessments. As a comparison group, we additionally assessed refugee children in the first grade. Children for that group were eligible if they had neither attended BPs nor any other ECD programs beforehand and their elementary schools were located within the same districts of the participating BPs. We conducted all assessments between summer 2017 and late summer 2018 in three phases: initially at T_1 (n_{T1} = 152), again five months later, T_2 (n_{T2} = 73, repeated measure sample), and the comparison group with the beginning of the new school year in late summer 2018, T_C (n_{TC} = 55). See Figure 1 for an illustration of this study's sample structure.

We chose an inter-assessment interval to balance our research goals (especially observing change) with practical demands. Regarding our goals, we planned an interval length that was at least comparable to previous studies on the development of young refugee children attending ECD programs (e.g., Metzler et al., 2019). At the same time, we practically intended to limit the expected rate of children stopping BP attendance during the inter-assessment interval, especially due to the start of the new academic year. Throughout the longitudinal data collection period, fifty percent of the overall T_2 sample (n_{T2} = 73)

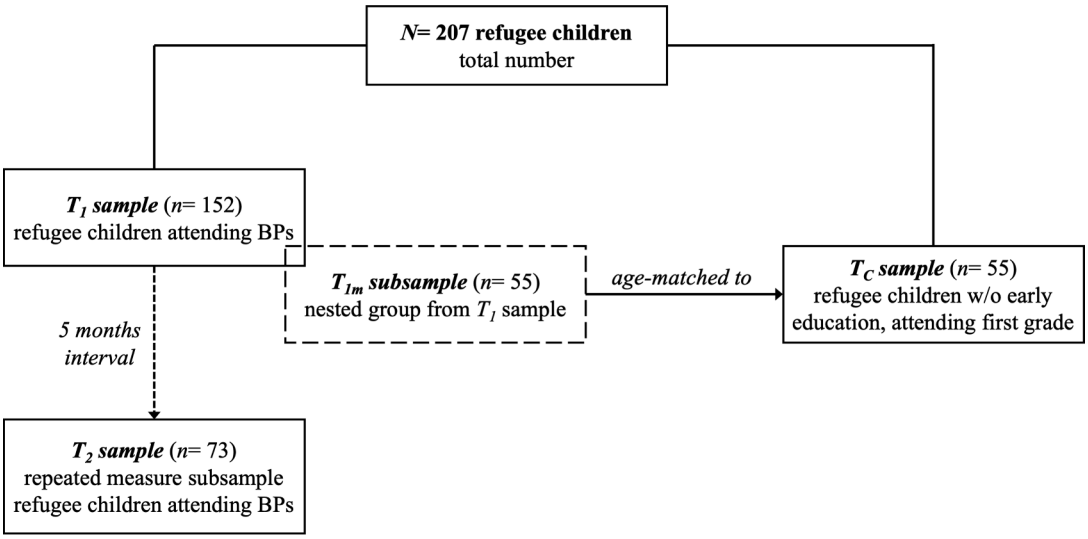

Figure 1. Illustration of study samples. The study design consisted of three approaches, involving different samples/subsamples. (A) we compared the T_1 sample to norm data and linked child outcomes to previous duration of BP attendance (cross-sectional approach); (B) we investigated changes from T_1 to T_2 throughout five months of additional attendance (longitudinal approach); (C) we selected an age-matched comparison group from T_1 and compared to the T_C sample (T_{1m} subsample, quasi-experimental approach).

transitioned from BPs into first grade mostly toward the end of the five-month period. Ensuring the inter-assessment interval of our research design, those children were still considered for the repeated assessment at T_2. Drop out between T_1 and T_2 was mostly due to families moving to distant locations. The repeated measurement subsample (T_2) did not differ regarding their age, gender-ratio, time of arrival in Germany, length of BP attendance, and region-of-origin composition from the T_1 sample. Refugee children from T_C were older and included fewer children from North African countries compared to the children in T_1 sample. For demographic information on the three study samples see Table 1.

Procedures

At the time of enrollment into the BPs and at transition into first grade, teachers informed refugee caregivers about our investigation. Caregivers who allowed their children to participate signed an informed consent form. Nine research assistants, mainly with Bachelor's degrees in psychology, were

Table 1. Sample characteristics.

	T_1 sample (n_{T1} = 152)	T_2 sample (n_{T2} = 73)	T_{1m} subsample (n_{T1m} = 55)	T_C sample (n_{TC} = 55)
Age in months[a], *mean* (*SD*)	69.4 (10.2)	70.4 (10.4)	78.49 (5.05)	86.13 (7.5)
Gender female (%)	50.4	47.0	51.0	35.9
Region of origin (%)				
Middle East	23.8	21.9	25.0	29.1
Southeastern Europe	34.7	41.1	55.0	34.6
North Africa	19.1	21.9	5.0	3.6
Subsaharan Africa	1.4	1.4	5.0	1.8
Unknown	21.1	13.7	10.0	30.9
Time since arrival in Germany[b], *mean* months (*SD*)	26.4 (21.1)	26.3 (20.0)	27.19 (20.75)	23.97 (16.2)
Previous length of BP attendance[c], *mean* months (*SD*)	6.5 (4.9)	6.9 (4.6)	6.14 (4.43)	–

SD, standard deviation. *N/n*, number of cases. T_1-T_2 inter-measurement interval is *M* = 160 days (Median 154 days).
[a]for T_1 and T_2 samples: age at T_1 assessments
[b]for T_1 and T_2 samples: absolute time since arrival until T_1 assessments
[c]for T_1 and T_2 samples: until T_1 assessment

trained in child direct assessment procedures. The overall 12-h training comprised a theoretical (e.g., reading about constructs and procedures) and practical phase (e.g., introduction into child assessments, conducting mock assessments). At the end of the training, each observer had to demonstrate proficiency in assessment procedures in a final test. Research assistants administered child direct assessments of refugee children's development in separate rooms during BP program hours or, for the comparison group, during morning hours in elementary schools. The individual testing of each child lasted around 30–40 minutes. All research assistants participated in regular debriefings and were supervised by the study authors. Teachers were asked to assess refugee children's socio-emotional behavior within education contexts. Additionally, we assessed the quality of the preschool-based BPs in structured field observations. The full study protocol was approved by the Ethics Committee of the Faculty of Psychology, Ruhr-University Bochum (Num. 381, 2017).

Measures

We addressed the four developmental domains cognition, (host country) language, motor skills, and socio-emotional behavior. We selected indicators with the rationale to balance between (1) those skills that are required for progressing academically from first grade onwards (i.e., reflecting school readiness) and (2) constraining method bias (i.e., applying Western assessment procedures to diverse refugee children). In none of the assessment domains language was critical for the instructions and task performance (except for the language domain itself). All scores obtained from assessments can be compared to norm data and/or converted to adjusted standard scores.

Language

A tablet-based version of the German adaptation of the Peabody Picture Vocabulary Test, 4th ed., was administered to measure receptive vocabulary skills in German (Dunn & Dunn, 2007). In this multiple-choice task, children have to tap on one of four pictures per trial based on an identifying word that is presented auditorily. The PPVT overall includes 228 sets of pictures with increasing difficulty. The number of correct responses is the raw score which can range between [0, 228].

Cognition

The "Object Assembly" subtest from the Wechsler Preschool and Primary Scale of Intelligence-III (Wechsler, 2002) was used to assess visual concept formation, which we considered as a proxy for the higher-order domain of nonverbal reasoning. That task requires understanding of part-whole relations and engagement in trial-and-error learning. The child has to assemble pieces into pictures with increasing difficulties (14 pictures at maximum within 90 seconds). The task yields raw scores based on the number of solved pictures and speed ranging between [0, 72].

Motor Skills

Subtests from the Intelligence and Development Scales were used to assess the functional domains of psychomotor development (gross motor skills, fine motor skills and visual-motor coordination; Grob et al., 2009). Gross motor skills were assessed in specific tasks including balancing, catching, throwing, and jumping. To assess fine motor skills, children strung beads on a thread under time pressure. To assess visual-motor coordination, children copied pictures of geometric figures by hand. For fine and gross motor subdomains, summarized raw scores reflecting performance can range between [0, 12] with higher values indicating better performance. For the visual motor subdomain, raw scores can range between [0, 16].

