Implementing aeioTU: quality improvement alongside an efficacy study—learning while growing

Milagros Nores, 1 Alexandra Figueras-Daniel, 2 Maria Adelaida Lopez, 3 and Raquel Bernal 4

1 National Institute for Early Education Research, New Brunswick, New Jersey. 2 Teachers College, Columbia University, New York, New York. 3 Fundación Carulla, Bogotá, Colombia. 4 Facultad de Economía, Universidad de los Andes, Bogotá, Colombia

Address for correspondence: Milagros Nores, National Institute for Early Education Research, 73 Easton Ave., New Brunswick, NJ 08901. mnores@nieer.org

Effectiveness trials of increasing childhood development interventions across low- and middle-income countries have shown significant variability. The strength and consistency of benefits for children are dependent on program quality, and this requires paying attention to program implementation. In this paper, we summarize findings on program quality and teacher practices and perceptions for the aeioTU program, a center-based Reggio-inspired program in Colombia, now serving more than 13,000 children. The research found engaged, committed staff who valued the emergent approach and understood the children as requiring opportunities to express themselves, being the source for the curriculum, and having relationships with the materials around them. Although the average classroom quality was low in 2011, it increased significantly by 2014, particularly in the language and reasoning and interactions items. Indicator-level analyses showed that higher-order interactions and language processes were observed in a large proportion of classrooms by 2014. Teachers’ self-reports on the environment and their teaching and learning showed high levels of quality by 2013. These findings illustrate the significance of process data for program improvement, especially when a program is young. Program quality can be raised after teachers improve their skills, have experience enacting a curriculum, and after training has been strengthened in response to information, while simultaneously scaling up the program.

Keywords: early childhood; preschool; implementation; Colombia; quality

Introduction

A growing body of research has shown the importance of early interventions to ameliorate risks, reduce disparities, and improve child developmental trajectories.1,2 Effectiveness trials across low- and middle-income countries (LMICs) have shown variations in effects across children’s various developmental dimensions.3,4 The strength, consistency, and persistency of the benefits children (and families) receive from the variety of early interventions that have been tried in the early years are strongly dependent on quality.2,4 While quality may mean different things in different contexts,5 quality requires that attention is paid to the coherence of a program, its content, its participants, trainings, monitoring, and evaluation (see Ref. 6), among other things. Understanding the quality of a program also requires understanding its implementation, e.g., any course corrections done early on, whether it was evaluated, how data were used, and what changes were made, for scalability and replication (see Ref. 7). Our study discussed in this paper is unique in that it reports a process evaluation in a program as it was growing to scale, with the goal of informing the scale-up in a continuous improvement cycle.

Colombia advanced a comprehensive strategy on early childhood development (ECD) in 2011.6,8,9 There is a national strategy known as De Cero a Siempre (DCAS) aimed at increasing access and improving the quality of early childhood services provided to poor children. Its objective is to deliver high-quality integrated early childhood services for 1.2 million in poverty under age 6.10 This strategy includes partnerships with the private sector for
found that 8 months into treatment, a program called aeioTU had produced positive effects on language, cognitive, motor, and social development from low-quality center-based programs. Similarly, Attanasio et al. reported positive effects of a quality enhancement of a parenting program targeting poor parents of children younger than age 2. The quality improvement included the use of a structured early stimulation curriculum, along with pre-service and in-service training of paraprofessional home visitors. Stimulation and parenting interventions for children under 5 in low-income countries have been recently reviewed by Baker-Henningham and Lopez-Boo, who conclude that there were generally positive program effects on cognitive, socio-emotional, and nutritional development.

AeioTU’s effectiveness was investigated between 2010 and 2016, and its results are still in the process of being published. The study measured children’s anthropometry, language, cognitive, motor, and socio-emotional development, executive function, and parenting. Children were identified through a census in the catchment area for two centers opening in northern Colombia. Each aeioTU center in the study had capacity for about 320 children ages 0–5. Eligible children were underserved and poor. Families were randomized using computer generated random lists to assign them to either treatment or an ordered waiting list, stratified by age and gender. Children were first assessed in late 2010, prior to random assignment and to the centers starting operation, assessed again about 8 months’ post-treatment, and then yearly after that.

Nores et al. found that 8 months into treatment, the program had already produced positive effects on infants and toddlers of 0.16 SD (P = 0.006) for language, 0.08 SD (P = 0.086) for cognitive development, and P = 0.048 for overall development. Effects were observed only for girls. The effects for girls reached up to 0.30 SD in language. No intervention effects were observed for nutritional outcomes or all other dimensions of socio-emotional development and home environment. These findings make understanding the implementation of this program relevant, as no other program has shown comparable effects in such a short period of time. For example, effects of effective interventions in Colombia are 0.20–0.25 SD only after at least a year or more of intervention.

Our research and evaluation assessed processes with a twofold purpose: (1) to understand the quality of the program as implemented and (2) to support a process of continuous improvement in the program. We understood we were evaluating
the program “before it was proud” as defined by Campbell (1987, as cited in Ref. 20) and therefore saw our role as one that would not only measure program impact, but also program quality and inform program improvement as aeioTU expanded country-wide. Below, we provide an account of the results of the evaluation in terms of program quality and the practices, perceptions, and knowledge of teachers, as well as the measures for program improvement that were taken in response to the findings as part of such a continuous improvement cycle process (Fig. 1).

