RESEARCH

Open Access

Implementation and quality of an early childhood education program for newly arrived refugee children in Germany: an observational study

Julian Busch[1,2]* , Thimo Buchmüller[1] and Birgit Leyendecker[1]

*Correspondence:
julian.busch@rub.de

[1] Faculty of Psychology, Child and Family Research, Ruhr-University Bochum, Bochum, Germany
[2] Center for Research on Education and School Development, TU Dortmund University, Room 208, Martin-Schmeißer-Weg 13, 44227 Dortmund, Germany

Abstract

Early childhood education [ECE] can foster the social-emotional adjustment and development of young refugee children. Still, the large numbers of newly arriving refugee families challenge the ECE capacities of host countries. In Germany, state authorities have subsidized flexible ECE programs for refugee children in response to this situation. The goal of this study was to examine the implementation and quality of these programs. In the first study phase, we categorized the seemingly heterogeneous ECE programs and assembled measures to assess their ECE quality. In the second study phase, we evaluated the ECE quality of a randomly selected sample of these ECE programs ($N = 42$) using standardized observation procedures. The ECE programs were implemented differently in temporary setups (caravans, tents), improvised settings (parish rooms, refugee accommodations), or education settings (preschools, elementary schools). To evaluate ECE quality, we created an observation tool for structural quality and coded dimensions from the Classroom Assessment Scoring System Pre-K for process quality. Overall, structural quality was acceptable but differed between implementation settings. Process quality was consistently high, independent of the settings. Our findings suggest that adaptive ECE programs under a flexible childcare policy could support young refugee children after they arrive in host countries. Still, such ECE programs do not compensate for center-based ECE services because of their more vigorous emphasis on children's social-emotional adjustment than pre-academic learning. Further research should consider adaptive assessment tools to assess ECE quality, taking into account heterogeneous program implementation strategies and the specific needs of refugee children.

The early years of life lay the foundations for subsequent human development over the lifespan. In conjunction with forced displacement, chaos, threat, and deprivation threaten positive early childhood development (ECD) (Fazel et al., 2012; Park et al., 2018). Findings from Germany yield that young children from newly arrived refugee families demonstrate higher rates of socio-emotional problems (Buchmüller et al., 2018) and lower levels of cognitive development (Busch et al., 2021). In Germany, more than

163,000 children under six fleeing from Syria, Afghanistan, and Iraq applied for asylum in 2016 (BAMF, 2017). Since then, the immigration of families with young children to Germany has continued and reached new heights during the war in Ukraine in 2022.

Given our knowledge of the consequences of early adversity due to forced displacement, there is a pressing need for action to support the ECD of refugee children after their arrival. In the largest state in Germany, North-Rhine Westphalia (NRW), federal stakeholders in childcare have addressed this need through a new policy that funds specialized early childhood education (ECE) programs. The policy provides only a few regulations such that the ECE programs can flexibly adapt to different settings and diverse needs. Our study assessed the programs' implementation strategies and their ECE quality.

Early childhood education programs promote positive youth development

Investing in early childhood development (ECD) through ECE services has demonstrated the most extensive benefits for children's development over the life span (Anders, 2013; Schweinhart et al., 1993). ECE services subsume a range of child- and caregiver-centered programs that foster ECD and pre-academic learning of children below school age. Therefore, ECE services not only stimulate children's motor, social, emotional, language, and cognitive development (Hancock et al., 2012; High, 2008), but also facilitate child behavioral adjustment and provide resources to families. In particular, disadvantaged children can disproportionately benefit from attending ECE programs because they help children reach their developmental potential (Grantham-McGregor et al., 2007; Sincovich et al., 2019; Weiland & Yoshikawa, 2013; Winsler et al., 2008).

Recently, the significant portion of young children among refugee and immigrant populations worldwide has drawn heightened attention of policymakers and researchers to ECE. While there is still relatively little evidence on the benefits of ECE for displaced refugee populations, a body of work on immigrant populations highlights the multiple benefits of ECE programs for immigrant children's developmental trajectories. Specifically, their attendance of ECE programs in a host country positively affects the social-emotional adjustment (i.e., behavioral and psychological responses to changing environments) and host country language acquisition in the short-term. In the longer term, ECE program attendance fosters their later academic and life achievements (Castro et al., 2011; Votruba-Drzal et al., 2015). Thus, ECE programs potentially mitigate developmental, educational, and socio-emotional disparities already found in young refugee children (Buchmüller et al., 2019; Busch et al., 2021).

While research shows positive long-term benefits of ECE programs for disadvantaged and immigrant populations, there is a debate on how promotive effects emerge in such at-risk populations and how these effects are determined. ECE programs likely support young children from disadvantaged and immigrant populations through direct and indirect pathways. Indirectly, programs improve the macrosystemic influences on children's ECD. These effects are primarily driven by child-directed collaborations between ECE staff and caregivers toward meeting children's needs for positive ECD (e.g., Lee et al., 2006; Marti et al., 2018). Regarding direct effects, ECE programs facilitate ECD by providing enriched learning environments, especially stimulating interactions with program organizers and other children (Winsler et al., 2008). Dosage effects, for example,

substantiate the direct pathways. The frequency of children's program attendance predicts developmental outcomes, especially for children from disadvantaged families (Zaslow et al., 2016).

There is preliminary evidence that the attendance of ECE programs could similarly support the ECD of refugee children via both pathways as well. In an interview study with African refugee mothers, New et. al. (2015) found that caregivers' attendance of child playgroups after arrival provided refugee families with social-emotional support and host countries' socio-cultural knowledge regarding their young children's education (e.g., school readiness requirements, access to the education systems). Studies measuring the direct effects of ECE programs on refugee children's ECD are scarce; yet they suggest that children's social-emotional adjustment, pre-academic skills, and language acquisition could be supported (Busch et al., 2021; Erdemir, 2021). The quality of ECE programs thereby requires additional consideration to understand better why, how, and under which circumstances ECD programs are effective for refugee children (Murphy et al., 2018). Studying ECE quality is essential in developing ECE standards and practices that also consider the specific needs of refugee children.

Early childhood education quality determines the impact on child development

Previous research demonstrated that the effectiveness of ECE programs depends on the overall ECE quality (Burchinal et al., 2000, 2010; Sammons et al., 2014). ECE quality subsumes the structural and process characteristics of a program. The structural quality of ECE environments are characterized by the physical (e.g., group, staff, and equipment), spatial (e.g., location), and temporal conditions (e.g., schedule and routines) (Thomason & Paro, 2009). Process quality encompasses children's social, emotional, and pre-academic learning experiences during their program attendance. In particular, staff–child interactions facilitate children's learning experiences (Howes et al., 2008). Process quality can be separated into instructional support (i.e., cognitive stimulation and pre-academic activity) and social-emotional support (i.e., feelings of comfort and security, positive social interactions). Previous studies demonstrated distinct effects of both structural and process quality on the academic and socio-emotional development of children from general populations (Anders et al., 2013; Bradley et al., 2001; Trawick-Smith et al., 2016). Beyond the main effects, structural quality is considered to lay the groundwork for the impact of high process quality (Burchinal, 2018). Process quality was moreover found to be the primary driver for positive ECD in ECE programs (Slot et al., 2015).

To date, few studies with mainly qualitative approaches specifically inform on the relevant structural and process characteristics of ECE programs for refugee children (Busch et al., 2018; Hurley et al., 2013, 2014). Those studies emphasized the heightened importance of specific structural characteristics, such as clear routines and schedules, frequent use of symbols for communication and self-expression, as well as links to the local social service providers for refugee children. Beyond the structural characteristics, ECE staff highlighted some components of process quality characteristics for refugee children. These were high responsiveness and supportive interactions due to children's increased risk for socio-emotional problems. Staff moreover mentioned that interactions with a focus on language are fundamental because refugee children are typically dual language

learners. However, the previous studies have mainly aggregated idiosyncratic evidence as they reflect the experiences of educational staff working with refugee children in diverse ECE programs. Moreover, the respective quality of the ECE programs under investigation is unknown and rarely considered in such study evidence.