Socio-emotional Behavior

To assess children's behavior within the BPs, we used the German version of the 25-item Strength and Difficulties Questionnaire in teacher-report form (SDQ, 4–17 years; Goodman, 1997). One teacher of each preschool-based BP reported on the children of their respective group using a 3-point Likert scale

(0, not true; 1, somewhat true; 2, certainly true), referring to child behaviors in the past two weeks. The SDQ comprises four 5-item problem subscales (hyperactivity/inattention, conduct problems, peer-relationship problems, emotional problems), a total problem score, and a positively worded scale for prosocial behavior. We overall adhered to the guidelines for SDQ scoring (sdqinfo.org) and used norm data for descriptive analyses (Youthinmind, 2006). Scales were summed for interpretation; subscales sum scores can range between [0; 10], the total problem score between [0; 40] with higher values indicating more problem behavior. The SDQ norm data consist of teacher reports on a representative youth population sample including N = 4801 children aged 5–10 years (Meltzer et al., 2000). Psychometric investigations on the SDQ demonstrated moderate to good psychometric quality and cultural invariances for the subscales in diverse samples (Woerner et al., 2004). Internal consistency in our sample was at least acceptable for all subscales ($\alpha > .70$) and thus comparable to estimates found in other studies (Achenbach et al., 2008).

Early Education Quality of Bridging Projects

To assess quality of the preschool-based BPs, we used standardized observation tools: (1) The Classroom Assessment Scoring System Pre-K (CLASS; La Paro et al., 2002) to examine teacher–child interaction quality as a proxy for process quality; (2) a self-developed observation tool for structural program quality of BPs that refers to quality standards of center-based preschool programs in Germany. We compared observation results against a sample of diverse BPs. Details on the tools and procedures to assess quality of BPs, as also used in this study, are reported elsewhere (Busch et al., 2021).

Analytical Approaches

To examine children's development on group-level, we transposed assessment scores into age-referenced T-scores based on norm data (partially also gender-referenced, depending on task) and calculated the median T-scores per measure. We interpret median T-scores over 60 as reflecting above-average ability, in the range [40, 60] medium ability, [30, 40] low ability, and below 30 very-low ability. To examine children's socio-emotional problems, we calculated mean raw scores and standard deviations of the samples and, additionally, mean difference scores for the SDQ subscales and the total problem score against gender-adjusted norm data.

To investigate study hypotheses using inferential statistics, we used raw scores of the tasks and unadjusted sum scores of the SDQ subscales. We set alpha-error probability to $p < .05$ in one-sided testing and applied Bonferroni-Holm corrections. We report uncorrected p-values for all model estimates. We applied corrections on model estimates that address our study questions and indicate significance only for those p-values by an asterisk. In regression-based models, we therefore based our interpretations on standardized regression coefficients ($ß$). In pre-analyses, we inspected all variables for normality, outliers, missing values and examined regression diagnostics, respectively, based on pre-modeling. Rates of missing test scores for developmental tasks ranged from 27% to 34% at T_1 and 19–28% at T_2 per domain. Rates of missing values on SDQ subscale-level were 20% at T_1 and 32% at T_2. We maintained the assumption of a missing-completely-at-random pattern for outcome data at T_1, T_1 to T_2 and T_C data based on Little's test on missing value patterns (Little, 1988). We hence applied multiple-imputation strategies and pooled model parameters according to Rubin's rules (e.g., Li et al., 1991). We ran all analyses in R (version 3.5.0; R Core Team, 2014).

For inference-based analysis of hypothesis (1), we examined links between periods of BP attendance in months and developmental learning foundations with two approaches. In a cross-sectional approach, we applied univariate multiple regression (UMR) modeling on cross-sectional data at the initial assessment (T_1) on developmental tasks and SDQ scales as criteria and "length of previous BP attendance in months" as the focal predictor. We additionally considered children's gender, age, and the length of stay in Germany as first-entered covariates. In a longitudinal approach, we examined within-subject changes (i.e., T_1 to T_2) in refugee children attending BPs using a one-group pretest-

posttest design. We therefore tested intra-individual changes among our indicators using linear mixed effect (LME) models, for which we centered the individual changes within subjects. For analysis of hypothesis (2), we compared indicators of children's development in a quasi-experimental between-group design. Specifically, we compared refugee children from BPs close to their transition into first grade and refugee children attending first grade without previous attendance of ECD programs. As we expected systematic age-differences between samples, we selected an age-matched and equal-sized subsample of refugee children at T_1 assessments with refugee children from the comparison group (T_C) using propensity scores. In the propensity score calculations, we considered children's age only and applied the "nearest neighbor matching" procedure as implemented in the "MatchIt" package (Ho et al., 2011). We then calculated several UMR models to test whether children from BPs would show better developmental foundations for academic learning. We therefore examined whether the dummy-coded predictor "previous BP attendance" was linked to higher scores in developmental tasks and lower scores on the SDQ scales.

Results

Early Child Development Programs

The 10 preschool-based BPs that children of our study attended demonstrated low variability regarding dosage and early education quality. Professional early childhood teachers ran BPs located in elementary school facilities with a high frequency (i.e., 4–5 days weekly, M = 3.27 hours daily). Structured observations yielded high structural quality of BPs. We found good premises and equipment, session structuring, well-cooperating teams and sufficient educational materials. Most domains of structural quality were better than in other BPs which were not part of the local initiative under investigation. Structured field observations using the CLASS also showed high teacher–child interaction quality in preschool-based BPs. That is, teachers achieved a positive climate, demonstrated medium to high levels of sensitivity and stimulated language development to a moderate extent. Notably, the preschool-based BPs had higher levels of learning productivity when compared to regular BPs. For more details on early education quality see Appendices (A) and (B).

Norm Data Comparison: Development of Recently arrived Refugee Children Attending Bridging Projects

We initially compared assessment results for recently arrived refugee children attending BPs (at T_1) to norm data. The total score for socio-emotional problems was higher than the respective value from norm data indicating heightened overall symptom levels. Peer-interaction problems showed largest differences, followed by conduct problems, while emotional problems, hyperactivity/inattention, and prosocial behavior did not substantially differ. For the other developmental domains assessed individually (at T_1), norm-referenced assessment results yielded a mixed picture. While German receptive language and also nonverbal reasoning were, on average, on very-low levels, children demonstrated low-level performances in the visual motor skill task and low-to-medium-level performances in fine and gross motor skill tasks. For detailed results on the child assessments see Table 2.

Cross-sectional Analysis: Links of Children's Development with Their Previous Duration of BP Attendance

To prepare our cross-sectional analysis of hypothesis (1), we applied a predictive mean matching procedure using the *mice* package (Van Buuren & Groothuis-Oudshoorn, 2011). The efficacy of the imputation process was assessed via the fraction of missingness criterion (FMI; i.e., the amount of information lost due to missing data). FMI for most predictors were moderate, while for the variables "time since arrival in Germany" and "duration of BP attendance" this criterion was moderately large

Table 2. Mean, median, and difference scores for behavior problems and indicators of child development.