As is argued in Nores and Fernandez, 6 continuous improvement is a central piece for scaling up any program and for strengthening quality. Our discussion below focuses on program quality and teacher’s self-efficacy, which are central to effectiveness. 4,22,23 Researchers often make a distinction between process and structural quality. Structural quality refers to the resources and elements in the space and environment that facilitate processes for learning, activities, and others, such as class size, child–adult ratio, teacher qualifications, and physical environment. 8 Process quality is associated with the quality and types of teacher–child interaction and child–child interactions, as well as processes related to the use of time, enactment of a curriculum, integration of curriculum, and scaffolding of children’s learning, among others. 8 Measures of process quality are critical aspects that promote child development. 24,25 Use of observational Western measures of quality in non-Western countries is not without debate; 26 however, the expanding use of these measures is one source to understand quality, despite their limitations. 27,28

Context

Colombia has 65% of its 4.3 million children between age 0 and 5 living in socioeconomically disadvantaged conditions, based on the national census and Colombia’s proxy means test for social programming, known as SISBEN. 29 Therefore, it has made a strong effort to increase early childhood care and education services, which currently reach about 40% of children. 30,31 Low-income children in Colombia show developmental gaps as early as

Figure 1. Distribution of ECERS-R total scores, 2011 and 2014. Note: $N = 17$ in 2011 and $N = 30$ in 2014. A two-sample Kolmogorov–Smirnov test was used for equality of ECERS-R distribution functions ($P = 0.040$).
12 months, and in language these are about 1 standard deviation by the age of 5 relative to high-income children. In 2011, the government launched its DCAS national intersectoral early childhood strategy with the goal of providing quality comprehensive services for 1.2 million children. As part of this strategy, center-based care grew from serving about 125,000 children in 2011 to about 370,000 children in 2014. This initiative became institutionalized into a national law in 2016 (see Nores and Fernandez).

The aeioTU program

AeioTU has been part of the DCAS early childhood national strategy from the very beginning, partnering with the national government to serve children through direct service provision. Starting its operations in 2009, it now serves over 13,300 children across Colombia. In 2010, we initiated an impact study in two centers opening in the northern coastal region serving poor children (as measured by Colombia’s SISBEN). Aligned with this impact study, process measures were included to understand implementation and support aeioTU as it scaled up.

AeioTU is a Reggio Emilia “inspired” educational program, as Reggio Emilia does not define itself as a curriculum but rather a philosophy. The aeioTU program started by providing center-based care to low-income children ages 0–5; later, it began to provide home-visiting services to infants and toddlers, and center-based care to children ages 2–5 across most of its centers. AeioTU developed its own curriculum based on learning through strategies such as play, exploration, and projects, and which builds upon child-initiated activities and balances these with teacher-directed activities. A day in an aeioTU center sees a child moving between fully outfitted classrooms and experientially-based classrooms; the latter may focus on areas such as art, music, natural sciences, and light and shade, among others.

The Reggio Emilia approach encompasses a set of philosophical and pedagogical assumptions, methods of school organization, and principles of environmental design to foster children’s intellectual development. It is built upon a constructivist framework inspired by John Dewey, Jean Piaget, Lev Vygotsky, and Jerome Bruner. This model sees children as co-constructing knowledge through their interactions with each other, with adults, and with the environment. The curriculum, therefore, is emergent in nature, as it is based on the interests of the children and the goal of co-construction. A recent evaluation of Reggio Emilia in three cities in Italy showed little or no effects on children and adolescents in relation to other center-based care, and positive program effects in relation to not having attended center-based care.

Children in aeioTU receive 70% of their nutritional intake needs in morning and afternoon snacks and a lunch. The program is year-round (11 months, about 213 days) and serves children 9 h per day. All the centers are supported by a central aeioTU office in charge of operations and financing. Educational requirements have been higher, generally, than those required by Colombia’s national guidelines (technical or professional degree), and aeioTU provides preservice (about 120 h) and in-service training (about 130 h per year), which is quite uncommon in other center-based services in the country.

Training and development is planned year-round, mostly at the center, and lead by the center’s director, with central office support and using materials developed by the central office in conjunction with input from staff across various aeioTU centers. This is complemented with regional training. Training combines theory-based, didactic, individual and collective, in-person and online series of supports. Training is aligned with onsite individualized feedback and in-classroom observations.

Staffing is done by the central office. Each center included (at the beginning of our study) teachers and assistant teachers in ratios of 8:2 for infants, 12:2 for toddlers, and 24:2 for preschoolers; kitchen personnel (one person per 50 children); cleaning personnel (one person per 50 children); one security person per center; one administrative assistant to the center director; the center director (or Pedagogista); one nutritionist; and a resident artist (shared between two centers). AeioTU’s requirements for teachers’ qualifications were a technical degree (2 years of tertiary education) or undergraduate degree (5 years of tertiary education with a professional degree). However, the partnership with the Instituto Nacional the Bienestar Familiar (National Family Welfare Agency, ICBF), in charge of early childhood services in Colombia, threatened this requirement, as their guidelines provided an exception to hire madres comunitarias, women with lower qualifications who provided home-based care.
under the national program (the hogares comunitarios). Screening for hiring included personality tests, background checks, and health checks. Wages for teachers varied between US$240 and $630 per month before benefits, depending on location, experience, and degree at the time our evaluation started; these were aligned with national guidelines. Meals in the centers were also included as part of the work. Child:teacher ratios are now fixed at 25:1.5 for children ages 3–5, a change due to costs over time.

As part of internal implementation practices, aeioTU included several instruments to monitor children and support teachers. These included daily reports on children’s experiences, biweekly reports to families about the projects children engaged in, quarterly reports of children’s progress, a collection of children’s productions, a planning tool for teachers, observation tools to assess how children progress, and center documentation of children’s projects and processes.

**Methods**

Early in 2011 and in mid-2014, we conducted observations of classroom quality. Understanding the limitations of these measures, these were complemented with teacher surveys on their teaching practices, perceptions and knowledge, and their qualifications in 2012. The observations included classrooms in the effectiveness study, as well as other classrooms in aeioTU centers in other cities. Teacher surveys were only done with teachers in the two centers in the effectiveness study. The participation on teacher surveys was optional.

The education and experience survey was developed by the research team and was administered to collect information on the characteristics of the teachers in 2012. A total of 39 questions were asked ranging from teachers’ sociodemographic data to stress, job satisfaction, and beliefs about teaching young children. In parallel, we included the Patient Health Questionnaire-9 (PHQ-9) depression screener to detect depression in the teachers, which has been shown to be associated with program quality.