Assessing quality in heterogenous ECE programs for refugee children is challenging

Measuring quality among specialized ECE programs for refugee families is challenging for several reasons. First, specialized ECE programs can have rather different conceptual orientations (e.g., center-based preschool programs versus child playgroups), especially given the immense diversity of refugee families and their respective living circumstances. While some specialized programs tend to emphasize the indirect effects of ECE with holistic orientations (i.e., playgroups: address families and thus support children's environments in which ECD occurs), others emphasize direct effects and are exclusively child-directed (i.e., center-based preschool programs: stimulating pre-academic learning). Second, ECE programs can be general in their goals (e.g., supporting school readiness) or rather specific (fostering specific competencies and coping with migration-related challenges such as language acquisition and social-emotional adjustment after arrival). When assessing the ECE quality of programs that address the specific needs of refugee children, we argue that the program's conceptual orientations should be considered. If not, results on ECE quality following non-adapted assessment tools are likely a function of the receptive ECE concept or implementation setting. Non-adapted quality assessments could moreover underestimate the specific needs of young refugee children.

The established ECE quality observation tools usually comply with universal ECE paradigms and embedding ECE systems. Specifically, quality observation tools reflect national program regulations by stakeholder authorities and (at least implicitly) assume specific implementation settings. For example, the Early Childhood Environment Rating Scale (ECERS-R; Harms et al., 2015) is a widely used observation tool designed to examine structural characteristics of state-subsidized and center-based preschool programs in Western, high-income countries (Betancur et al., 2021). The ECERS-R thus is less applicable to evaluating the ECE quality of playgroups, typically a more informal type of ECE programs. Playgroups are usually more flexibly organized, set up in less equipped settings, and engage caregivers more frequently than center-based preschool programs (Sincovich et al., 2019). Moreover, playgroups emphasize social-emotional adjustment (e.g., connecting caregivers and children with the community, fostering a sense of belonging) and joyful activities over children's progress in pre-academic learning.

Substantially less work has investigated standardized tools for measuring ECE quality among playgroups (Commerford & Robinson, 2016). One reason could be that it is more difficult to propose univocal guidelines given the diverse ECE concepts and goals among playgroups. In one attempt, researchers proposed principles of a high-quality playgroup using focus groups (Commerford & Hunter, 2017; Jackson, 2013). The postulated principles are to appropriately stimulate early childhood experiences, increase parental knowledge on ECD and learning, facilitate social networks, support transitioning into education, and provide resources as well as referral to appropriate services. Overall, these principles reflect findings on the benefits of ECE programs that were distinctively

pronounced for refugee populations. Consistently, international ECE initiatives recently generated sets of items measuring the ECE quality of playgroup-like services for refugee populations (e.g., Russo et al., unpublished; UNESCO et al., 2017). Those sets, however, were developed along with specific curricula by funding agencies, designed for emergency contexts, or blended different quality criteria for easy administration. We need further research on how to adaptively measure the quality of diverse ECE programs that address the needs of specific populations (in our case recently arrived refugee children) in different implementation settings and with different ECE program orientations and concepts.

Flexible ECE programs for refugee children—Bridging Projects in Germany

The challenge to set up and effectively regulate ECE programs for refugee children has been emerging in Germany since 2016. The Ministry of Children, Families, Refugees and Integration (MKFFI) of the largest German state, North-Rhine Westphalia (NRW), then introduced an ECE policy to support the ECD of newly arriving refugee children. Local stakeholders in ECE, such as the Communal Youth Welfare offices and private ECE agencies, were granted flexibility in implementing a range of ECE programs, so-called "Bridging Projects" (BPs), each tailored to the local circumstances and the diverse needs of young refugee children and their families in their respective reach. Based on that policy, the state ministry has annually funded more than 1000 ECE programs with an overall capacity of more than 10,000 children. On average, BPs offered enrollment to 8.6 (SD = 4.05) children per group, had a duration of 33.5 weeks (SD = 14.23), and a caretaking time of 10.41 h per week (SD = 8.27; own calculations based on registration data for BPs provided by the state authorities). Attendance is subsidized as BP organizers receive a flat rate of €30 per hour for the caretaking of one to five children. The few regulations request that at least one staff member per group has a qualification in ECE (i.e., formal training or a degree in an ECE-related subject), and the staff–child ratio should be 1:5 or better. Volunteers are encouraged to support trained staff. BP organizers are free to choose the location, time, frequency, and age range of children before school entry and the involvement of parents. Consequently, the implementation of such specialized ECE programs can range from highly structured preschool programs to low-barrier mother–child playgroups.

Study aim

Studying the implementation of specialized ECE programs contributes to generating meaningful ECE strategies for refugee children. Specifically, studying the BPs can inform stakeholders (1) on variations between ECE programs when policies provide only a few regulations and (2) on how to refine program guidelines when programs are created locally and regulated at scale. Our study contributes to these pending issues by investigating the implementation and ECE quality of the BPs. Using a two-phase study approach, we (A) explored heterogeneous implementation strategies for BPs and generated a set of measurements to assess ECE quality among heterogeneous BPs. We then (B) evaluated the ECE quality of BPs.

Method

Study design

We subdivided our study into two phases. In the first study phase, we identified a hypothetical scheme to categorize the heterogeneous BPs, and we selected a set of indicators to assess ECE quality among BPs. To achieve these goals, we reviewed BP registration data provided by state authorities, conducted explorative field observations, and reviewed available observation tools along with guidelines on ECE quality. In the second study phase, we conducted structured field observations in a randomly selected sample of BPs. Based on this data, we examined the newly created observation measures on ECE quality, evaluated the program quality of BPs, explored whether ECE quality varied across the hypothetical categorization scheme (from study phase 1), and compared the process quality of BPs to childcare centers in NRW. We report the results of study phase 1 as embedded within our study method.

Study phase 1

Categorization of Bridging Projects

The state authorities provided us with a registration list of BPs that included brief and unstructured descriptions of the BPs. The first and second authors and two research assistants independently reviewed this registration list and identified characterizing attributes of BPs. Subsequently, the team structured the content during group discussions and decided on the final categories. BPs were described as located in settings for education, improvised settings, or more flexibly organized in mobile and temporary setups (see Table 1 for details). The classification scheme could be substantiated in six explorative field observations of seemingly different BP types we selected based on the registration list. Located *in non-education facilities*

Approaches to observing childcare quality of Bridging Projects

We also explored ECE quality characteristics following an open-observation report form during the explorative field observations. We decided a priori to split ECE quality into structural and process components. Accordingly, the report form included the three

Table 1 Scheme to distinguish Bridging Projects by their implementation settings

Type	Rationale	Setting	Examples
Education settings	Located in facilities of education-services	Facilities of childcare and early childhood education	Daycare centers Preschools
		Facilities of primary education	Elementary schools
Improvised settings	Located in non-education facilities	Neighborhood centers	Parish rooms/halls Community centers Municipal libraries
		Refugee centers	Refugee residences Asylum camps
Mobile concepts and temporary setups	Located in provisional areas and rooms	Outside areas	Parks Playgrounds
		Temporary facilities	Converted caravans Tents

Categorization was generated based on Bridging Project registration lists provided by state authorities and probed in explorative field observations

domains: (1) a general project description, (2) structural quality, and (3) process quality. Each domain included a few guiding questions to organize the unsystematic exploration process (e.g., number of children and staff members, parents' participation; goals of the BP [general]; characteristics of the location; description of the program setup [structural quality]; children's activities, staff–child interactions, the role of language [process quality]). The report form supported comparative group discussions of ECE quality indicators across BPs to identify the best approach to evaluating structural and process quality among the various BPs. The first study author facilitated group discussions, while the second study author mainly conducted documentation of the meetings. Additionally, between two and four research assistants with at least bachelor's degrees participated in all group discussions.

Structural quality   The available instruments on structural quality in ECE research, such as the ECERS-R (Harms et al., 2015) and the "Child Care Checklist Physical Environment Checklist" (NICHD, 2006), were considered not equally applicable to the heterogeneous BPs. In the group discussions, some indicators for key concepts of structural quality were thus identified and adapted concerning the widely used observation inventories. Furthermore, we examined the fit of selected indicators in further explorative field observations. We set the following guidelines for the development and identification of structural quality indicators: we intended to (1) maximize the applicability of indicators to different BPs and their relevance to the ECE principles proposed by stakeholder authorities in NRW (MKFFI, 2016; initiator and funder of the refugee ECE policy), and (2) create tool structure and content as comparable as possible to the established measures of ECE quality. The measurement should also be feasible with good reliability and internal validity. We describe the generated observation tool for structural quality in "Measures" section of study phase 2.