	T_1 sample, refugees from BPs				T_2 sample, refugees from BPs				T_C sample, refugees w/o early education			
	Raw Mean (SD)	*Δ Mean*	*Md T*	n_{T1}	*Raw Mean (SD)*	*Δ Mean*	*Md T*	n_{T2}	*Raw Mean (SD)*	*Δ Mean*	*Md T*	n_{TC}
Behavior problems												
Emotion problems	1.47 (1.87)	−0.15		127	1.70 (2.41)	0.22		59	2.54 (2.87)	1.04		39
Hyperactivity/Inattention	3.15 (2.59)	0.18		127	3.63 (2.62)	0.64		60	5.03 (2.81)	1.80		39
Conduct problems	1.37 (1.85)	0.50		127	1.68 (2.26)	0.75		59	2.28 (2.34)	1.30		39
Peer interaction problems	2.65 (2.21)	1.27		127	2.63 (2.28)	1.25		59	2.85 (2.13)	1.45		39
Total problems	8.63 (5.53)	1.75		127	9.70 (7.12)	2.88		59	12.69 (7.51)	5.55		39
Prosocial behavior[a]	7.15 (2.25)	−0.19		127	7.17 (2.62)	−0.16		58	5.36 (2.27)	−1.81		39
Child development												
Nonverbal reasoning	23.98 (15.18)		26.67	118	33.48 (18.86)		33.33	69	24.67 (9.46)		20	54
Receptive vocabulary	30.73 (21.69)		27.00	111	43.48 (23.95)		27	69	50.15 (22.61)		27	55
Gross motor skills	4.59 (2.04)		46.67	119	5.58 (2.27)		50	73	6.06 (2.08)		43.33	54
Fine motor skills	6.95 (2.52)		46.67	118	7.81 (2.28)		50	73	5.93 (2.26)		36.67	55
Visual motor coordination	4.72 (2.51)		40.00	117	5.59 (2.61)		41.67	73	6.32 (2.65)		36.42	54

Raw Mean (SD), mean raw score and standard deviation for behavior problems per SDQ scale, for child development per measure. *ΔMean*, gender-adjusted mean difference scores for SDQ subscales (direction: norm value subtracted from reported value; Youthinmind, 2006). *Md T*, median *T*-value based on gender- and age-adjusted norm data. *n*, number of cases per analysis

[a]Prosocial behavior scores were inversed, that is, higher values indicated more prosocial behavior

or large. Because the model estimates of the pre-models did not considerably change after exclusion of one potential outlier, we conducted our analyses based on the complete data set. For the influence of our covariates on the SDQ subscales, we found an influence of gender on the hyperactivity/inattention subscale ($\beta = 0.195$, $p = .003$). Regarding our focal predictor, a longer duration of BP attendance did not predict overall values on SDQ subscales. Only prediction of the subscale score for peer-interaction problems was on a trending level ($\beta = -0.184$, $p = .046$), yielding lower scores for a longer duration of BP attendance. For the child direct measures, we found the covariate age to positively influence raw scores for all assessments and, for predicting the fine motor scale, covariates gender ($\beta = 0.601$, $p = .002$; higher values for female gender) and time since arrival ($\beta = -0.327$, $p = .001$; shorter time since arrival linked to better scores) showed additional influence. A longer duration of BP attendance predicted better German receptive language skills only ($\beta = 0.257$, $p = .008$*). For detailed results see Table 3 and for full models Appendices (C) and (D).

Longitudinal Analysis: Changes in Children's Development Throughout 5 Months of BP Attendance

For pre-analyses of the longitudinal approach for hypothesis (1), we computed attrition analysis to examine longitudinal selection bias in participants of the repeated measurement subsample T_2. Using T_1 sample data, we therefore compared those refugee children who were considered in the repeated measure subsample assessments (T_2) and those who dropped out before ($n_{drop} = 79$, threshold at $p < .10$). Those refugee children who dropped out tended to show higher levels of "conduct problems" at T_1 ($t(120) = -1.90$, $p = .06$). They, however, did not differ regarding socio-demographic and migration-related characteristics (age at T_1, time since arrival in Germany, previous length of BP attendance, region of origin and gender). Note that several children of the T_1 sample transitioned to first grade shortly before the repeated assessment (T_2). We thus additionally compared children at T_2 assessments who previously transitioned into first grade to those children who remained in BPs (threshold at $p < .10$). Children who had transitioned were older, demonstrated better development regarding all domains, yet tended to have more emotional problems ($t(46) = 2$, $p = .05$). For

Table 3. Associations of Bridging Project attendance with behavior problems and indicators of child development.

	Cross-sectional analysis (at T_1) Predictor "previous length of BP attendance"				Longitudinal analysis (T_1 to T_2) Within-subject changes			Quasi-experimental design (T_{1m} vs. T_C) Between-group comparison		
Child outcome	*β*	*Std err*	p^b	ΔR^2	*β*	*Std err*	p^b	*β*	*Std err*	p^b
Behavior problems										
Emotion problems	−0.095	0.092	.302	0.011	0.371	0.223	.097	0.316	0.235	.182
Hyperactivity/Inattention	−0.037	0.096	.699	0.003	0.278	0.306	.365	0.631	0.244	**.011***
Conduct problems	0.007	0.010	.942	0.003	0.437	0.256	.088	0.378	0.236	.113
Peer-interaction problems	−0.184	0.091	.046	0.035	0.040	0.289	.889	0.093	0.236	.696
Total problems	−0.122	0.093	.192	0.017	1.128	0.771	.144	0.531	0.221	.018
Prosocial behavior[a]	0.044	0.091	.630	0.003	−0.022	0.294	.941	−0.706	0.231	**.003***
Child development										
Nonverbal reasoning	0.107	0.095	.266	0.015	8.257	1.599	**.001***	−0.702	0.226	**.003***
Receptive vocabulary	0.257	0.096	**.008***	0.070	9.777	1.923	**.001***	0.078	0.202	.701
Gross motor skills	−0.054	0.098	.581	0.007	0.921	0.196	**.001***	−0.045	0.218	.837
Fine motor skills	0.156	0.099	.120	0.010	0.529	0.243	.030	−0.663	0.212	**.002***
Vis motor coordination	−0.029	0.103	.781	0.005	0.700	0.227	**.001***	0.187	0.230	.418

Summarized results of the effects of "length of BP attendance" in regression analyses (study 1), the pre- post-effects in linear-mixed effect models (study 2) and effects of group-belongingness to BP attenders (=1) vs. non-attenders of early education (=0, study 3). See Tables C–F in the Appendix for detailed model reports.

β, pooled estimator for standardized regression weights; *Std err*, standard error for pooled estimator. *p*, *p*-value for pooled estimator (two-sided, no alpha-error correction); ΔR^2, change in adjusted R^2 by adding predictor "BP attendance."

[a]For the Prosocial Behavior scale higher values indicate more prosocial behavior

[b]Significant results (*) are highlighted in bold based on one-sided testing and Bonferroni-Holm corrections per study approach ($p < .05$*)

longitudinal analyses, we imputed data using the multilevel imputation procedure described by Grund et al. (2016). Imputation yielded that FMI for time effects varied from moderate to large across the models. We found that SDQ subscales scores did not show a significant decrease throughout the 5-month period. Notably, positive regression coefficients of all SDQ subscales even suggested tendencies of further increase (e.g., total problems score: $\beta = 1.128$, $p = .144$). Scoring on all developmental indicators increased from low levels at baseline throughout five months of attendance (note that p-value of the fine motor skills indicator, $\beta = 0.529$, $p = .030$, became non-significant after alpha-error correction). Compared to norm data, however, especially performance in the nonverbal reasoning task (Median $T = 33.33$) and German receptive vocabulary task (Median $T = 27.00$) were still on low to very-low levels. For details on the norm comparisons at T_2 see Table 2, on the within-subject T_1-T_2 changes see Table 3.