The ECERS-R (Early Childhood Environmental Rating Scale Revised) was administered in May 2011 and March–April 2014, in 3-hour visits to each classroom. ECERS-R is an observation scale through which routines, environment, and interactions that occurred in a preschool classroom are observed and coded. Quality of education and care has been found to be positively related to infant and child cognitive and social development, as well as to social-emotional skills. The ECERS-R provides a global measure of preschool classroom quality, with 43 items that cover a broad range of quality considerations from safety to teacher–child interaction to parent involvement. The total ECERS-R score is an average of the scores on the 43 items. A rating of 1 indicates inadequate quality, 3 indicates minimal quality, 5 indicates good quality, and 7 indicates excellent quality. This measure has been used extensively in the field and has well-established validity and reliability. It has been used in a wide range of countries with different cultures and economic contexts, including countries in the region similar to Colombia, for example Ecuador, Peru, Bolivia, Chile, and Brazil. We used the official Spanish translation of the ECERS-R. The ECERS-R is based on seven subscales: Space and Furnishings, Personal Care Routines, Language-Reasoning, Activities, Interaction, Program Structure, and Parents and Staff. Most of the items examine the quality of what children actually experience in the program. Assessors were trained by a group of expert observers from the National Institute for Early Education Research at Rutgers University. Observers were shadow scored in practice assessments until they reached 80% or higher reliability with trainers.

The Teacher Survey of Early Education Quality (TSEEQ), introduced in September 2012, is a self-report survey for ECD teachers on classroom practices. Teachers are asked to reflect on several aspects of the curriculum and classroom practices, including: literacy, math, science, physical education, art curriculum, curriculum in general, instruction, assessment, physical environment, interaction and emotional climate, leadership and supervision, and family involvement. There are 105 questions on the survey mostly on a 5-point Likert Scale, with a few exceptions that include questions on either a 3-point Likert Scale response or a yes/no response. A modified version was used with teachers in infant classrooms.

Below, we report descriptive analyses for the surveys; total scores and subscales for the ECERS-R and the Kolmogorov–Smirnov test for equality comparing distributions across years; and compare
Implementing aeioTU—learning while growing

Nores et al.

Environment Rating Scales (ERS) scores to other programs in the region.

In addition, we also report indicator-level descriptive analyses for the language and reasoning and the interactions subscales, more closely related to children’s learning and development, including their socio-emotional development. The ERS scores are composed of subscales, which in turn are composed of items, which in turn are composed of indicators. Indicators are arranged into four progressive levels including “inadequate”, “minimal”, “good”, and “excellent”. Observations entail a binary system of coding each of these indicators as either “occurring” or “not”, to arrive at a final score. This means that scores can be derived from the presence of various combinations of indicators met across different items; and so two classrooms could receive the same total score regardless of which specific indicators are met. If a higher order interaction or process is present but not co-occurring with another indicator under the same scoring level, the ERS score does not consider that level as attained, and instead a lower score is provided for that item. For example, item 16 “Encouraging children to communicate” is composed of nine indicators. Indicators 3.1, 3.2, and 3.3 (level 3) are required to co-occur for a “3” to be considered attained, and if only 50% of the indicators are observed the score is a “2” instead. No attention is paid to which specific indicators are met, only the number of them. If less than 50% is observed, then the score is a “1” and no indicator observed in this level or above is considered in the total score, even if observed. Indicator-level analyses look at the percentage of classrooms meeting the different indicators in the scale at any level. This type of indicator analyses allows capturing which higher order indicators are met and further promotes better guidance on what specifically needs to improve.

Results

Teacher sociodemographic characteristics

As part of our 2012 data collection, we included rich details about the sociodemographic characteristics of teachers, including education, qualifications, and working conditions. The majority of the teachers described themselves as “Mestizo.” Mestizo refers to a mixture of white and afrocolombian, which is the majority of self-reported ethnicity in the country. Also, at least 70% having specialized in ECD (N = 24; see Supplementary Table S1 online).

Most of the respondents to the survey were classroom or area teachers. Additionally, most of the teachers (63.6%) reported working 41–50 hours per week, and 20.5% reported working more hours than scheduled in a week. Although not shown in the tables, we computed the average annual gross salary of US$4,778 (converted at the average exchange rate of September 2012); and 100% reported having health insurance. The largest percent of teachers (72.7%) were working with large class sizes (21–25 children), but most (79.5%) reported working with one additional teacher in the classroom. This would follow the ratios established for infants, versus preschoolers versus area teachers (this information is reported in more detail in Supplementary Table S2 online).

Most teachers reported being satisfied with their salaries (72.7%), responsibilities (81.8% “do not agree with their job being difficult”), and working environment among colleagues (100.0% agreed “that personnel frequently share ideas among each other”). Similarly, most or all of the teachers reported feeling committed to the early childhood field (95.4%) and content with their career choice (i.e., 0.0% agreed with “feeling stuck in early childhood”). Generally, teachers enjoyed their work (97.7%), found their work relevant (79.6%), and appreciated the training and education opportunities offered to them (89.5%). These questions were asked in a 5-point Likert-type question that ranged between “strongly disagree” and “strongly agree” (summarized in Supplementary Table S3 online).

Central to the implementation of the Reggio philosophy is understanding and working with an emergent curriculum. The research team asked teachers their beliefs about an emergent curriculum. While 29.5% of teachers agreed that implementing the emerging curriculum is interesting but difficult, 61.3% of teachers disagreed with this statement. The biggest challenge reported by teachers refers to the requirement of documenting individual and daily work (59.2%). Other challenges, such as follow-up, time management, training, and alternative curricula, were not reported to be of concern. In general, 93.2% of teachers reported feeling that the emergent curriculum was interesting and valuable and worth the effort and time required to implement. Furthermore, teachers disagreed with “having a structured curriculum” (77.2%) or a more
Table 1. Emergent curriculum and beliefs about child development

| Emergent curriculum                                                                 | Disagree | Indifferent | Agree  | Missing |
|--------------------------------------------------------------------------------------|----------|-------------|--------|---------|
| It is very interesting but difficult to implement.                                   | 61.3%    | 4.5%        | 29.5%  | 4.5%    |
| It requires a lot of work because we must document individually and daily.          | 31.8%    | 4.5%        | 59.1%  | 4.5%    |
| It requires lots of follow up and there is not enough time for that.                | 75.0%    | 11.4%       | 9.1%   | 4.5%    |
| It is so interesting and valuable that it is worth the effort and time required.    | 2.3%     | –           | 93.2%  | 4.5%    |
| I do not completely understand how to elaborate a work plan within an emergent curriculum. | 90.9%    | 2.3%        | 2.3%   | 4.5%    |
| I do not feel sufficiently prepared or trained to work under this methodology.     | 95.4%    | –           | –      | 4.5%    |
| I feel it would be better to have a structured curriculum with preestablished goals and activities. | 77.2%    | 13.6%       | –      | 9.1%    |
| I feel my work would be easier with specific directives or content about how to work with children. | 63.7%    | 29.5%       | 2.3%   | 4.5%    |