Process quality   For process quality, there was team consensus that key concepts of constructs and tools that focus on staff–child interactions yielded overall applicability to BPs. However, we acknowledged that BPs overall emphasize children's social-emotional adjustment rather than the support of pre-academic learning. Therefore, as a measure of process quality in BPs, we decided to use a subset of dimensions of the Classroom Assessment Scoring System Pre-K [CLASS] (La Paro et al., 2002), a widely established and flexible observation tool of process quality in ECE programs. We describe our selection of process quality indicators in "Measures" section of study phase 2.

Study phase 2
Sample of Bridging Projects
We randomly drew BPs from the registration list and requested their participation. We stopped the recruitment after $N = 50$ BPs consented. At this point, we had contacted a total of 153 BPs. Of those BPs that did not participate, some organizers did not respond. Others reported no active BP due to the relocation of families or stated concerns that the study might disturb the safe space atmosphere in groups. We lost two BPs before data collections had started (one group closed as scheduled, and the other was closed due to decreasing numbers of participants). We used the initial 6 BP visits of study phase 2 for

preparing instruments, piloting across heterogeneous BPs, and observer trainings. The analytical study sample for structured field observations consisted of $n = 42$ BPs. Among these BPs, $n = 14$ were set up in education settings, $n = 22$ in improvised settings, and $n = 5$ in temporary setups or following mobile concepts.

During field visits in the $n = 42$ BPs, an average of $M = 1.56$ ECE staff (SD = 0.50; median = 2) and $M = 5.65$ children (SD = 2.07, median = 6) were present. The average staff–child ratio during the visits was 1:3.56 (SD = 1.35, Median = 1:3.50). The most frequent caregiving constellations that we observed were "*activities with 2 to 6 children*" (27.38%), followed by "*one-on-one*" interactions (25%) and "*activities with more than 6 children*" (25%). For 32 BPs, we obtained additional information on staff and the country of origin of the participating children. In total, 452 children attended those BPs regularly. The children's major countries of origin were Syria (40.39%), followed by South-Eastern European countries (13.27%), Iraq (13.05%), and Afghanistan (12.61%). The reporting staff was on average $M = 41.33$ years old (SD = 11.84, Median = 40.50, range = [20, 61]) and 12.5% were male. Regarding staff education levels, 15.6% had a college degree or had completed an ECE-related subject in tertiary education. 46.9% had received ECE-related vocational training; 12.5% were childcare assistants. 21.9% of teachers did not report any ECE-related qualification, and 3.1% omitted this information.

Procedure

Teams of two observers conducted structured field observations in BPs between February and April 2017. The entire observation team for BPs consisted of five graduate students, each with a bachelor's or master's degree in psychology. The first and the second study authors instructed and supervised the team on all observation procedures. Additionally, four observers (the same research assistants who participated in study phase 1) and the first and second study author were officially trained and licensed in the CLASS. One person assessed structural quality during the BP observations, while the other assessed process quality.

All BP staff involved in the present study provided written informed consent beforehand. All parents of children who attended the participating BPs during our study received written information on the study and verbal information from staff. Families were asked not to attend the BP on the day of the observation if they felt uneasy about the study. No child-level data were analyzed in this study. The Internal Review Board of the *<Faculty-University>* approved the study protocol (2016-298) following the ethical guidelines of the German Psychological Society.

Measures

Structural quality   Based on study phase 1 results, we created the "Bridging Project Evaluation Scale" (BREVIS) to observe structural quality in diverse ECE environments. BREVIS consists of 24 indicators of structural quality, which are assigned to five dimensions: (1) *premises*, covering structural aspects of the setting such as availability of space for activities, an area for relaxation, or sanitary facilities, (2) *equipment*, covering the availability of movable furniture and their suitability for young children, (3) *structuring of a session*, covering the formal structure of the program, including clearly indicated start and end times, establishment of rituals, rules, and routines, (4) *team coherence*, charac-

teristics of team climate and the degree of effective staff cooperation, and (5) *educational materials* for pre-academic activities and play, as well as for language facilitation in multi-lingual groups. We provide an overview of the BREVIS tool with exemplary evaluations in Table 2. Completion of the BREVIS by a single observer on-site took around 30 min. Each indicator was rated on a three-point Likert scale (1—*inadequate*, 2—*acceptable*, 3—*very good*). Anchors for each indicator facilitated ratings. Observers could additionally comment on their ratings in a separate column. When observers felt uncertain about some ratings, we discussed those ratings in subsequent group meetings.

Process quality The CLASS Pre-K observation tool assesses different aspects of caregiver interactions with preschool-aged children on-site by an independent observer. Given the overall emphasis of BPs on social-emotional and behavioral adjustment, we omitted 4 CLASS dimensions that were more strongly linked to pre-academic learning. The selected 6 dimensions for our study were "*positive climate*" (e.g., relationships, positive affect), "*negative climate*" (e.g., punitive control, disrespect), "*teacher sensitivity*" (e.g., awareness, responsiveness), "*behavior management*" (e.g., redirection of misbehavior, clear expectations), and "*productivity*" (e.g., preparation, transitions from one activity to another). In addition, we selected the CLASS dimension "*language modeling*" (e.g., frequent conversation, self- and parallel talk) because the acquisition of basic host language skills is especially relevant to behavioral adjustment and navigating social situations. We further added the dimension "*teacher involvement*" suggested by Agache et. al. (2018). Higher ratings on *teacher involvement* indicated more active engagement and greater attention to the children's activities. All 6 CLASS dimensions, as well as the additional dimension *teacher involvement*, were rated on scales ranging from "1" indicating low, over "4" indicating moderate, to "7" indicating high staff–child interaction quality (i.e., process quality). One observer per BP conducted two CLASS observation cycles of 15 min each. The four licensed CLASS observers passed the official online reliability test with average inter-rater agreement rates between 80 and 94% compared to gold-standard raters. Internal consistency of CLASS scales for our sample of BPs was acceptable (average $\alpha = 0.75$, thresholds according to Tavakol & Dennick, 2011).

Analytical approaches

Reliability and internal validity of the BREVIS observation tool Unlike the CLASS, BREVIS was newly created for this study. We examined estimates for reliability and internal validity of the BREVIS based on the sample of BPs from study phase 2. Specifically, we calculated inter-rater reliability for the BREVIS indicators in four double coding sessions using two-way consistency, single-measure intra-class correlations (ICC) with random effects. We expected at least moderate ICC coefficients on domain levels (ICC ≥ 0.50; thresholds according to Koo & Li, 2016). For internal validity, we calculated the two following estimates based on the entire sample. First, we used Cronbach's Alpha to evaluate internal consistency for each BREVIS domain. We expected at least satisfactory values to justify sum score calculations of BREVIS domain-level scores ($\alpha > 0.70$, Tavakol et al., 2011). Second, we used Spearman rank correlations to evaluate inter-domain relations based on intercorrelations of BREVIS domain-level sum scores. We expected moderate intercorrelations ($r = 0.20$–0.50), assuming that different ECE quality criteria are partially linked yet reflect different facets of ECE quality.

Table 2 Structural quality according to the "Bridging Project Evaluation Scale"