Group Comparison with Refugee Children without ECD Program Attendance

For hypothesis (2), we conducted a between-group comparison with children from the T_C sample. After propensity-score matching, the subsample of T_1 was on average $M = 78.49$ months old ($SD = 5.05$), had attended BPs for $M = 6.1$ months ($SD = 4.43$), had been in Germany for $M = 27.19$ months ($SD = 20.75$) and was gender-balanced (51% female). Most children of the age-matched T_{1m} subsample transitioned into first grade within 5 months after assessments (90.6%). Since both groups still differed regarding their age- and gender-distribution after matching, we added both as covariates in our UMR models with the dummy-coded focal predictor "BP attendance vs. no program attendance." We applied missing data imputation analogously to study 1. For most predictors, FMIs were low to moderate. We found moderate to large FMIs for the predictors "time since arrival in Germany" in all models and for "gender" in models on indicators for development. For our covariates, we found male gender to be linked to more conduct problems ($\beta = -0.412$, $p = .043$) and female gender to more prosocial behavior ($\beta = 0.482$, $p = .019$). Regarding the focal predictor for hypothesis (2), refugee children with BP attendance demonstrated less hyperactivity/inattention ($\beta = 0.631$, $p = .011$*) and more prosocial behavior ($\beta = -0.706$, $p = .003$*) compared to refugee children without previous ECD program attendance. For child direct assessment measures, older age ($\beta = 0.058$, $p = .001$) and longer time since arrival ($\beta = 0.231$, $p = .034$) were linked to better German receptive language; female gender ($\beta = 0.579$, $p = .005$) and longer time since arrival ($\beta = 0.108$, $p = .048$) to better fine motor skills, and higher age ($\beta = 0.049$, $p = .002$) to better gross motor skills. Regarding the focal predictor for hypothesis (2), children attending BPs showed better performances in tasks on nonverbal reasoning ($\beta = -0.702$, $p = .003$*) and fine motor skills ($\beta = -0.663$, $p = .002$*). See Table 3 for detailed results and for full models Appendices (E) and (F).

Discussion

We investigated developmental foundations for academic learning in recently arrived refugee children and links to children's attendance of preschool-based ECD programs (BPs) in Germany. We examined child outcomes among refugee children attending BPs, tested links to program attendance in cross-sectional, and longitudinal designs. Further, we compared those from BPs to other refugee children without any ECD program attendance at the transitioning to first grade. Overall, our findings suggest that recently arrived refugee children demonstrate low levels of developmental foundations. We observed inconsistent evidence on the links between refugee children's attendance of BPs with child outcomes.

Assessment of the Developmental Foundations for Academic Learning

We observed higher levels of socio-emotional behavior problems and low to very-low level performances in assessments on cognitive skills, German language, and partly also motor skills. Our findings support previous evidence and are consistent with our initial expectations that recently arrived refugee children are at risk of not achieving the expected developmental foundations before transitioning into first grade (Bouchane et al., 2018). While the developmental indicators we examined the established predictors of academic learning, the use of standardized and norm-based assessments with recently arrived refugee children need to be discussed. That is, test procedures and also reference data reflect the normative experiences of children from Western populations. The extent of refugee children's deficits could thus be overestimated due to method bias (Van de Vijver & Tanzer, 2004). Initially lacking early education experiences of Western contexts, refugee children might need time to familiarize themselves with the materials and activities used in the assessment procedures as these adopt common practices and materials from such contexts. Large gains in child direct assessments found in the longitudinal approach (i.e., repeated measurement) support this notion. Second, we only assessed German receptive language as the host countries' language is essential for academic learning after resettlement. Although host-country language skills are essential for academic success, they might only constitute a fraction of refugee children's overall linguistic abilities. Children's family language skills are a resource that could also be related to later academic achievement via transfer effects (Prevoo et al., 2016). Given such caveats, our findings still inform on a set of developmental learning foundations assessed as relative to their non-refugee peers, who will become their classmates in first grade.

Our findings revealed high levels of socio-emotional problems that are overall consistent with previous evidence on recently arrived and preschool-aged refugee children in high-income countries (Almqvist & Broberg, 1999; Buchmüller et al., 2018). Our findings, however, more strongly suggest externalizing behavior and peer-interaction problems. Such differences could be due to observer bias across studies. While Buchmüller et al. (2018) surveyed refugee children's parents and their preschool teachers, we exclusively focused on teacher assessments. As teachers experience child behavior in groups within education settings, they tend to focus on children's externalizing behaviors (Achenbach & Rescorla, 2001). Additionally, our findings reflect young children's situation in center-based programs during post-migration periods. As we found highest scores for peer-interaction problems, especially social situations with diverse peers might challenge refugee children within education settings. This context-related effect is also supported by other studies, similarly reporting distinct peer-interaction problems in young refugee children when attending kindergarten or center-based early education programs in Germany (Buchmüller et al., 2020; Chwastek et al., 2021).

Links of the Assessment Results to Bridging Project Attendance: Cognitive, Motor, and Language Development

We hypothesized links between BP attendance and child development. With regard to nonverbal reasoning, motor skills and German language skills, findings inconsistently supported our expectations that BP attendance would link to better development. In the cross-sectional analysis on time of previous BP attendance, we only found links to improvements in German language development. In the longitudinal design, however, we found that cognitive, language, and motor domains of development (except for the fine motor skills subscale) improved after 5 additional months of BP attendance. Those findings are consistent to previous evidence on promoting effects of preschool attendance for non-refugee populations (Weiland & Yoshikawa, 2013; Winsler et al., 2008). Similarly, such findings also correspond to emerging evidence on ECD interventions specific for refugee children including play or preschool activities in humanitarian contexts (e.g., Metzler et al., 2019; Rousseau et al., 2009). However, how BPs could promote refugee children's development has yet to be substantiated in more detailed research. For general ECD programs in high-income countries, better program quality could

be linked to improvements in children's pre-academic skills (Burchinal et al., 2000) – although evidence is inconsistent (Kohl et al., 2020). As we found high program quality among the BPs under investigation, especially high process quality with a focus on host-country language development could have contributed to the positive changes observed. As the variability across the program quality domains was limited, we were not able to further explore the impact of quality variations on developmental domains.

Notably, German language was the only domain with consistent links to BP program attendance among cross-sectional *and* longitudinal analyses. Effects of BP attendance on language development could be strongest for two reasons. First, as second language learners are especially sensitive to language exposure during early development, young refugee children likely benefit from German language exposure in BPs. Second and related to this, BPs facilitate language immersion as German is the connecting language between diverse refugee children and their teachers. While cognitive and motor domains might be stimulated in other contexts as well, the BPs were likely to provide a major German language context for young refugee children.

In our longitudinal research design, it is difficult to disentangle development-stimulating effects of BP attendance from maturation. Having used raw scores from child direct assessments, the positive changes observed might, at least partially, reflect maturation. That notion is supported by our cross-sectional approach where we found influence of age (added as a covariate) on raw scores of language, cognition, and motor tasks. However, findings from the between-group comparison with refugee children without ECD program experience attending first grade more strongly supported our initial expectations. While group differences in the fine motor and nonverbal reasoning task reached significance, all domains yielded on average better raw scores for those children attending BPs (note that child age was also added as a covariate). Still, those raw score comparisons could have even underestimated true differences because the T_C sample was on average older. We, however, cannot preclude that those refugee families who had previously placed their children into BPs might systematically differ from other refugee families whose children transitioned into first grade without any ECD program attendance (e.g., regarding families' educational aspirations).

Links of the Assessment Results to Bridging Project Attendance: Socio-Emotional Behavior

We further hypothesized that BP attendance would generally be linked to less behavior problems in refugee children. The cross-sectional approach yielded null-effects and the longitudinal suggested even increasing levels of behavior problems. While those findings overall contradicted our initial expectations, such inconsistencies among our study results could refer to the different methodological approaches. In the longitudinal approach, first, a distinctive increase of behavior problems was possibly undetected, as attrition analyses yielded a higher likelihood to dropout before T_2 for those children who exhibited higher levels of conduct problems at T_1. Second, several refugee children were assessed around their transition into first grade. Here, the longitudinal approach (T_1-T_2) was potentially not able to disentangle distressing effects of transitioning into first grade (e.g., behavioral adjustment to new classroom settings) from promoting effects of BP attendance (i.e., continuous BP attendance contributing to decreased behavior problems).