| Beliefs about child development                                                     |          |             |        |         |
|--------------------------------------------------------------------------------------|----------|-------------|--------|---------|
| Children must have some control over their learning.                                 | 75.0%    | 2.3%        | 11.3%  | 11.4%   |
| Children should be able to develop through experiences of touch, movement, sound and vision. | 9.1%     | –           | 84.1%  | 6.8%    |
| Children are never wrong.                                                            | 68.2%    | 2.3%        | 22.7%  | 6.8%    |
| Children should have unlimited opportunities of expression.                         | 25.0%    | –           | 70.5%  | 4.5%    |
| The ideas of the child are the source of the curriculum.                             | 2.3%     | –           | 86.3%  | 11.4%   |
| Children have a relationship with materials around them that they should be allowed to explore. | –        | –           | 99.6%  | 2.3%    |

directive one (63.7%). Overall, teachers showed commitment with and understanding of the importance of the curriculum they were implementing. These questions were asked in a 5-point Likert-type question that ranged between “strongly disagree” and “strongly agree”, and are aggregated in Table 1.

The use of an emergent curriculum depends on teachers' mastery and understanding of child development. Table 1 also reports how teachers agreed or disagreed with statements about child development and learning. While teachers did not necessarily agree with children having “some control over their own learning process” (75%), they answered favorably to questions about children “having the opportunity to express themselves limitlessly”, with the ideas of the children being the source for the curriculum, and with children having a relationship with the materials around them and exploring these, all concepts embedded in the aeioTU model. These questions were asked in a 5-point Likert-type question that ranged between “strongly disagree” and “strongly agree” and were aggregated into three levels.

The final section of the survey asked teachers questions about stress and depression. These factors are important as they can affect teacher demeanors and approach to working with young children. Some teachers reported feeling the symptoms of stress and depression. The PHQ-9 Scores though showed that the majority of the teachers (84.1%) did not report signs of depression.

Classroom observations

As mentioned above, the ECERS-R is an observation and rating instrument for preschool classrooms serving children aged 3–5; a rating of 1 indicates inadequate quality, 3 indicates minimal quality, 5 indicates good quality, and 7 indicates excellent quality. The ECERS-R was completed on a total of 17 classrooms in 2011 and 30 classrooms in 2014. Cronbach’s for the scale were 0.93 in both years. Table 2 reports the mean, range, and 90th percentile

---

*Infant and toddler classrooms were also observed using the Infant and Toddler ERS measure (ITERS), but the sample was $N = 3$ in 2011, and they were not observed in 2014; therefore, they are excluded from this report.
for the ECERS-R and each of its subscales. In 2011, the program averaged a quality of 2.3, and in 2014 2.9. The language and reasoning and interaction activities were among the highest scoring, 2.04 and 3.16, respectively, and they grew to 2.80 and 3.89, respectively. These two areas focus on the activities related to child learning, and less so to program structure. The highest rated classrooms (90th percentile) scored 3.38 in 2011 and 3.61 in 2014. For language and reasoning, the 90th percentile was 3.50 in 2011 and increased to 4.10 by 2014. For interactions, the 90th percentile was 4.80 in 2011 and 5.72 in 2014.

After the first set of scores, aeioTU put in place a set of reforms (described in detail further below) to improve quality as the program was scaled up. These investments were reflected in the quality indicators. Figure 1 illustrates improvements in the distribution of quality in aeioTU classrooms, comparing 2011 and 2014. By 2014, aeioTU increased the levels of both lower and higher scoring classrooms, and resulted in good quality levels in several classrooms. The increase in the total score was statistically significant (Kolmogorov–Smirnov, \( P = 0.040 \)). For the subscales, there were statistically significant (or marginally so) improvements for personal care routines (\( P = 0.004 \)), language and reasoning (\( P = 0.093 \)), activities (\( P = 0.068 \)), interactions (\( P = 0.083 \)), and program structure (\( P = 0.001 \)).

While these scores are low relative to high-quality programs in developing countries, they are not low in relation to other programs using ERS scales in the region. Figure 2 compares the 2014 aeioTU ECERS-R scores with ERS scores from four other programs in the region. Higher quality is harder to find in the region, and the counterfactual to these programs is so low that their impact, as in the case of aeioTU, comes from providing a substantial improvement over alternative programs.

### Interactions and indicator-level data

Research has indicated that global measures of classroom quality, like the ECERS-R, may not be the most highly predictive of student outcomes and that, rather, interaction-specific constructs are more predictive of student outcomes than the total score. The ECERS-R does contain several items that are “interaction-specific.” For example, “Language and Reasoning” documents and quantifies evidence on specific teacher–child interactions that build on children’s vocabulary. This subscale has been noted by researchers as the most predictive subscale of the ECERS-R as related to child outcomes. The following descriptive analyses look into this and into the interactions subscales at a much more granular level, in order to describe which indicators were and were not observed in 2011 and how this changed over time. As part of a continuous improvement cycle with programs, knowing which indicators are met throughout concretely identifies areas for improvement—identifying the frequency with which classrooms are meeting each indicator under each “minimal, good, and excellent” level of quality. Such detail information provides specific information to support professional development efforts more strategically.