BREVIS dimension and indicators	Examples of indicators rated as "*very good*"
Premises	
Lighting	The main room has at least one window/the lighting is sufficient to compensate for lacking daylight
Indoor climate	There is sufficient heating; no signs of critical humidity, and lacking ventilation (e.g., mold/fogged windows) in the main room
Cleanliness	There is no waste on the floor, no intrusive smell, surfaces and materials appear clean
Cond. of indoor area[a]	There are no/or few signs of wear and tear (e.g., wallpapers, windows, lighting)
Cond. of outdoor area	There is an outdoor area with sufficient space for gross motor activities (e.g., soccer, bobby car driving); staff is capable of ensuring the safety of several children at once by supervision
Space for children	The room size is at least 2.5 m^2 per child with freedom of movement and few obstructing furniture
Cond. of sanitary facilities	There is a child-suited sanitary facility (child-friendly toilet seat/changing table for infants and toddlers, if necessary) with an equipped washbasin (freshwater, soap, towel)
Relaxation area	There is space for children to calm down and withdraw from demanding activities (need not be sufficient space for all children at once)
Equipment	
Interior equipment	There is sufficient furniture for play, learning, and care relative to the group size; there is sufficient storage space for materials
Child-friendly furniture	The size/height of chairs/tables is child-friendly; there is sufficient child-friendly furniture
Cond. of furniture[a]	Tables, chairs, and cabinets are functional
Safety of furniture	There are no potential hazards (e.g., sharp edges, open electrical outlets), no accessible exterior doors or unattended kitchen areas
Structuring of a session	
Beginning session	The staff greets each arriving child individually, supports in cases when separation difficulties occur, and facilitates finding an initial activity
Structure of daily routine	There is a recognizable session structure; an adaptive balance between the staff's responsivity to children's needs and interests and following a fixed session structure
Ending session	Staff says goodbye to each child at the end of the session; the group ends the session together
Routine/rituals/rules	The staff employs recurring elements (including do's/dont's) in the session structure (both within and between sessions); children appear to recognize these elements, and the staff reinforces these elements
Team coherence	
Atmosphere within team	Staff express mutual appreciation and support to each other; they act following the same rules, especially towards the children
Cooperation of staff	Staff effectively divides tasks with apparently clear responsibilities; their actions are consistent and appear coordinated
Educational materials	
Toys	There is sufficient and age-appropriate material for joyful activities relative to the group size (e.g., play cars, dolls, board or card games)
Language support	There are age-appropriate books/picture books and engaging language games of a sufficient amount relative to the group size
Artistic activities	There are age-appropriate coloring books, copies (e.g., mandalas), materials to color (e.g., watercolors, crayons)/there is plasticine, both of sufficient amounts relative to the group size
Fine motor skills	There are age-appropriate puzzles, beads, and handwork materials of a sufficient amount relative to the group size
Quantitative reasoning	There are age-appropriate materials such as beads, rulers, slide rules, shapes, and geometric objects of sufficient amount relative to the group size
Language facilitation	Staff uses pictograms, images, and sets of symbols to bridge language barriers

Structural quality   We descriptively analyzed BREVIS ratings in two ways to evaluate the structural quality of BPs. First, we computed means, confidence intervals, and ranks for each BREVIS indicator and calculated sum scores on the dimension-level. We defined thresholds for interpretation based on the numeric equivalents of the BREVIS scoring scheme (0 = inadequate, 1 = acceptable, 2 = very good). We thus used the following thresholds for domain-level mean scores: < 0.5 = "inadequate", > 0.5 to < 1.5 = "acceptable", > 1.5 = "very good". Second, we analyzed the frequency of "inadequate" ratings on the indicator-level across all BPs. We defined thresholds regarding the number of BPs that showed ratings better-than-inadequate as follows: the relative portion of BPs better-than-inadequate is over 85% (achieved in almost all BPs); between 85% and over 70% (achieved in many BPs); between 70% and over 55% (inconsistently achieved in BPs); 55% or less (only occasionally achieved in BPs).

Process quality and comparison to daycare centers   We followed the rating scheme by the tool manual (Pianta et al., 2008) to descriptively analyze and interpret CLASS ratings. Specifically, we computed means for each CLASS dimension and a second stratum score for overall *social support* by summarizing dimension ratings of *positive climate*, *negative climate*, *teacher sensitivity*, *behavior management*, and *productivity*. We considered thresholds for interpretation based on the numeric equivalents of the CLASS scoring scheme, i.e., mean scores < 2.5 = low, > 2.5 and < 5.5 = medium, > 5.5 = high. We compared CLASS ratings for BPs to a representative sample of ECE groups in daycare centers in NRW ($N = 177$). In those groups, the average group size was $M = 21$ children with an average staff–child ratio around 1:6.49 (SD = 3.60, median = 5.75). On average, BPs showed a better staff–child ratio, $t(227.96) = -9.44$, $p < 0.001$, $d = 0.99$). For more information on this comparison sample, see the publication by Bihler et al. (2018).

ECE quality between different types of BPs   To investigate the structural and process quality of differently implemented BPs (i.e., mobile concepts, improvised settings, education settings), we compared aggregated-to-dimension BREVIS and CLASS ratings separated by BP types.

Interpretation of all inferential parameters followed two-sided testing at an Alpha-error level of 5%. We additionally report Cohen's *d* with pooled variances to evaluate the effect sizes. In correlation tables, we indicate statistical significance based on a $p < 0.05\%$ threshold to better account for multiple testing. All analyses were run in R using the packages *Hmisc*, *lsr*, *coin*, and *psych* (3.5.0; R Core Team, 2014).

Results

Reliability and construct validity of the BREVIS

We calculated reliability estimates of the BREVIS indicators based on observations in 41 BPs; one observation was incomplete due to early closing and therefore excluded. The average *ICC* for the BREVIS was overall good (mean ICC = 0.724), and ICCs for the BREVIS dimensions, respectively, showed moderate-to-excellent inter-rater reliability (ICC range = [0.56; 1.00]). On internal validity, Cronbach's Alphas showed overall good ($\alpha = 0.80$), and moderate-to-good internal consistency on dimension

levels. We found moderate-to-substantial correlations between BREVIS domain-level scores. See Table 3 for details.

Structural quality

We summarize our findings on BREVIS observations for each dimension (*premises, equipment, structuring of a session, team coherence, educational materials*). Detailed results are depicted in Table 4. The quality of *premises* was overall acceptable. Only conditions of sanitary facilities for children were inadequate in some BPs, and areas for relaxation were inconsistently available. The quality of *equipment* was, on average, very good. Almost all of the BPs provided different kinds of equipment of acceptable quality, at least. On the dimension *structuring of a session*, observations indicated on average acceptable quality. We observed inadequate routines/rituals/rules or an unclear session ending only in a few BPs. On the dimension *team coherence*, ratings indicated very good quality for our sample of BPs. Almost none of the BPs showed inadequate ratings on this dimension. On the dimension *educational materials*, we observed acceptable quality overall. Almost all BPs of our sample provided at least acceptable educational materials. Only some BPs lacked material for language facilitation in multilingual settings. Moreover, BPs provided children with relatively fewer materials for quantitative reasoning in comparison to other types of materials.

Process quality and comparison to daycare centers

We analyzed process quality based on the CLASS observations in 41 BPs. Detailed results are depicted in Table 5. All of the socio-emotional dimensions (*positive climate, negative climate, teacher sensitivity, behavior management,* and *productivity*) were, on average, rated in the high-quality range for our sample of BPs (mean range $= [5.55, 6.87]$). *Language modeling* was overall rated within the medium range. Ratings of the additional domain *teacher involvement* revealed that staff in BPs was frequently engaged in activities with the children ($M = 5.07$; $SD = 1.16$). We further compared CLASS ratings on the BPs to ECE groups in daycare centers (see Table 4). For the BPs, we found fewer *negative interactions* ($t(80.24) = 2.78$, $p < 0.01$, $d = 0.40$), higher *productivity* ($t(51.97) = 3.12$, $p < 0.01$, $d = 0.62$), and better *language modeling* ($t(48.81) = 3.86$, $p < 0.01$, $d = 0.88$). We found no differences in the domains *positive climate, teacher sensitivity,* and *behavior management*. The second-stratum dimension, *social support*, yielded better ratings for BPs than for daycare centers ($t(52.19) = 2.42$, $p < 0.05$, $d = 0.47$).