Longitudinal and cross-sectional analyses along with norm comparisons jointly support that refugee children's externalizing behavior problems could persist in the short term of BP attendance. We found that refugee children's behavior problems persisted after BP enrollment and tended to unfold even further throughout 5 additional months of program attendance. As the BPs offered structured learning activities in preschool settings, such behavior problems could reflect children's socio-emotional adjustments to education environments. Picchio and Mayer (2019) correspondingly described refugee children's enrollment in early education as a challenging "double transition," that is, both settling in a new country and subsequently attending context-specific education services. Beyond the academic learning goals, refugee children need to establish positive relationships with German teachers and diverse peers within early education while regulating their own needs and emotions more

independently (Busch et al., 2018). Adding to that, accumulating stressful experiences outside the BPs during resettlement periods could also have challenged children's socio-emotional adjustment (Almqvist & Broberg, 1999; Montgomery, 2008). Alternatively, increasing levels of externalizing problems reported by teachers might reflect an observation bias. BP teachers get to know refugee children better over time and more likely recognize socio-emotional distress. Findings by Chwastek et al. (2021) support this notion albeit with a different explanation. They found that preschool teachers who had longer work experience with refugee children were more likely to have negative stereotypes, which in turn were linked to reporting more externalizing behavior problems among refugee children.

Amid inconsistent evidence from the longitudinal and cross-sectional approaches, findings comparing refugee children attending first grade with and without previous BP attendance revealed lower levels of hyperactivity/inattention and more prosocial behavior for children with previous BP attendance. These findings more strongly suggest the promoting effects of BP attendance on socio-emotional behavior and, also, that such effects could become first observable during transitioning points (i.e., elementary school enrollment). Refugee children who have attended a BP are more likely to be familiar with school settings. Therefore, previous attendance of BPs could distinctively facilitate transitioning, hence buffering a disproportional increase of externalizing and peer-interaction problems among refugee children. In sum, linking recently arrived refugee children's ECD program attendance immediately to decreasing levels of behavior problems could be a too simplified model. Instead, future research should consider effects of transitioning, especially when refugee children are studied within education settings.

Limitations, Strengths, and Future Research

Using the multiple approaches and samples enables different perspectives on our research objectives. However, these approaches also have notable caveats that warrant starting points for future research. Cross-sectional and group comparisons cannot preclude sample selection bias thus impairing external validity. Although the longitudinal approach could better address such limitations, effects of repeated assessments and longitudinal selection could bias results. Beyond, causal inferences are limited by an overall non-experimental research approach. Future studies, at best randomized controlled trials, need to identify those determinants that explain progress in the developmental learning foundations of refugee children during resettlement periods. Specifically, research should address (A) whether recently arrived young refugee children experience developmental gains from program attendance and (B) identify specific program components that distinctively drive ECD program impact on this specific group. Previous findings by Gagné et al. (2018) suggest that fostering communication or socio-emotional functioning differently shape refugee children's academic trajectories. Our investigation moreover focused mainly on children's short-term trajectories within early childhood until transitioning into first grade. While the beneficial effects of ECD attendance might become first visible at transitioning points, longer-term trajectories require further attention. Aghajafari et al. (2020) found that academic outcomes of refugee children aged between 5 and 12 years improved during school attendance while differential trajectories of refugee children who attended early education services (versus others who did not) are still unknown.

Policy Relevance and Conclusion

Our findings reiterate and substantiate previous claims after promoting refugee children's ECD for positive education trajectories (Park et al., 2018; Sirin & Rogers-Sirin, 2015). Designed context-specific by local stakeholders based on a statewide ECD policy, preschool-based BPs therefore blend specialized ECD programs with center-based early education services to prepare children for their transition into first grade. Situated at such a conceptual borderline and given our findings, one must critically reflect on the relations between BPs and regular ECD services in Germany. While the intention to establish BPs was compensatory in nature (i.e., tackling shortage in early education services amid a high demand by refugee families), BPs are overall less formalized, have heterogeneous quality

standards, and on average lower resources compared to regular state-funded German preschools or kindergartens. Apart from the potential benefits of establishing BPs during years with peaking demands, stakeholders should generally work toward a timely inclusion of refugee children into regular early education services and thereby counteract the emergence of parallel early education services for refugee- and non-refugee children.

Overall, our study moves research empirically forward in two regards. First, findings emphasize the importance of specifically addressing the needs of recently arrived refugee children, as they are at risk for low developmental learning foundations. Second, findings provide first evidence that flexibly organized preschool-based ECD programs could support refugee children's successful transition into first grade.

Acknowledgments

We gratefully thank Sandy Chwastek and Felix Schideck for their valuable contributions and also thank all participating students, teachers, and schools who have supported this study.

Data Availability Statement

The dataset supporting the conclusions of this article is available from the corresponding author or Birgit Leyendecker on reasonable request.

Disclosure Statement

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

Funding

This study was supported by the educational trust RuhrFutur (Mercator Foundation) through a grant to Birgit Leyendecker.

ORCID

J. Busch http://orcid.org/0000-0003-4626-7045
N. Cabrera http://orcid.org/0000-0002-1558-7356
F. Ialuna http://orcid.org/0000-0002-2159-7198
T. Buchmüller http://orcid.org/0000-0002-8002-5608
B. Leyendecker http://orcid.org/0000-0003-1301-8069

References

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H.-C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*(3), 251–275. https://doi.org/10.1111/j.1469-7610.2007.01867.x

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. University of Vermont, Research Center for Children, Youth, and Families.

Aghajafari, F., Pianorosa, E., Premji, Z., Souri, S., & Dewey, D. (2020). Academic achievement and psychosocial adjustment in child refugees: A systematic review. *Journal of Traumatic Stress*, *33*(6), 908–916. https://doi.org/10.1002/jts.22582

Almqvist, K., & Broberg, A. G. (1999). Mental health and social adjustment in young refugee children 3.5 years after their arrival in Sweden. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*(6), 723–730. https://doi.org/10.1097/00004583-199906000-00020

Bailey, A. L., & Butler, F. A. (2007). A conceptual framework of academic English language for broad application to education. In A. L. Bailey (Ed.), *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.

Bouchane, K., Yoshikawa, H., Murphy, K. M., & Lombardi, J. (2018). *Early childhood development and early learning for children in crisis and conflict*. Paper commissioned for the 2019 GEM Report, Migration, displacement and education: Building bridges, not walls. UNESCO.

Buchmüller, T., Lembcke, H., Busch, J., Kumsta, R., & Leyendecker, B. (2018). Exploring mental health status and syndrome patterns among young refugee children in Germany. *Frontiers in Psychiatry*, *9*, 212. https://doi.org/10.3389/fpsyt.2018.00212

Buchmüller, T., Lembcke, H., Ialuna, F., Busch, J., & Leyendecker, B. (2020). Mental health needs of refugee children in specialized early education and care programs in Germany. *Journal of Immigrant and Minority Health, 22*(1), 22–33. https://doi.org/10.1007/s10903-019-00896-4

Burchinal, M. R., Roberts, J. E., Riggins, R., Jr., Zeisel, S. A., Neebe, E., & Bryant, D. (2000). Relating quality of center-based child care to early cognitive and language development longitudinally. *Child Development*, *71*(2), 338–357. https://doi.org/10.1111/1467-8624.00149

Burrage, M. S., Ponitz, C. C., McCready, E. A., Shah, P., Sims, B. C., Jewkes, A. M., & Morrison, F. J. (2008). Age- and schooling-related effects on executive functions in young children: A natural experiment. *Child Neuropsychology*, *14*(6), 510–524. https://doi.org/10.1080/09297040701756917

Busch, J., & Leyendecker, B. (2019). Socialization and development of refugee children: Chances of childcare. In T. Tulviste, D. L. Best, & J. L. Gibbons (Eds.), *Children's social worlds in cultural context* (pp. 187–200). Springer International Publishing. https://doi.org/10.1007/978-3-030-27033-9_14

Busch, J., Bihler, L.-M., Lembcke, H., Buchmüller, T., Diers, K., & Leyendecker, B. (2018). Challenges and solutions perceived by educators in an early childcare program for refugee children. *Frontiers in Psychology*, *9*, 1612. https://doi.org/10.3389/fpsyg.2018.01621

Busch, J., Buchmüller, T., & Leyendecker, B. (2021). *An observational study on the implementation and quality of a policy-based early childhood education program for newly arrived refugee children in Germany*. ResearchSquare. https://doi.org/10.21203/rs.3.rs-587074/v1

Carlson, A. G., Rowe, E., & Curby, T. W. (2013). Disentangling fine motor skills' relations to academic achievement: The relative contributions of visual-spatial integration and visual-motor coordination. *The Journal of Genetic Psychology*, *174*(5), 514–533. https://doi.org/10.1080/00221325.2012.717122

Chwastek, S., Leyendecker, B., Heithausen, A., Ballero Reque, C., & Busch, J. (2021). Pre-school teachers' stereotypes and self-efficacy are linked to perceptions of behavior problems in newly arrived refugee children. *Frontiers in Psychiatry*, *11*, 574412. https://doi.org/10.3389/fpsyt.2020.574412

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). NCS Pearson.