Figures 3–5 show that between 2011 and 2014 there were large increases in the percentage of classrooms meeting the indicators under each item. The indicators that rely on interactions improved significantly. There was growth in the frequency with which classrooms were meeting indicators, and also meeting the higher order indicators for each of the

### Table 2. ECERS-R total and subscales ratings for the aeioTU program

| ECERS-R          | Mean (Min-Max) | 90th % | Cr. alpha |
|------------------|----------------|--------|-----------|
| **Total**        | 2.30 (1.24–3.90) | 3.38   | 0.93      |
| **Space and furnishing** | 2.18 (1.25–4.25) | 3.05   | 0.73      |
| **Pers. care routines** | 1.85 (1.17–2.60) | 2.44   | 0.94      |
| **Language reasoning** | 2.04 (1.00–4.50) | 3.50   | 0.90      |
| **Activities**   | 1.74 (1.00–3.00) | 2.31   | 0.90      |
| **Interactions** | 3.16 (1.20–6.60) | 4.80   | 0.93      |
| **Program structure** | 1.60 (1.00–4.00) | 2.47   | 0.76      |
| **Parent and staff** | 3.88 (1.67–5.67) | 5.23   | 0.85      |
| **2011 (n = 17)** |                |        |           |
| **2014 (n = 30)** |                |        |           |
| **Mean (Min-Max)** | 2.89 (1.56–5.72) | 3.61   | 0.93      |
| **90th %**       |                |        |           |
| **Cr. alpha**    |                |        |           |
items (those in the “excellent” level). These findings are critical in light of research on the importance of these kinds of interactions being imperative for child development.52

Figure 3 shows an increase in staff who balanced listening and talking, and a fourfold increase in staff who linked children’s spoken communication with written language. These increases are not observed in the total score, as they would be held back by the percentages of classrooms attaining level 3 (“good”) indicators. This is aligned with growth in the percentage of classrooms where children talked about logical relationships and concepts, which also increased significantly, as well as the appropriate introduction of concepts (Supplementary Fig. S2 online). Modest growth was also observed in the percentage of classrooms where staff talked about logical relationships as children engage with materials, and in the percentage of classrooms where children are encouraged to explain their reasoning (Supplementary Fig. S2 online).

Figure 4 shows another example where indicators demonstrated large increases between 2011 and 2014 that would not be fully reflected in the total scores, as the increases were in some, but not all, the indicators under each level. For example, there was a large increase in the percentage of classrooms where language was centrally used with children and staff, encouraging communication between children, as well as in individual conversations with children.

Figure 5 focuses on staff–child interactions; indicators in this area were already met for a high percentage of the classrooms in 2011. Nevertheless, by 2014, all indicators were observed in 60% of the classrooms, and all but the last indicator was met by over 80% of the classrooms. Interactions among children (Supplementary Fig. S3 online), basically showing increases between 2011 and 2014, are also evident. In this case, almost all higher indicators were observed in more than over 60% of the classrooms.

A last set of indicators described has to do with discipline (Supplementary Fig. S4 online). This is another area where various higher order indicators were met by over half the classrooms, but the scores in the ERS were held back because some, but not
implementing aeioTU—learning while growing

nores et al.

Figure 3. Percentage of classrooms meeting indicators on encouraging children to communicate.

all, indicators were met. Staff react consistently to children’s behavior and involve children in solving conflicts and problems are two areas that are central to the aeioTU model.

Tseeq

The Tseeq asks questions related to the physical environment and routine aspects of teaching (including content areas), planning, interactions, and leadership to teachers; it was used in September of 2012. Out of 26 teachers, 14 completed the Tseeq (54%). This, a measure developed to capture self-reported classroom quality, is summarized by seven composites that vary between 1 (minimal quality) and 5 (high quality). Figure 6 reports Tseeq scores by subscale. Average quality across these was 4.10 (0.38 SD) over a 5-point scale. That is, self-reported quality for responding teachers as measured by the Tseeq was good, particularly in terms of assessment of children, teaching, and interactions. Internal consistency (Cronbach’s α) for the Tseeq total scores was 0.96 (0.64 for assessment, 0.56 for physical structure, 0.82 for family participation, 0.82 for teaching, 0.91 for planning, 0.78 for interactions, and 0.76 for leadership).

Discussion

The evaluation of the aeioTU program was conceived as utilization-based, combining both implementation and child outcome parts. The implementation evaluation was meant to provide information on how the program was working and information that could guide program improvement, particularly as it was progressively scaling up. This was central, as aeioTU was committed to scale and quality. The evaluation relied on a combination of qualitative and quantitative measures. The outcome evaluation focused on the impacts of the program on participating children, as well as on program improvement strategies, as it highlighted need for supports in specific areas. The implementation evaluation focused on processes, from understanding the teachers, the training they got, the quality of interactions in their classrooms, to their perception of their work. Therefore, it
Our report shows that quality, as measured by the ERS, was quite low in the first year of observations, and that the actions implemented by the program were meaningful enough to increase quality significantly over a period of 3 years. Analyses at the indicator level show that higher level indicators were present in a significant percentage of classrooms by 2014. Teacher surveys revealed that they had a strong commitment to the Reggio inspired emergent curriculum, while they felt challenged with the amount of documentation required to follow children's progress. Moreover, the survey on self-reported quality practices in early childhood (TSEEQ) showed that teachers' perceptions of their practices, and their access to materials, resources, and supports, were high by 2013.

The results of our study illustrate that when a program is young, process and outcome data for children can be central to inform program improvement. Quality can only be raised after teachers have had their skills improved and experience enacting a curriculum, and in this case, after trainings were strengthened in response to what was observed needing further support. Other necessary improvements may be structural. Overall, this means that quality may not be as high as expected initially, and that implementation evaluation of the programs and processes can support continuous quality improvement. Below, we discuss the program improvements that were triggered by this evaluation, as well as general lessons for the field.

**Program improvement**

AeioTU has a centralized system of supports with curricular leadership, which allowed for decisions to be made from the central administration to support not only the centers in the study but also to put in place a system to strengthen quality. The

| Indicator                                                   | 2011 | 2014 |
|-------------------------------------------------------------|------|------|
| Some staff/child conversation                               | 82   | 80   |
| Children allowed to talk much of the day                    | 88   | 87   |
| Many staff–child conversations during free play/routines   | 35   | 47   |
| Language primarily used to exchange info w/children         | 47   | -    |
| Staff add info to expand children's ideas                   | 12   | 17   |
| Staff encourage communication among children                | 29   | 53   |
| Staff have individual conversations with most children      | 0    | 27   |
| Staff ask questions to encourage longer/more complex answers| 6    | 17   |

Figure 4. Percentage of classrooms meeting indicators on informal use of language.
Implementing aeioTU—learning while growing

Nores et al.