Comparing ECE quality between different Bridging Project types

We excluded the subsample of BPs with mobile concepts or temporary setups from between-type inferential comparisons due to its small subsample size ($n = 5$). Focusing BREVIS on domain levels, we analyzed structural quality comparatively for the different BP types (Table 4). Except for *equipment* ($t(33.28) = -0.93$, $p > 0.05$, $d = 0.30$), BPs in education settings tended to have higher scores on structural quality dimensions when compared to those in improvised settings. We found largest differences

Table 3 Intercorrelations for the domain-level scores of "Bridging Project Evaluation Scale" and "Classroom Assessment Scoring System Pre-K" ratings

Dimension-level ratings	1	2	3	4	5	6	7	8	9	10	11	ICC	α
BREVIS													
Premises (1)	–											0.50	0.71
Equipment (2)	0.73*	–										1	0.75
Structuring of a session (3)	0.40	0.11	–									0.95	0.80
Team coherence (4)	0.52*	0.44	0.52*	–								0.67	0.74
Educational materials (5)	0.45*	0.52*	0.36	0.31	–							0.69	0.77
CLASS													
Positive climate (6)	0.29	0.22	0.22	−0.05	0.11	–						–	0.76
Negative climate (7)	−0.06	0.09	0.07	−0.16	0.16	0.41	–					–	0.58
Teacher sensitivity (8)	0.22	0.18	0.30	0.04	0.30	0.54*	0.44	–				–	0.87
Behavior management (9)	0.20	0.16	0.35	−0.05	0.37	0.27	0.48*	0.65*	–			–	0.78
Productivity (10)	0.05	−0.05	0.53*	0.12	0.50*	0.17	0.18	0.40	0.33	–		–	0.67
Language modeling (11)	0.13	0.15	0.21	0.00	0.16	0.71*	0.57*	0.58*	0.46*	0.29	–	–	0.78
Teacher involvement (12)	0.19	0.17	0.37	−0.05	0.36	0.45*	0.34	0.65*	0.43*	0.49*	0.54*	–	0.79

Method is Spearman rank correlation for pairwise complete cases; N = 41

ICC intra-class correlations (two-way, consistency, single); internal consistency per dimension is assessed via Cronbach's alpha (α)

*p < 0.005

Table 4 Structural quality according to the "Bridging Project Evaluation Scale"

BREVIS dimension	Inadequate (%)	Acceptable (%)	Very good (%)	M	95% CI	Rank
Premises				1.43	[1.32, 1.55]	
Lighting	0	7.50	92.50	1.93	[1.84, 2.01]	1
Indoor climate	4.88	19.51	75.61	1.71	[1.53, 1.88]	4
Cleanliness	2.44	36.59	60.98	1.59	[1.41, 1.76]	9.5
Cond. of indoor area[a]	0	48.78	51.22	1.51	[1.35, 1.67]	13
Cond. of outdoor area	10.00	37.50	52.50	1.43	[1.21, 1.64]	15
Space for children	12.20	43.90	43.90	1.32	[1.10, 1.53]	18
Cond. of sanitary facilities	25.64	38.46	35.90	1.10	[0.85, 1.36]	22
Relaxation area	37.50	35.00	27.50	0.90	[0.64, 1.16]	24
Equipment				1.54	[1.43, 1.65]	
Interior equipment	2.44	31.71	65.85	1.63	[1.47, 1.80]	7.5
Child-friendly furniture	12.20	19.51	68.29	1.56	[1.34, 1.79]	11.5
Cond. of furniture[a]	0	43.90	56.10	1.56	[1.40, 1.72]	11.5
Safety of furniture	0	60.98	39.02	1.39	[1.23, 1.55]	17
Structuring of a session				1.36	[1.17, 1.55]	
Beginning session	7.69	38.46	53.85	1.46	[1.20, 1.72]	14
Structure of daily routine	10.26	38.46	51.28	1.41	[1.19, 1.63]	16
Ending session	19.23	30.77	50.00	1.31	[0.99, 1.63]	19
Routine/rituals/rules	18.42	39.47	42.11	1.24	[0.99, 1.48]	20
Team coherence				1.71	[1.55, 1.88]	
Atmosphere within team	0	16.65	82.35	1.82	[1.69, 1.96]	3
Cooperation of staff	2.86	28.57	68.57	1.66	[1.47, 1.84]	6
Educational materials				1.48	[1.36, 1.61]	
Toys	0.00	17.07	82.93	1.83	[1.71, 1.95]	2
Language support	7.32	19.51	73.17	1.66	[1.46, 1.85]	5
Artistic activities	2.44	31.71	65.85	1.63	[1.47, 1.80]	7.5
Fine motor skills	4.89	31.71	63.42	1.59	[1.40, 1.77]	9.5
Quantitative reasoning	12.20	60.98	26.83	1.15	[0.95, 1.34]	21
Language facilitation[b]	25.64	46.15	28.21	1.03	[0.79, 1.27]	23

This table represents the list of indicators for the BREVIS observation tool generated in study phase 1. Relative proportions of categorical quality evaluations ($N = 41$); descending ranks, arithmetic means (*M*) and 95% confidence intervals (95% CI) are based on sum scores per indicator (*inadequate* = 0, *acceptable* = 1, *very good* = 2); average sum scores and 95% CIs are additionally provided for each dimension

[a] Cond., conditions

[b] Materials for bridging communication (e.g., through pictograms/symbols/pictures)

for *structuring of a session* ($t(32.70) = -3.71$, $p < 0.001$, $d = 1.19$). Descriptively, BPs with mobile concepts or in temporary setups consistently tended to have the lowest ratings on structural quality indicators. See Table 6 for detailed results. For process quality, we compared CLASS ratings and ratings on *teacher involvement* between BP types. Dimensions of socio-emotional support did not differ between BPs in settings for education or improvised settings except for *productivity* ($t(30.50) = -2.50$, $p < 0.05$, $d = 0.76$). The dimensions *language modeling* and *teacher involvement* did not show differences between the two types. CLASS ratings tended to be slightly lower for BPs with mobile concepts or in temporary setups. See Table 7 for detailed results.

Table 5 Comparing process quality ratings between Bridging Projects and regular daycare centers

CLASS dimension	BPs	Daycare	BPs vs. daycare			
	(*n* = 41)	(*n* = 177)	*t*	*df*	*p*	*d*
Positive climate	6.16 (0.83)	6.04 (0.72)	0.42	52.53	0.674	0.61
Negative climate	6.87 (0.25)	6.74 (0.33)	2.78	80.24	0.007*	0.40
Teacher sensitivity	5.75 (0.91)	5.42 (0.79)	1.94	55.89	0.058	0.37
Behavior management	5.83 (1.10)	5.69 (0.77)	0.56	50.68	0.580	0.12
Productivity	5.55 (1.19)	4.92 (0.99)	3.12	51.97	0.003*	0.62
Social support	6.03 (0.66)	5.76 (0.56)	2.42	52.19	0.019*	0.47
Language modeling	3.69 (1.34)	2.83 (0.83)	3.86	48.81	0.010*	0.88

Means (standard deviations) per subscale and program type (*lowest* = 1, *medium* = 4, *highest* = 7); Welch test due to variance heterogeneity; *n* = 41

BPs bridging projects, *daycare* regular and center-based daycare programs (*n* = 177)

*$p < 0.05$

Table 6 Structural quality by implementation type

BREVIS dimension	Mobile	Improv	Education	Improv vs. education			
	(*n* = 5)	(*n* = 22)	(*n* = 14)	*t*	*df*	*p*	*d*
Premises	0.74 (0.31)	1.45 (0.24)	1.66 (0.24)	− 2.60	27.54	0.015*	0.89
Equipment	1.05 (0.21)	1.57 (0.35)	1.66 (0.25)	− 0.93	33.28	0.361	0.30
Structuring of a session	0.83 (1.04)	1.18 (0.51)	1.74 (0.39)	− 3.71	32.70	0.001*	1.19
Team coherence	1.20 (0.76)	1.69 (0.43)	1.96 (0.14)	− 2.43	22.36	0.023*	0.77
Educational materials	1.30 (0.65)	1.42 (0.41)	1.66 (0.20)	− 2.33	32.47	0.026*	0.69

Means (standard deviations) per subscale and implementation type (*inadequate* = 0, *acceptable* = 1, *very good* = 2); Welch test due to variance heterogeneity; *n* = 41. *Mobile* groups are not considered in inferential testing due to the small subsample

Mobile mobile concepts or temporary setups, *Improv* improvised settings, *Education* education settings

*$p < 0.05$

Discussion

In the largest German state (NRW), a state ministry established specialized ECE programs for newly arrived refugee children (BPs) through a liberal ECE policy. In this study, we evaluated the implementation of the BPs in two study phases. In study phase 1, we created a hypothetical categorization scheme for BPs and generated a set of observation measures on the ECE quality of different BPs. We distinguished BPs based on their implementation settings. We found that process quality indicators were applicable across differently implemented BPs, and we assembled a specific set of structural quality indicators. In study phase 2, we used the set of observation measures to evaluate ECE quality in a sample of BPs. Overall, we found that structural quality in BPs was acceptable or better but differed systematically between implementation types. As can be expected, our data supported that those BPs located in education settings were most likely to provide good structural quality. Process quality was consistently high and independent of the implementation setting, also when compared to center-based ECE programs in NRW.