Dybdahl, R. (2001). Children and mothers in war: An outcome study of a psychosocial intervention program. *Child Development*, *72*(4), 1214–1230. https://doi.org/10.1111/1467-8624.00343

Gagné, M., Janus, M., Milbrath, C., Gadermann, A., & Guhn, M. (2018). Early emotional and communication functioning predicting the academic trajectories of refugee children in Canada. *Educational Psychology*, *38*(8), 1050–1067. https://doi.org/10.1080/01443410.2018.1475627

Gambaro, L., Liebau, E., Peter, F. H., & Weinhardt, F. (2017). Viele Kinder von Geflüchteten besuchen eine Kita oder Grundschule: Nachholbedarf bei den unter Dreijährigen und der Sprachförderung von Schulkindern [Refugee children attending pre- and primary schools]. *German Institute for Economic Research – Weekly Report*, *84*(19), 379–386. https://www.diw.de/documents/publikationen/73/diw_01.c.557905.de/17-19-1.pdf

German Federal Agency for Migration and Refugees. (2020). *Aktuelle Zahlen zu Asyl-Ausgabe: Juni 2020*. https://www.bamf.de/SharedDocs/Anlagen/DE/Statistik/AsylinZahlen/aktuelle-zahlen-juni-2020.pdf?__blob=publicationFile&v=4

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Graham, H. R., Minhas, R. S., & Paxton, G. (2016). Learning problems in children of refugee background: A systematic review. *Pediatrics*, *137*(6), e20153994–e20153994. https://doi.org/10.1542/peds.2015-3994

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2009). *Intelligence and development scales (IDS). Intelligenz- und Entwicklungsskalen*. Hans Huber.

Grund, S., Lüdtke, O., & Robitzsch, A. (2016). Multiple imputation of multilevel missing data: An introduction to the R package pan. *SAGE Open*, *6*(4), 215824401666822. https://doi.org/10.1177/2158244016668220

Halle, T., Hair, E., Wandner, L., McNamara, M., & Chien, N. (2012). Predictors and outcomes of early versus later English language proficiency among English language learners. *Early Childhood Research Quarterly*, *27*(1), 1–20. https://doi.org/10.1016/j.ecresq.2011.07.004

High, P. C. (2008). School Readiness. *Pediatrics*, *121*(4), e1008–e1015. https://doi.org/10.1542/peds.2008-0079

Ho, D., Imai, K., King, G., & Stuart, E. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8), 1–28. https://doi.org/10.18637/jss.v042.i08

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, *23*(1), 27–50. https://doi.org/10.1016/j.ecresq.2007.05.002

Kien, C., Sommer, I., Faustmann, A., Gibson, L., Schneider, M., Krczal, E., Jank, R., Klerings, I., Szelag, M., Kerschner, B., Brattström, P., & Gartlehner, G. (2019). Prevalence of mental disorders in young refugees and asylum seekers in European Countries: A systematic review. *European Child & Adolescent Psychiatry, 28*(10), 1295–1310. https://doi.org/10.1007/s00787-018-1215-z

Kohl, K., Bihler, L.-M., Willard, J. A., Agache, A., & Leyendecker, B. (2020). Linking quantity and quality of early childhood education and care to children's socio-emotional adjustment: A German cross-sectional study. *Early Education and Development*, *31*(2), 177–199. https://doi.org/10.1080/10409289.2019.1650543

La Paro, K., Pianta, R. C., Hamre, B., & Stuhlman, M. (2002). *Classroom assessment scoring system (CLASS). Pre-K version*. University of Virginia.

Ladd, G. W., Birch, S. H., & Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten: Related spheres of influence? *Child Development*, *70*(6), 1373–1400. https://doi.org/10.1111/1467-8624.00101

Lakin, J. M. (2012). Assessing the cognitive abilities of culturally and linguistically diverse students: Predictive validity of verbal, quantitative, and nonverbal tests. *Psychology in the Schools*, *49*(8), 756–768. https://doi.org/10.1002/pits.21630

Li, K. H., Raghunathan, T. E., & Rubin, D. B. (1991). Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution. *Journal of the American Statistical Association*, *86*(416), 1065–1073. https://doi.org/10.2307/2290525

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

Meltzer, H., Gatward, R., Goodman, R., & Ford, F. (2000). *Mental health of children and adolescents in Great Britain*. The Stationery Office.

Metzler, J., Diaconu, K., Hermosilla, S., Kaijuka, R., Ebulu, G., Savage, K., & Ager, A. (2019). Short- and longer-term impacts of child friendly space interventions in Rwamwanja refugee settlement, Uganda. *Journal of Child Psychology and Psychiatry*, *60*(11), 1152–1163. https://doi.org/10.1111/jcpp.13069

Montgomery, E. (2008). Long-term effects of organized violence on young Middle Eastern refugees' mental health. *Social Science and Medicine*, *67*(10), 1596–1603. https://doi.org/10.1016/j.socscimed.2008.07.020

Morantz, G., Rousseau, C., Banerji, A., Martin, C., & Heymann, J. (2013). Resettlement challenges faced by refugee claimant families in Montreal: Lack of access to child care. *Child and Family Social Work*, *18*(3), 318–328. https://doi.org/10.1111/j.1365-2206.2012.00848.x

Murphy, K. M., Yoshikawa, H., & Wuermli, A. J. (2018). Implementation research for early childhood development programming in humanitarian contexts: ECD implementation research-humanitarian settings. *Annals of the New York Academy of Sciences*, *1419*(1), 90–101. https://doi.org/10.1111/nyas.13691

Murrah, W. M. (2010). *Comparing self-regulatory and early academic skills as predictors of later math, reading, and science elementary school achievement* [Dissertation abstracts International Section A: Humanities and Social Sciences]. ProQuest Information & Learning.

Park, M., Katsiaficas, C. M., & McHugh, M. (2018). *Responding to the ECEC needs of children of refugees and Asylum Seekers in Europe and North America*. Migration Policy Institute.

Picchio, M., & Mayer, S. (2019). Transitions in ECEC services: The experience of children from migrant families. *European Early Childhood Education Research Journal*, *27*(2), 285–296. https://doi.org/10.1080/1350293X.2019.1579552

Prevoo, M. J., Malda, M., Mesman, J., & van IJzendoorn, M. H. (2016). Within-and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: A meta-analytical study. *Review of Educational Research*, *86*(1), 237–276. https://doi.org/10.3102/0034654315584685

R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science*, *4*(1), 2–14. https://doi.org/10.1207/S1532480XADS0401_1

Reed, R. V., Fazel, M., Jones, L., Panter-Brick, C., & Stein, A. (2012). Mental health of displaced and refugee children resettled in low-income and middle-income countries: Risk and protective factors. *The Lancet*, *379*(9812), 250–265.

Rousseau, C., Benoit, M., Lacroix, L., & Gauthier, M.-F. (2009). Evaluation of a sand-play program for preschoolers in a multiethnic neighborhood. *Journal of Child Psychology and Psychiatry*, *50*(6), 743–750. https://doi.org/10.1111/j.1469-7610.2008.02003.x

Rousseau, C., Drapeau, A., & Corin, E. (1996). School performance and emotional problems in refugee children. *American Journal of Orthopsychiatry*, *66*(2), 239–251. https://doi.org/10.1037/h0080175

Sabol, T. J., & Pianta, R. C. (2012). Patterns of school readiness forecast achievement and socioemotional development at the end of elementary school. *Child Development*, *83*(1), 282–299. https://doi.org/10.1111/j.1467-8624.2011.01678.x

Sirin, S. R., & Rogers-Sirin, L. (2015). *The educational and mental health needs of Syrian refugee children*. Migration Policy Institute.