Figure 5. Percentage of classrooms meeting indicators on staff–child interactions.

results of the implementation process evaluation summarized here, as well as the impact evaluation mentioned earlier, have been feeding since 2011 a continuous improvement process for aeioTU. There have been program improvements along two main sets of action based on the information provided by the evaluation: (1) changes in pre- and in-service training; and (2) program improvements in relation to structural supports, curriculum, child evaluation, and fidelity control systems. These improvements have occurred as aeioTU continued to progressively scale up, and now are part of the scaled aeioTU systems serving more than 13,000 children; and they continue to be used and revisited every year.

Improvements in pre- and in-service training. The results of the study were shared internally within the aeioTU teams of teachers and directors to initiate processes that would strengthen quality not only in the centers that were included in the evaluation but also across all the centers. The pre- and in-service training processes were reorganized as follows.

A professional development experience was implemented early in the school year (acercamiento a la experiencia educativa, which translates as “getting close to the education experience”) to revisit the key elements of implementing the aeioTU Reggio inspired experience. This emphasizes the central fundamentals and components of the experience and relates it to everyday pedagogical practices. This professional development experience is meant to refresh and strengthen the knowledge and skills of existing staff, and to build and integrate new staff.

A set of 11 professional development workshops were developed on (1) positive discipline in the aeioTU experience and the importance of relationships and interactions; (2) play as a learning strategy in the aeioTU experience; (3) pedagogical tools and their use in classrooms, sensory areas, and common areas; (4) assessment of children’s development and learning in the aeioTU experience; (5) sensory areas, their dynamics, intention, and contribution to the development and learning of children; (6) documentation as a learning tool in the aeioTU educational experience; (7) transition moments and their contributions to children’s learning and development; (8) thought and language abilities through the learning strategies in the aeioTU educational experience; (9) family and community
participation in the aeioTU educational experience; (10) promoting healthy lifestyles; and (11) the importance of generating habits. In addition,

- Other professional development experiences, such as gatherings and meetings were redefined and refined as needed; this also included further work on the preservice trainings established for all new centers opening as the operations scaled up;
- Internal pedagogical committees were established in all centers to assess the child monitoring and evaluation processes, as well as the implementation of learning strategies;
- Security processes were established to strengthen monitoring and observation processes for classrooms to assess risks related to the infrastructure and the appropriate use of materials; this was supported by protocols on personal care routines for children.

**Structural, curricular, and evaluation improvements.** Another set of decisions had to do with revisiting the adequacy of materials, as well as solidifying a curriculum to better scaffold teachers in centers and classrooms. After the first year’s process data became available, the support center put together a curriculum and pedagogy quality support team to work with the centers to support and monitor quality. As part of this support process, three additional quality support roles were added: zonal coordinators (one per department); regional pedagogical coordinators (one per region); one curriculum professional and two arts and culture professionals. This expanded quality team initiated eight activities.

The first, following findings in the ERS analysis, was the revision and subsequent improvement of learning materials and furniture in centers. In addition, quality control processes were organized and standardized across all schools and centers. Third, curriculum tools and strategies were revised in response to what was being observed in the impact evaluation in terms of children’s language and cognitive development. A software to monitor children’s development (ConecTU) was developed and implemented across all aeioTU centers. Internal observation and evaluation tools were created to assess the quality of processes in all centers inspired.
by the ERS measures (Faro Operación Sana and Faro Pedagógico). A toolbox (called Cartografía Curricular, see Supplementary Image S1 online) was developed to scaffold teaching and learning processes. This toolbox includes a book on how to create pedagogical spaces that promote exploration and learning; a book on learning strategies (e.g., thought abilities, play, exploration, and projects); a book on aeioTU at home (supports the home visiting program developed afterward serving mothers and children younger than age 2); a positive discipline tool; a catalog on pedagogical tools for the different ages; didactic tools for language and mathematical processing; a support plan for children with special needs; a planning tool for teachers; and tools to strengthen the use of transition times throughout the day and throughout the different spaces. In addition, leadership and process trainings were integrated across the systems in a systemic approach.

The understanding and appropriation of the documents and instruments of the Cartografía Curricular is done in in-person trainings, where teachers are trained on the material; its content; and what, why, and how to use it in practice. This is followed by onsite visits, observations, collaborative and one-on-one meetings, providing multiple opportunities for feedback. This is the core of the aeioTU program, which can be scaled to other regions.

Lessons learned
There are no specific curricular guidelines for early education in Colombia. The Board for Early Childhood has emphasized the principle of curricular freedom, and national standards are intentionally broad. That means that teachers and other providers of services are expected to adapt the learning standards to their own classrooms and contexts.\(^8\)

AeioTU is based on the concept of an emergent curriculum. Traditional teaching in Colombia is not. While teachers were highly receptive with the conceptualization and the model, they also appeared to struggle with elements of its implementation. Classroom observations revealed that some of the ERS scores were lower because children’s choices were sometimes controlled. Consequently, the aeioTU curriculum support staff invested in further scaffolding for their teachers in ways that published curriculums such as High Scope\(^53\) or Tools of the Mind\(^54\) commonly do, with a variety of tools and structure to help support the practice in the classrooms. This became a necessary ingredient in the development of a strong curriculum that could be enacted across all of the aeioTU centers. This is now the “toolbox” described earlier, published as a set, and is the core of the aeioTU model.

The process of understanding and working through the strengths and weaknesses of the model allowed the program to build a stronger system of supports for all its centers. This system now also provides professional development and technical assistance to other early education providers in Colombia, in partnership with the public sector or through market-based arrangements. Professional development and technical assistance are now central components of the aeioTU strategy to promote high-quality early education services in Colombia beyond their centers. Scaffolding their own model has also allowed Colombia to generate the inputs to support other ECD programs across Colombia and the region. The curriculum, the tools for monitoring progress of children, and the tools to support teachers’ processes in the classroom and with children are transferable across contexts.