Table 7 Process quality by implementation type

CLASS dimension	Mobile	Improv	Education	Improv vs. education			
	(*n* = 5)	(*n* = 22)	(*n* = 14)	*t*	*df*	*p*	*d*
Positive climate	5.70 (1.61)	6.18 (0.89)	6.14 (0.84)	0.13	29.07	0.896	0.05
Negative climate	6.90 (0.22)	6.84 (0.28)	6.89 (0.21)	− 0.63	32.95	0.536	0.20
Teacher sensitivity	5.10 (0.89)	5.64 (1.05)	6.00 (0.62)	− 1.31	33.89	0.200	0.40
Behavior management	5.20 (1.35)	5.66 (1.08)	6.18 (1.05)	− 1.43	28.53	0.164	0.49
Productivity	5.50 (1.47)	5.21 (1.35)	6.07 (0.65)	− 2.50	30.50	0.018*	0.76
Social support	5.95 (0.58)	5.89 (0.77)	6.26 (0.47)	− 1.75	32.77	0.090	0.55
Language modeling	3.10 (1.39)	3.61 (1.32)	3.96 (1.42)	− 0.74	26.26	0.464	0.26
Teacher involvement	4.40 (1.25)	5.09 (1.24)	5.61 (0.92)	− 1.43	33.03	0.163	0.46

Means (standard deviations) per subscale and implementation type (*lowest* = 1, *medium* = 4, *highest* = 7); Welch-test due to variance heterogeneity; *n* = 41. *Mobile* groups are not considered in inferential testing due to the small subsample

Mobile mobile concepts or temporary setups, *Improv* improvised settings, *Education* education settings

*p < 0.05

Assessing and interpreting ECE quality in heterogeneous settings

In study phase 1, we found that process quality can be assessed using the dimensions of an established tool. However, measuring structural quality among BPs required more tool adaptation efforts. Our study experience shows that structural indicators could concomitantly depend on the implementation context as they cover the physical characteristics of ECD program environments. With the BREVIS, we selected a set of observable indicators for structural quality across heterogeneous settings and additionally considered refugee children's specific needs. Based on our data, readers should still be aware that variations in structural characteristics could either link to program concepts (i.e., purposeful design) or reflect a lack of general structural quality. We consider both aspects in the interpretations of our observational data.

ECE quality of heterogeneous Bridging Projects

In study phase 2, we evaluated ECE quality among a heterogenous sample of BPs. Overall, structural quality indicators showed acceptable to high quality despite implementation heterogeneity. On several indicators, however, structural quality varied between BP settings. BPs with mobile concepts or in temporary setups were more likely to lack *relaxation areas*, *materials for quantitative reasoning*, or *language facilitation*. The overall findings suggest that ECE staff in BPs could ensure fundamental necessities for ECE under a liberal policy and within different settings. Still, some dependency of structural quality on BP settings also suggests that BPs in more challenging locations for ECE (i.e., mobile concepts/temporary setups) could require additional resources to compensate for the structural disadvantages of the settings.

More specifically, we found the most considerable differences between implementation types for the dimension *structuring of a session*. This finding offers different interpretations. First, the observed differences could be due to transactional costs, thus requiring additional resources. BP staff working in improvised settings or temporary setups could have needed more time to prepare sessions or arrange different activities during a session. A second interpretation is that BPs in improvised settings or temporary setups might generally be less likely to apply curricula or fixed schedules. Lower

ratings in the dimension *structuring of a session* could thus reflect a larger amount of flexible session times. For both interpretations, it is important to consider that ECE staff in Germany had little previous experience working with refugee families (Chwastek et al., 2021). Here, the superior structural premises of education settings could have better supported BP organizers to plan, structure, and flexibly adapt sessions to unexpected demands by the refugee children. In improvised settings or temporary setups, however, this could be overall more challenging and reflected in the lower BREVIS ratings.

Unlike structural quality, BPs yielded moderate to high process quality with only a few links to the implementation types. Levels of process quality were comparable to center-based ECE programs in Germany or even slightly better. There are several explanatory approaches to this finding concerning group characteristics, teacher involvement, and ECE dosage. First, BPs demonstrated a good staff–child ratio with small groups and high staff involvement. Such group characteristics are considered important preconditions for better process quality (Pianta et al., 2005; Slot et al., 2015). Second, the consistently high staff involvement observed among heterogeneous BPs could have contributed to the high process quality. That interpretation is supported by the positive correlations between ratings on the dimension "teacher involvement" with several CLASS dimensions in our data. A study by Singer et. al. (2014) further supports this interpretation as they accordingly found links between higher teacher involvement and better process quality. Third, staff in BPs could show less fatigue than staff in other state-subsidized ECE services because BPs offer a lower ECE dosage, on average. Still, we did not include staff mental health data in our study that could support this third explanatory approach.

We found one link between a process quality dimension with the implementation type. The CLASS dimension of *productivity* (e.g., effectively transitioning between activities) was rated lower for improvised settings. There are two, not mutually exclusive, interpretations for this finding. First, BPs in improvised settings had more flexible concepts and thus put less emphasis on *productivity*. Specifically, BPs in improvised settings might be less likely to prepare sessions in advance and establish re-occurring procedures to facilitate transitions. Second, BPs in improvised settings could have had more difficulties retaining families for an extended period. Consequently, new refugee children and families entered the BPs and needed to learn and adapt to the group routines. Beyond anecdotal evidence from study phase 1, the second interpretation is also backed by a study that identified challenges in ECE with refugee children, namely infrequent attendance, tardiness, and fluctuation of refugee children attending BPs (Busch et al., 2018).

Our findings of overall high process quality are noteworthy in the light of previous evidence on links between staff professional training and better process quality (Slot et al., 2015). While we found overall high process quality in BPs, the staff, on average, had low levels of training, and volunteers usually supported the trained staff. Two different explanations could account for our findings, which seemingly oppose the previous evidence. First, the links between professional training and process quality are stronger for instructional domains of staff–child interactions (Pelatti, 2016). These domains were, however, not prioritized in many BPs and were not investigated in our study. Second, professional staff training could be less critical for achieving high process quality among the BPs. The study by Pelatti (2016) suggests a moderating effect of the staff–child ratio. In their study, the link between staff professional training and high process quality was

stronger for larger groups with few staff involved (likely reflecting groups with a relatively worse staff–child ratio, predominant large group activities, or low levels of teacher involvement). Among the BPs, however, we generally found small group sizes, that staff is primarily engaged in one-on-one or small group activities with children, and that staff has high levels of teacher involvement.

Considering our findings on structural and process quality together, BPs in education and improvised settings differ in their temporal frameworks and predictable procedures, routines, and rituals. Those differences are reflected in the interrelated domains *structuring of a session* (BREVIS) and *productivity* (CLASS). From integrative perspectives, these findings might reflect a specific characteristic of adaptive ECE programs designed for refugee children and families: the staff must balance efforts to establish routines and structures versus flexibly reacting to the individual needs of the diverse participants. BPs with different implementation strategies likely differ in approaching this conflict of goals. Consistently, 'establishing flexible routines' was described by Swedish ECE staff as a critical yet challenging strategy to prepare young refugee children for transitioning into preschool, kindergarten, or first grade (Lunneblad, 2017). Still, our findings and interpretations require further study to better understand the ECE quality across the heterogeneous implementation strategies and to posit evidence-based policy recommendations subsequently.

High process quality in BPs could support refugee children's socio-emotional adjustment

The high process quality of BPs might foster refugee children's social-emotional adjustment. Previous studies demonstrated links between high process quality in ECE programs and better child-related outcomes (Burchinal et al., 2008; Mashburn et al., 2008; Slot et al., 2015). Precisely, direct and indirect effects of ECE could convey the impact of high process quality on refugee children in BPs. Regarding the direct effects, high process quality provides refugee children with stimulating interactions that especially serve their social-emotional needs. Considering the target group, the high process quality of BPs could mitigate increased levels of child behavior problems among newly arrived refugee children (Buchmüller et al., 2018; Busch et al., 2021). Regarding the indirect effects, high process quality in BPs could facilitate trustful relationships of ECE staff with refugee families and provide them with ECD- and education-related information (Busch & Leyendecker, 2019). Both indirect effects were previously described for refugee families attending transitional ECE services in Canada (Poureslami et al., 2013). Further evidence is necessary to substantiate such links of high process quality to child-level outcomes among young refugee children.