United Nations International Children's Emergency Fund. (2019). *UNICEF global databases- child migrants and refugees*. https://data.unicef.org/wp-content/uploads/2017/01/Child-migrants-and-refugees_Dec2019.xlsx

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, *54*(2), 119–135. https://doi.org/10.1016/j.erap.2003.12.004

Walpole, S. C., Abbara, A., Gunst, M., & Harkensee, C. (2018). Cross-sectional growth assessment of children in four refugee camps in Northern Greece. *Public Health*, *162*(1), 147–152. https://doi.org/10.1016/j.puhe.2018.05.004

Wechsler, D. (2002). *The Wechsler preschool and primary scale of intelligence, Third Edition (WPPSI-III)*. The Psychological Corporation.

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, *84*(6), 2112–2130. https://doi.org/10.1111/cdev.12099

Wilson, A., Piek, J. P., & Kane, R. (2013). The mediating role of social skills in the relationship between motor ability and internalizing symptoms in pre-primary children. *Infant and Child Development*, *22*(2), 151–164. https://doi.org/10.1002/icd.1773

Winsler, A., Tran, H., Hartman, S. C., Madigan, A. L., Manfra, L., & Bleiker, C. (2008). School readiness gains made by ethnically diverse children in poverty attending center-based childcare and public school pre-kindergarten programs. *Early Childhood Research Quarterly*, *23*(3), 314–329. https://doi.org/10.1016/j.ecresq.2008.02.003

Woerner, W., Fleitlich-Bilyk, B., Martinussen, R., Fletcher, J., Cucchiaro, G., Dalgalarrondo, P., & Tannock, R. (2004). The strengths and difficulties questionnaire overseas: Evaluations and applications of the SDQ beyond Europe. *European Child and Adolescent Psychiatry*, *13*(2), 47–54. https://doi.org/10.1007/s00787-004-2008-0

Yousafzai, A. K., Rasheed, M. A., Rizvi, A., Armstrong, R., & Bhutta, Z. A. (2014). Effect of integrated responsive stimulation and nutrition interventions in the lady health worker programme in Pakistan on child development, growth, and health outcomes: A cluster-randomised factorial effectiveness trial. *The Lancet*, *384*(9950), 1282–1293. https://doi.org/10.1016/S0140-6736(14)60455-4

Youthinmind. (2006). *British norms for the strengths and difficulties questionnaire*. http://sdqinfo.org/norms/UKNorm4.pdf

Appendices

Appendix A

Table A. Structural quality of preschool-based Bridging Projects according to the "Bridging Project Evaluation Scale."

	M (*SD*)	
Observed structural quality per dimension	Preschool BPs	Various BPs
Premises *Structural aspects of the setting such as availability of sufficient space for activities, areas for relaxation, or sanitary facilities*	1.48 (0.22)	1.45 (0.24)
Equipment *Availability and condition of the movable interior and its suitability for preschool-aged children*	1.48 (0.23)	1.57 (0.35)
Structuring of a Session *Formal structure of the program, e.g., clearly indicated start and ending times, establishment of rituals, rules, and routines*	1.70 (0.43)	1.18 (0.51)
Team Coherence *Characteristics of team climate and the degree of effective cooperation among staff*	1.86 (0.23)	1.69 (0.43)
Educational materials *For pre-academic activities and play, as well as for language facilitation in multilingual groups*	1.67 (0.27)	1.42 (0.41)

Mean ratings (*M*) and standard deviations (*SD*) for structured observations using the "Bridging Project Evaluation Scale" (BREVIS) in the preschool-based BPs of our sample ($N = 10$) and other BPs with various implementation strategies ($N = 22$). Only aggregated values on domain level are reported. BREVIS overall consists of 24 quality indicators to be ranked on a 3-point Likert scale (*Inadequate* = 0, *Moderate* = 1, *Excellent* = 2). All observers were trained in using the BREVIS and conducted observations in all BPs. Corresponding author provides the full observation protocol on request.

Appendix B

Table B. Teacher–child interaction quality of preschool-based Bridging Projects according to the "Classroom Assessment Scoring System Pre-K."

	M (*SD*)	
Observed interaction quality per dimension	Preschool-based BPs	Various BPs
Positive climate	6.38 (0.63)	6.18 (0.89)
Negative climate	6.93 (0.11)	6.84 (0.28)
Teacher sensitivity	5.93 (0.68)	5.64 (1.05)
Behavior management	6.28 (0.62)	5.66 (1.08)
Productivity	6.15 (0.70)	5.21 (1.35)
Language Modeling	3.68 (1.26)	3.61 (1.32)

Mean ratings (*M*) and standard deviations (*SD*) for structured observations using the "Classroom Assessment Scoring System Pre-K" (CLASS; La Paro et al., 2002) in the preschool-based BPs of our sample ($N = 10$) and other BPs with various implementation strategies ($N = 22$). CLASS dimensions were rated on a 3-point Likert scale (lowest = 1, medium = 4, highest quality = 7). All observers were licensed by Teachstone® at the time of assessments and proved reliability in the official CLASS trainings.

Appendix C

Table C1. Cross-sectional approach, univariate multiple regression models predicting behavior problems.

	Emotion Problems			Hyperactivity/ Inattention			Conduct Problems		
	β	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*
Model 1									
Intercept	−0.392	0.598	.514	−0.095	0.632	.881	0.442	0.601	.471
Age	0.004	0.009	.685	0.006	0.009	.527	−0.006	0.009	.530
Gender	0.281	0.188	.138	−0.574	0.193	.004	−0.145	0.180	.425
Time since arrival	0.028	0.109	.798	0.039	0.105	.710	0.102	0.097	.291
Model 2									
Intercept	−0.396	0.598	.509	−0.097	0.634	.879	0.444	0.611	.469
Age	0.004	0.009	.661	0.006	0.009	.520	−0.006	0.009	.527
Gender	0.253	0.189	.182	−0.585	0.195	.003	−0.142	0.183	.441
Time since arrival	0.041	0.109	.710	0.044	0.107	.679	0.102	0.098	.300
Length BP attendance	−0.095	0.092	.302	−0.037	0.096	.700	0.007	0.010	.942
Model Comparisons									
Base vs. Model 1	F (3, 143) = 0.843, p = .473, R^2 = 0.042			F (3, 143) = 2.889, p = .038, R^2 = 0.086			F (3, 143) = 0.676, p = 568, R^2 = 0.023		
Model 1 vs. Model 2	F (1, 142) = 1.074, p = .302, ΔR^2 = 0.011			F (1, 142) = 0.150, p = .699, ΔR^2 = 0.003			F (1, 142) = 0.005, p = .942, ΔR^2 = 0.003		

Univariate regression models predicting subscales of the Strengths and Difficulties Questionnaire on T_1 sample (N = 152). β, standardized beta-coefficients for predictors. *SE*, standard error for standardized beta-coefficient. p, two-sided alpha-error probability. ΔR^2, changes in adjusted R^2 for intercept-only model (base), model 1 (without predictor "length of BP attendance") and model 2 (with "length of BP attendance").

Table C2. Cross-sectional approach, univariate multiple-regression models predicting behavior problems.