An example of this work is the support and scaffolding provided in the State of Quintana Roo, Mexico, where aeioTU supported the processes necessary to start operations in two government-sponsored centers. Similarly, in 2015 and 2016, aeioTU provided training and professional development to 300 and 600 early childhood (non-aeioTU) centers in Colombia, as part of a quality improvement effort and in contract with the ICBF. Pre- and post-treatment evaluations of classroom quality showed classroom improvements in these programs, which serve 60,000 children.

AeioTU provides a case study of a program that progressively scale up while carefully increasing quality. Careful monitoring of quality and continuous strengthening of professional development and scaffolding provided to teachers is not unique to the aeioTU program. Other programs have implemented similar continuous improvement cycles as they have scale-up, particularly in developed countries. This is true of the Abbott preschool program in the United States in the state of New Jersey,\(^55,56\) which is now a top ranked preschool program in the United States.\(^57\) A study of professional development that combined didactic and coaching components in Chile’s Un Buen Comienzo showed similar effects in classroom quality (although these did not
translate into child outcomes). Ultimately, good quality early childhood interventions require time to mature, processes to measure what is working and what is not, and capacity to implement changes to improve quality continuously. Scaling while increasing quality, as aeioTU did and can do, requires robust processes of assessment, information, and responsive decision-making.

Acknowledgments

This paper was invited to be published individually and as one of several others as a special issue of Ann. N.Y. Acad. Sci. (1419: 1–271, 2018). The special issue was developed and coordinated by Aisha K. Yousafzai, Frances Aboud, Milagros Nores, and Pia Britto with the aim of presenting current evidence and evaluations on implementation processes, and to identify gaps and future research directions to advance effectiveness and scale-up of interventions that promote young children’s development. A workshop was held on December 4 and 5, 2017 at and sponsored by the New York Academy of Sciences to discuss and develop the content of this paper and the others of the special issue. Funding for open access of the special issue is gratefully acknowledged from UNICEF and the New Venture Fund.

We are very thankful to aeioTU for their commitment to early childhood and opening the doors to our evaluation team; the Jacobs Foundation and the UBS Optimum Foundation for their generous grants that made this evaluation possible; the Inter-American Development Bank for their financial contribution; iQuartil for their wonderful work managing data collection on site; and our data collectors who worked under very difficult conditions. Any views expressed are those of the authors and do not necessarily represent those of the funders. This research was approved by the IRB at Rutgers University (USA; Internal Review Board, Protocol No. E09-568c) and Universidad de los Andes (Colombia; Acta No. 032–2009).

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1. Continuous quality improvement cycle.
Figure S2. Percentage of classrooms meeting indicators on interactions among children.
Figure S3. Percentage of classrooms meeting indicators on interactions among children.
Figure S4. Percentage of classrooms meeting indicators on discipline.
Table S1. Teacher ethnicity and qualifications.
Table S2. Work status.
Table S3. Working environment.
Image S1. Cartografía curricular.

Competing interests

The authors had financial support from the Jacobs Foundation, the UBS Optimum Foundation, and the Inter-American Development Bank for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work.

References

1. Black, M.M., S.P. Walker, L.C.H. Fernald, et al. 2017. Early childhood development coming of age: science through the life course. Lancet 389: 77–90.
2. Marope, P.T.M. & Y. Kaga. 2015. Investing against evidence: the global state of early childhood care and education. UNESCO Publishing, New York. Accessed February 8, 2018. http://unesdoc.unesco.org/images/0023/002355/233558E.pdf.
3. Nores, M. & W.S. Barnett. 2010. Benefits of early childhood interventions across the world: (under) investing in the very young. Econ. Educ. Rev. 29: 271–282.
4. Britto, P.R., S.J. Lye, K. Proulx, et al. 2017. Nurturing care: promoting early childhood development. Lancet 389: 91–102.
5. Barnett, W.S. & M. Nores. 2015. Investment and productivity argument for ECCE. In Investing Against Evidence: The Global State of Early Childhood Care and Education. P.T.M. Marope & Y. Kaga, Eds.: 73. New York: UNESCO Publishing.
6. Nores, M. & C. Fernandez. 2018. Building capacity in health and education systems to deliver interventions that strengthen early child development. Ann. N.Y. Acad. Sci. 1419: 57–73.
7. Yousafzai, A., F.E. Aboud, M. Nores & R. Kaur. 2018. Reporting guidelines for implementation research on nurturing care interventions designed to promote early childhood development. Ann. N.Y. Acad. Sci. 1419: 26–37.
8. Yoshikawa, H. & S.B. Kabay. 2015. The evidence base on early childhood care and education in global contexts (background paper for the Education for All Global Monitoring Report 2015). UNESCO, Paris. Accessed February 8, 2018. http://unesdoc.unesco.org/images/0023/002324/232456e.pdf.
Implementing aeioTU—learning while growing

Nores et al.

9. Vargas-Barón, E. 2015. Institutional frameworks and governance for early childhood systems: multisectoral coordination and integration. In Investing Against Evidence: The Global State of Early Childhood Care and Education. P.T.M. Marope & Y. Kaga, Eds.: 269. New York: UNESCO Publishing.

10. Comisión Inter-Sectorial para la Primera Infancia. 2013. De Cero a Siempre: Atención Integral a la primera infancia. Estrategia de Atención Integral a la Primera Infancia. Fundamentos políticos, técnicos y de gestión. Colombia. Presidencia de la República. Accessed February 8, 2018. http://www.decoraosiempre.gov.co/QuienesSomos/Documents/Fundamentos-politicos-tecnicos-gestion-de-cero-a-siempre.pdf.

11. Kagitcibasi, C., S. Bekman & A. Goksel. 1995. A multi-purpose model of nonformal education: the mother–child education programme. Early Childhood Counts: Programming Resources for Early Childhood Care. The Consultative Group on ECCD, Canada. Accessed February 8, 2018. https://www.researchgate.net/publication/254197817_A_MULTIPURPOSE_MODEL_OF_NONFORMAL_EDUCATION_THE_MOTHER-CHILD_EDUCATION.

12. Chang, S.M., S.P. Walker, S. Grantham-McGregor & C.A. Powell. 2002. Early childhood stunting and later behaviour and school achievement. J. Child Psychol. Psychiatry 43: 775–783.