Heterogenous Bridging Projects could follow different program concepts

The implementation heterogeneity among BPs underlines that these ECE programs differ from other policy-based ECE services in Germany in several regards. First, BPs were designed to address the developmental and educational needs of a specific target group characterized by their living circumstances, i.e., children from refugee families during post-migration periods (see Busch et al., 2018). Second, state subsidies for BPs were not tied to certain ECE standards or specific implementation settings beforehand. The implementation heterogeneity among BPs could hence mirror different ECE concepts

under the flexible regulatory policy. Slot et. al. (2017) consistently found that structural characteristics in ECE link to program concepts. BPs in settings for education could bridge a demand for transitional ECE services, such as during the kindergarten year before the transition to first grade. With better structural quality, they might focus on the direct effects of ECE in their concepts (e.g., fostering ECD and facilitating educational transitions). Beyond, BPs in improvised settings and with mobile concepts might be adaptive outreach work to overcome contact barriers and initiate trust in ECE providers among diverse refugee families (Morantz et al., 2013; Quintero, 1999). In such BPs, structural quality would be less relevant because they focus primarily on the indirect effects of ECE (e.g., provide information on ECE and education systems).

Moreover, we overall found weak to moderate correlations between structural quality and process quality dimensions, with only a few significant links. Previous evidence on other ECE programs supported links between structural and process quality, yet also with substantial between- and within-study variability among such links (Cabell et al., 2013; Singer et al., 2014). Again, different ECE program concepts (especially regarding activities and program settings) could account for such variability. Singer and colleagues observed process quality among playgroups and found weak links only. Cabell et. al. (2013) studied different ECE classroom-based situations and found stronger but variable links. In their study, stronger links emerged for the instructional domains of process quality when observed during pre-academic learning activities in large group settings. Many BPs of our sample seemed overall more similar to playgroups with an emphasis on social-emotional support. Given the previous evidence, our study findings support the idea that process quality in social-emotional support domains could be more invariant to structural quality characteristics, implementation settings, and ECE concepts.

ECE quality assessments contribute to ECE policy development

Effective ECE policy development fosters the impact of ECE programs through assuring high quality standards (Melhuish & Gardiner, 2019). Structural characteristics are, therefore, the best regulable determinants of ECE and thus often the main subject of policy regulations. Still, the challenge of measuring structural quality among heterogeneous BPs we faced in our study mirrors a general issue of ECE policy regulations: how to set adequate quality standards across arrays of different ECE programs (see Melhuish, 2016)? While most of the research on ECE quality has focused on center-based preschool programs, other ECE program types lack empirical investigations regarding ECE quality, especially those in improvised settings. However, adequate measurement of ECE quality would require the consideration of implementation settings (i.e., environments and concepts), at least during the interpretation of the observation results. The dearth of proper assessment tools is especially critical for specific target groups, such as newly arrived refugee families, because they likely require more adaptive ECE services (Lunneblad, 2017; Morantz et al., 2013). Thus, advancing ECE quality assessments among diverse ECE services is essential; it enables researchers and policymakers to better understand the links between ECE policies, implementation strategies, and program impact for children in diverse living circumstances.

Limitations and future research

Some methodological challenges and limitations of our investigation should be considered. We did not use sample stratification. In consequence, the sample was imbalanced across the different implementation types. Additionally, we could not examine a potential participation bias, e.g., that BPs with lower ECE quality were less likely to participate in our study. Our observation tool for structural quality (BREVIS) requires further validation in subsequent studies on heterogeneous ECE services. For process quality, the selected CLASS dimensions narrow the focus on social-emotional support and language modeling in BPs. However, some BPs in education settings are likely to address children's pre-academic learning skills as well. Subsequent studies should consider adaptive ECE programs' quality and implementation characteristics to predict young refugee children's socio-emotional development and language acquisition.

We did not consider systematic information on the BPs' conceptualizations, context-related premises and challenges, while BPs were differently implemented regarding the local circumstances and refugee children's needs. In-depth analyses of BPs across different implementation types could be a starting point to further exploring our interpretations, e.g., understanding the links between program concepts, ECE quality characteristics, and implementation strategies.

Conclusion

Our study provides a new assessment tool for ECE quality and offers insights into policy-based ECE programs for recently arrived refugee children characterized by heterogeneous implementation settings. We found preliminary evidence that adaptive ECE programs might not necessarily require strict regulatory policies and education settings to achieve acceptable ECE quality. Still, findings also suggest that program settings matter for ECE quality. Given the implementation heterogeneity of the BPs, such specialized ECE programs cannot generally provide a compensating alternative to ECE programs at daycare centers. However, the strength of the BPs is that they offer adaptive ECE services for post-resettlement contexts and that many of the programs are easily accessible for refugee children and their families. The BPs can thus inspire ECE stakeholders in refugee-hosting countries to set up initial actions at scale for mitigating the detrimental impact of refugee experiences during the early years of life.

Declarations

References

Agache, A., Kohl, K., Bihler, L., Willard, J., & Leyendecker, B. (2018, February). Kita-Qualität: Neue Einblicke in ein dynamisch-multikausales System aus Sicht der Netzwerkanalyse. In K. Blatter, A. Agache, & K. Kohl (Chairs), *Kita-Förderung, Kita-Qualität und Sprachkompetenzen von Kindern – erste Ergebnisse der SEIKA-NRW-Studie*. 6th Convention of the "Gesellschaft für empirische Bildungsforschung", Basel, Schweiz.

Anders, Y. (2013). Stichwort: Auswirkungen frühkindlicher institutioneller Betreuung und Bildung. *Zeitschrift Für Erziehungswissenschaft, 16*(2), 237–275. https://doi.org/10.1007/s11618-013-0357-5

Anders, Y., Grosse, C., Rossbach, H.-G., Ebert, S., & Weinert, S. (2013). Preschool and primary school influences on the development of children's early numeracy skills between the ages of 3 and 7 years in Germany. *School Effectiveness and School Improvement, 24*(2), 195–211. https://doi.org/10.1080/09243453.2012.749794

BAMF [Bundesamt für Migration und Flucht]. (2017). *Auszüge aus dem Ausländerzentralregister für 2017*. Bundesamt für Statistik.

Betancur, L., Maldonado-Carreño, C., Votruba-Drzal, E., & Bernal, R. (2021). Measuring preschool quality in low-and middle-income countries: Validity of the ECERS-R in Colombia. *Early Childhood Research Quarterly, 54*, 86–98.

Bihler, L.-M., Agache, A., Kohl, K., Willard, J. A., & Leyendecker, B. (2018). Factor analysis of the classroom assessment scoring system replicates the three domain structure and reveals no support for the bifactor model in German preschools. *Frontiers in Psychology, 9*, 1232. https://doi.org/10.3389/fpsyg.2018.01232

Bradley, R. H., Corwyn, R. F., Burchinal, M., McAdoo, H. P., & García Coll, C. (2001). The home environments of children in the United States Part II: Relations with behavioral development through age thirteen. *Child Development, 72*(6), 1868–1886.

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. (2019). Mental health needs of refugee children in specialized early education and care programs in Germany. *Journal of Immigrant and Minority Health*. https://doi.org/10.1007/s10903-019-00896-4

Burchinal, M. (2018). Measuring early care and education quality. *Child Development Perspectives, 12*(1), 3–9.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., et al. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction. *Applied Developmental Science., 12*, 140–153. https://doi.org/10.1080/10888690802199418

Burchinal, M. R., Roberts, J. E., Riggins, R., 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.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*(2), 166–176. https://doi.org/10.1016/j.ecresq.2009.10.004

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*, 1621. https://doi.org/10.3389/fpsyg.2018.01621

Busch, J., Cabrera, N., Ialuna, F., Buchmüller, T., & Leyendecker, B. (2021). Refugee children's early development during attendance of specialized preschool programs and transition into first grade in Germany. *Early Education and Development*. https://doi.org/10.1080/10409289.2021.1970427

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

Cabell, S. Q., DeCoster, J., LoCasale-Crouch, J., Hamre, B. K., & Pianta, R. C. (2013). Variation in the effectiveness of instructional interactions across preschool classroom settings and learning activities. *Early Childhood Research Quarterly, 28*(4), 820–830. https://doi.org/10.1016/j.ecresq.2013.07.007

Castro, D. C., Páez, M. M., Dickinson, D. K., & Frede, E. (2011). Promoting language and literacy in young dual language learners: Research, practice, and policy: Promoting language and literacy in young DLLs. *Child Development Perspectives, 5*(1), 15–21. https://doi.org/10.1111/j.1750-8606.2010.00142.x

Chwastek, S., Leyendecker, B., Heithausen, A., BalleroReque, 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*, 1621. https://doi.org/10.3389/fpsyt.2020.574412

Commerford, J., & Hunter, C. (2017). *Principles for high quality playgroups: Examples from research and practice.* Australian Institute of Family Studies.