	Peer-interaction problems			Total problems			Prosocial behavior		
	β	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*
Model 1									
Intercept	0.030	0.595	.959	−0.023	0.597	.961	−0.756	0.611	.219
Age	−0.003	0.008	.744	0.001	0.009	.893	0.009	0.009	.307
Gender	0.285	0.180	.116	−0.113	0.184	.540	0.264	0.178	.142
Time since arrival	0.001	0.099	.996	0.063	0.103	.541	−0.044	0.105	.675
Model 2									
Intercept	0.022	0.586	.970	−0.035	0.596	.953	−0.757	0.612	.219
Age	−0.002	0.008	.788	0.002	0.009	.861	0.009	0.009	.313
Gender	0.232	0.178	.197	−0.148	0.184	.422	0.277	0.181	.128
Time since arrival	0.252	0.099	.798	0.080	0.104	.444	−0.050	0.106	.638
Length BP attendance	−0.184	0.091	.046	−0.122	0.093	.192	0.044	0.091	.630
Model Comparisons									
Base vs. Model 1	F (3, 143) = 0.830, p = .479, R^2 = 0.061			F (3, 143) = 0.235, p = .872, R^2 = 0.028			F (3, 143) = 1.003, p = .394, R^2 = 0.036		
Model 1 vs. Model 2	F (1, 142) = 4.081, p = .045, ΔR^2 = 0.035			F (1, 142) = 1.31, p = 1.725, ΔR^2 = 0.017			F (1, 142) = 0.233, p = .630, ΔR^2 = 0.003		

See Table C1

[a]Scale on prosocial behavior is reversed compared to problem scales.

Appendix D

Table D1. Cross-sectional approach, univariate multiple-regression models predicting indicators for childhood development.

	Nonverbal Reasoning			German receptive vocabulary			Gross motor skills		
	β	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*
Model 1									
Intercept	−2.741	0.590	.001	−2.318	0.594	.001	−3.089	0.593	.001
Age	0.040	0.008	.001	0.034	0.009	.001	0.044	0.008	.001
Gender	−0.160	0.179	.374	−0.091	0.183	.622	0.056	0.180	.756
Time since arrival	−0.059	0.094	.533	0.159	0.106	.137	−0.150	0.101	.141
Model 2									
Intercept	−2.734	0.591	.001	−2.301	0.587	.001	−3.097	0.595	.001
Age	0.040	0.008	.001	0.033	0.008	.001	0.044	0.008	.001
Gender	−0.128	0.180	.478	−0.016	0.181	.928	0.041	0.182	.823
Time since arrival	−0.074	0.095	.442	0.125	0.104	.235	−0.144	0.103	.164
Length BP attendance	0.107	0.095	.266	0.257	0.096	.008	−0.054	0.098	.581
Model Comparisons									
Base vs. Model 1	F (3, 143) = 7.410, p = .001, R^2 = 0.205			F (3, 143) = 5.273, p = .002, R^2 = 0.225			F (3, 143) = 8.656, p = .001, R^2 = 0.237		
Model 1 vs. Model 2	F (1, 142) = 1.251, p = .266, ΔR^2 = 0.015			F (1, 142) = 7.226, p = .008, ΔR^2 = 0.070			F (1, 142) = 0.307, p = .581, ΔR^2 = 0.007		

Univariate regression models predicting assessment scores from performances in developmental tasks on T_1 sample (N = 152). *β*, standardized beta-coefficients for predictors. *SE*, standard error for standardized beta-coefficient. *p*, uncorrected two-sided alpha-error probability. ΔR^2, changes in adjusted R^2 for intercept-only model (base), model 1 (without predictor "length of BP attendance") and model 2 (with "length of BP attendance").

Table D2. Cross-sectional approach, univariate multiple-regression models predicting indicators for childhood development.

	Fine motor skills			Visual motor skills		
	β	*SE*	*p*	*β*	*SE*	*p*
Model 1						
Intercept	−1.770	0.652	.008	−2.763	0.658	.001
Age	0.021	0.009	.023	0.038	0.009	.001
Gender	0.555	0.190	.005	0.166	0.196	.399
Time since arrival	−0.306	0.095	.002	−0.103	0.105	.331
Model 2						
Intercept	−1.762	0.647	.008	−2.766	0.660	.001
Age	0.021	0.009	.026	0.039	0.009	.001
Gender	0.601	0.191	.002	0.158	0.198	.429
Time since arrival	−0.327	0.094	.001	−0.099	0.108	.363
Length BP attendance	0.156	0.099	.120	−0.029	0.103	.781
Model Comparisons						
Base vs. Model 1	F (3, 146) = 3.95, p = .010, R^2 = 0.104			F (3, 143) = 5.812, p = .001, R^2 = 0.185		
Model 1 vs. Model 2	F (1, 145) = 1.31, p = .254, ΔR^2 = 0.014			F (1, 142) = 0.078, p = .781, ΔR^2 = 0.005		

See Table D1

Appendix E

Table E1. Between-group comparison, univariate multiple-regression models predicting children's behavior problems.

	Emotion problems			Hyperactivity/ Inattention			Conduct problems		
	β	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*
Intercept	−0.060	1.265	.962	1.723	1.298	.188	1.147	1.255	.363
Age	−0.002	0.016	.921	−0.023	0.0161	.159	−0.014	0.016	.373
Gender	0.046	0.217	.834	−0.395	0.211	.065	−0.412	0.201	.043
Time since arrival	−0.046	0.111	.683	−0.015	0.116	.897	−0.043	0.106	.687
BP vs. no program	0.316	0.235	.182	0.631	0.244	.011	0.378	0.236	.113

Univariate regression models predicting subscales of the Strengths and Difficulties Questionnaire on age-matched sample and comparison group of refugee children without previous early education attendance (T_{1m} vs. T_C $N = 55 + 55$). The predictor of interest "BP attendance vs. no ECD program" was added (i.e., dichotomous, between-person variable). *β*, standardized beta-coefficients for predictors, lower estimates are associated with better outcomes for children from BPs (except for the prosocial behavior). *SE*, standard error for standardized beta-coefficient. *p*, uncorrected two-sided alpha-error probability.

Table E2. Between-group comparison, univariate multiple-regression models predicting children's behavior problems.

	Peer-interaction problems			Total problem score			Prosocial behavior[a]		
Predictors	*β*	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*
Intercept	0.971	1.280	.450	1.363	1.189	.254	−0.963	1.238	.439
Age	−0.012	0.016	.453	−0.018	0.015	.217	0.013	0.015	.390
Gender	−0.104	0.209	.621	−0.310	0.199	.123	0.482	0.201	.019
Time since arrival	0.026	0.114	.819	−0.029	0.109	.790	0.049	0.110	.660
BP vs. no program	0.093	0.236	.696	0.531	0.221	.018	−0.706	0.231	.003

See Table E1

[a]Scale on prosocial behavior is reversed compared to problem scales.

Appendix F

Table F1. Between-group comparison, univariate multiple-regression models predicting childhood development.

	German Receptive Language			Nonverbal Reasoning			Fine Motor Skills		
Predictors	*β*	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*
Intercept	−4.874	1.102	.001	−1.441	1.254	.254	−0.835	1.186	.483
Age	0.058	0.014	.001	0.022	0.015	.154	0.011	0.015	.449
Gender	0.118	0.201	.557	−0.045	0.207	.829	0.579	0.203	.005
Time since arrival	0.231	0.107	.034	−0.170	0.112	.132	−0.216	0.108	.048
BP vs. no program	0.079	0.202	.701	−0.702	0.226	.003	−0.663	0.212	.002

Univariate regression models predicting assessment scores from performances in developmental tasks on age-matched sample and comparison group of refugee children without previous early education attendance (T_{1m} vs. T_C $N = 55 + 55$). The predictor of interest "BP attendance vs. no ECD program" was added (i.e., dichotomous, between-person variable). *β*, standardized beta-coefficients for predictors, lower estimates are associated with better outcomes for children from BPs (except for the prosocial behavior). *SE*, standard error for standardized beta-coefficient. *p*, uncorrected two-sided alpha-error probability.

Table F2. Between-group comparison, univariate multiple-regression models predicting childhood development.

	Gross motor skills			Visuo-motor skills		
Predictors	*β*	*SE*	*p*	*β*	*SE*	*p*
Intercept	−4.114	1.212	.001	−1.633	1.272	.202
Age	0.049	0.015	.002	0.018	0.016	.266
Gender	0.223	0.218	.309	0.232	0.226	.306
Time since arrival	−0.104	0.118	.382	−0.187	0.113	.103
BP vs. no program	−0.045	0.218	.837	0.187	0.230	.418

See Table F1