Commerford, J., & Robinson, E. (2016). *Supported playgroups for parents and children: The evidence for their benefits. Child Family Community Australia, Report No. 40.* Australian Institute of Family Studies.

Erdemir, E. (2021). Summer preschools for Syrian refugee and host community children in Turkey: A model of contextually sensitive early intervention. *Early Education and Development.* https://doi.org/10.1080/10409289.2021.1961426

Fazel, M., Reed, R. V., Panter-Brick, C., & Stein, A. (2012). Mental health of displaced and refugee children resettled in high-income countries: Risk and protective factors. *The Lancet, 379*(9812), 266–282. https://doi.org/10.1016/S0140-6736(11)60051-2

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., & Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. *The Lancet, 369*(9555), 60–70. https://doi.org/10.1016/S0140-6736(07)60032-4

Hancock, K., Lawrence, D., Mitrou, F., Zubrick, S., Zarb, D., Berthelsen, D., & Nicholson, J. (2012). The association between playgroup participation, learning competence and social-emotional wellbeing for children aged four-five years in Australia. *Australasian Journal of Early Childhood, 37*(2), 72–81.

Harms, T., Clifford, R., & Cryer, D. (2015). *Early childhood environment rating scale* (3rd ed.). Teachers College Press.

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

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

Hurley, J. J., Saini, S., Warren, R. A., & Carberry, A. J. (2013). Use of the Pyramid Model for supporting preschool refugees. *Early Child Development and Care, 183*(1), 75–91. https://doi.org/10.1080/03004430.2012.655242

Hurley, J. J., Warren, R. A., Habalow, R. D., Weber, L. E., & Tousignant, S. R. (2014). Early childhood special education in a refugee resettlement community: Challenges and innovative practices. *Early Child Development and Care, 184*(1), 50–62. https://doi.org/10.1080/03004430.2013.769214

Jackson, D. (2013). Creating a place to "be": Unpacking the facilitation role in three supported playgroups in Australia. *European Early Childhood Education Research Journal, 21*(1), 77–93. https://doi.org/10.1080/1350293X.2012.760345

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

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

Lee, L., Halpern, C. T., Hertz-Picciotto, I., Martin, S. L., & Suchindran, C. M. (2006). Child care and social support modify the association between maternal depressive symptoms and early childhood behaviour problems: A US national study. *Journal of Epidemiology and Community Health, 60*(4), 305–310. https://doi.org/10.1136/jech.2005.040956

Lunneblad, J. (2017). Integration of refugee children and their families in the Swedish preschool: Strategies, objectives and standards. *European Early Childhood Education Research Journal, 25*(3), 359–369. https://doi.org/10.1080/1350293X.2017.1308162

Marti, M., Merz, E. C., Repka, K. R., Landers, C., Noble, K. G., & Duch, H. (2018). Parent involvement in the getting ready for school intervention is associated with changes in school readiness skills. *Frontiers in Psychology, 9*, 759. https://doi.org/10.3389/fpsyg.2018.00759

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

Melhuish, E. (2016). Longitudinal research and early years policy development in the UK. *International Journal of Child Care and Education Policy, 10*(1), 3. https://doi.org/10.1186/s40723-016-0019-1

Melhuish, E., & Gardiner, J. (2019). Structural factors and policy change as related to the quality of early childhood education and care for 3–4 year olds in the UK. *Frontiers in education.* (Vol. 4). Frontiers Media SA. https://doi.org/10.3389/feduc.2019.00035

MKFFI [Ministerium für Kinder, Familie, Flüchtlinge und Integration]. (2016). *Grundsätze zur Bildungsförderung für Kinder von 0 bis 10 Jahren in Kindertagesbetreuung und Schulen im Primarbereich in Nordrhein-Westfalen* (pp. 1–148). Ministerium für Kinder Familie Flüchtlinge und Integration des Landes Nordrhein-Westfalen.

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 & 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

New, R., Guilfoyle, A., & Harman, B. (2015). Children's school readiness: The experiences of African refugee women in a supported playgroup. *Australasian Journal of Early Childhood, 40*(1), 55–62. https://doi.org/10.1177/183693911504000108

NICHD [National Institute of Children Health and Human Development]. (2006). *The NICHD study of early child care and youth development (SECCYD): Findings for children up to age 4 1/2 years.* Government Printing Office. Retrieved from https://www.nichd.nih.gov/sites/default/files/publications/pubs/documents/seccyd_06.pdf

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* (pp. 1–70). Migration Policy Institute.

Pelatti, C. Y. (2016). Examining quality in two preschool settings: Publicly funded early childhood education and inclusive early childhood education classrooms. *Child youth care forum.* Springer.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., et al. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child–teacher interactions? *Applied Developmental Science, 9*, 144–159. https://doi.org/10.1207/s1532480xads0903_2

Pianta, R. C., LaParo, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system manual: Pre-K.* Brookes.

Poureslami, I., Nimmon, L., Ng, K., Cho, S., Foster, S., & Hertzman, C. (2013). Bridging immigrants and refugees with early childhood development services: Partnership research in the development of an effective service model. *Early Child Development and Care, 183*(12), 1924–1942. https://doi.org/10.1080/03004430.2013.763252

Quintero, E. (1999). The new faces of head start: Learning from culturally diverse families. *Early Education & Development, 10*(4), 475–497. https://doi.org/10.1207/s15566935eed1004_4

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

Russo, V., Barata, C., Joana, A., Leitao, C., & de Sousa, B. (n.d.). Assessing playgroup quality: Development and validation of the playgroup environment rating scale (PERS). (**Unpublished manuscript**).

Sammons, P., Sylva, K., Melhuish, E., Siraj, I., Taggart, B., Toth, K., & Smees, R. (2014). *Influences on students' GCSE attainment and progress at age 16: Effective pre-school, primary & secondary education project* (pp. 1–96). Department for Education.

Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant benefits: The high/scope perry preschool study through age 27*. High/Scope Press.

Sincovich, A., Gregory, T., Harman-Smith, Y., & Brinkman, S. A. (2019). Exploring associations between playgroup attendance and early childhood development at school entry in Australia: A cross-sectional population-level study. *American Educational Research Journal*. https://doi.org/10.3102/0002831219854369

Singer, E., Nederend, M., Penninx, L., Tajik, M., & Boom, J. (2014). The teacher's role in supporting young children's level of play engagement. *Early Child Development and Care, 184*(8), 1233–1249. https://doi.org/10.1080/03004430.2013.862530

Slot, P. L., Boom, J., Verhagen, J., & Leseman, P. P. M. (2017). Measurement properties of the CLASS Toddler in ECEC in The Netherlands. *Journal of Applied Developmental Psychology, 48*, 79–91. https://doi.org/10.1016/j.appdev.2016.11.008

Slot, P. L., Leseman, P. P., Verhagen, J., & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Childhood Research Quarterly, 33*, 64–76. https://doi.org/10.1016/j.ecresq.2015.06.001

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thomason, A. C., & Paro, K. M. L. (2009). Measuring the quality of teacher–child interactions in Toddler child care. *Early Education and Development, 20*(2), 285–304. https://doi.org/10.1080/10409280902773351

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

UNESCO, UNICEF, Brookings Institution, & World Bank. (2017). *Overview: Measuring early learning quality and outcomes (MELQO)*.

Votruba-Drzal, E., Coley, R. L., Collins, M., & Miller, P. (2015). Center-based preschool and school readiness skills of children from immigrant families. *Early Education and Development, 26*(4), 549–573. https://doi.org/10.1080/10409289.2015.1000220

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

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

Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Daneri, P., Green, K., Cavadel, E. W., Tarullo, L., Burchinal, M., & Martinez-Beck, I. (2016). Quality thresholds, features, and dosage in early care and education: Introduction and literature review. *Monographs of the Society for Research in Child Development, 81*(2), 7–26. https://doi.org/10.1111/mono.12236

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.