13. Engle, P.L., L.C. Fernald, H. Alderman, et al. 2011. Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. Lancet 378: 1339–1353.

14. Bernal, R., O.P. Attanasio, X. Peña, et al. 2014. The effects of the transition from home-based childcare to center-based childcare in Colombia. Universidad de los Andes and London, UK, Institute for Fiscal Studies, Bogotá, Colombia. Accessed February 8, 2018. https://ideas.repec.org/p/col/000089/014067.html.

15. Attanasio, O.P, H. Baker-Henningham, R. Bernal, et al. 2017. Early stimulation: the impacts of a scalable intervention. Universidad de los Andes, Bogotá, Colombia.

16. Baker-Henningham, H. & E. López Bóo. 2010. Early childhood stimulation interventions in developing countries: a comprehensive literature review. Inter-American Development Bank, Washington, DC. Accessed February 8, 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1700451.

17. Nores, M., R. Bernal & W.S. Barnett. Center-based care for infants and toddlers: the aeioTU randomized trial. Under review. New Brunswick, NJ: National Institute for Early Education Research & Bogotá, Colombia: Universidad de los Andes.

18. Bernal, R., C. Fernández, C.E., Flórez, et al. 2009. Evaluación de impacto del Programa Hogares Comunitarios de Bienestar del ICBF. Universidad de los Andes, Bogotá, Colombia. Documento CEDE no. 16. Accessed February 8, 2018.

19. Attanasio, O.P., C. Fernández, E.O. Fitzsimons, et al. 2014. Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial. BMJ 349: 5785.

20. McCall, R.B. 2009. Evidence-based programming in the context of practice and policy. SRCD Social Policy Report, Vol. 23, pp. 3–19.

21. Frede, E. 2005. Assessment in a continuous improvement cycle: New Jersey’s Abbott Preschool Program. Invited paper for the National Early Childhood Accountability Task Force with support from The Pew Charitable Trusts, the Foundation for Child Development, and the Joyce Foundation. Accessed February 8, 2018. http://nieer.org/docs.

22. Yoshikawa, H., C. Weiland, J. Brooks-Gunn, et al. 2017. Investing in our future: the evidence base on preschool education. Society for Research in Child Development, Washington, DC. Accessed February 8, 2018. https://www.srcd.org/sites/default/files/documents/washington/mb_2013_10_16_investing_in_children.pdf.

23. Camilli, G., S. Vargas, S. Ryan & W.S. Barnett. 2010. Meta-analysis of the effects of early education interventions on cognitive and social development. Teachers College Record, Vol. 112, pp. 579–620.

24. Araujo, M.C. & N. Schady. 2015. The early years. Child well-being and the role of public policy. Inter-American Development Bank, Washington, DC. Accessed February 8, 2018. https://steinhardt.nyu.edu/global-ties/early_childhood/melqo.

25. Aboud, F., N. Rao, H. Yoshikawa, et al. 2015. Measuring the early learning environment. Observational measure of quality for low and middle-income countries commissioned by the Measuring Early Learning and Quality Outcomes (MELQO) project (co-sponsored by the Brookings Institution, UNESCO, the World Bank, and UNICEF). Accessed February 8, 2018. https://steinhardt.nyu.edu/global-ties/early_childhood/melqo.

26. Aboud, F.E. 2006. Evaluation of an early childhood preschool program in rural Bangladesh. Early Child. Res. Q. 21: 46–60.

27. Leyva, D., C. Weiland, M. Barata, et al. 2015. Teacher–child interactions in Chile and their associations with prekindergarten outcomes. Child Dev. 86: 781–799.

28. Lopez boo, F., A.M. Caridad & R. Tomé. 2016 How is child care quality measured? A toolkit. Inter-American Development Bank, Washington, DC. http://pubdocs.worldbank.org/en/576511485965608183/How-to-measure-child-care-quality-WEBINAR-BBL-ELP-May-24.pdf.

29. Bernal, R. & A. Camacho. 2012. A early childhood policy in the context of equity and social mobility in Colombia. Universidad de los Andes, Bogotá, Colombia. Documento CEDE no. 33. Accessed February 8, 2018. https://economia.uniandes.edu.co/component/booklibrary/478/view/46/Documentos%20CEDE/131/la-politica-de-primer-infancia-en-el-contexto-de-la-equidad-y-movilidad-social-en-colombia.

30. World Bank. 2013. Colombia. Early childhood development. SABER country report. Washington, DC. The World Bank. Accessed February 8, 2018. http://documents.worldbank.org/curated/en/574501468028793407/SABER-early-childhood-development-country-report-Colombia-2013.

31. Bernal, R. & A. Camacho. 2011. La Importancia de los Programas de Primera Infancia en Colombia. Imprenta Nacional de Colombia, Bogotá, Colombia.
Implementing aeioTU—learning while growing

Nores et al.

Accessed February 8, 2018. https://economia.uniandes.edu.co/centros-de-investigacion/cede/web-de-proyectos-cede/primera-infancia/estudios/la-importancia-de-los-programas-para-la-primera-infancia-en-colombia.

32. Bernal, R., M.A. Martínez & C. Quintero. 2015. *Situación de Niñas y Niños Colombianos Menores de Cinco Años entre 2010 y 2013*. Bogotá, Colombia: Editorial Kimpres. Accessed February 8, 2018.

33. Edwards, C.P., L. Gandini & G.E. Forman, Eds. 1998. *The Hundred Languages of Children: The Reggio Emilia Approach—Advanced Reflections*. Greenwood Publishing Group.

34. Gandini, L. 1993. Fundamentals of the Reggio Emilia approach to early childhood education. *Young Children* 49: 4–8.

35. Dodd-Nufrio, A.T. 2011. Reggio Emilia, Maria Montessori, and John Dewey: dispelling teachers’ misconceptions and understanding theoretical foundations. *Early Child. Educ. J.* 39: 235–237.

36. Birolí, P., D. Del Boca, J.I., Heckman, et al. 2017. Evaluation of the Reggio approach to early education (No. w23390). IZA Institute of Labor Economics, Bonn, Germany. http://www.nber.org/papers/w23390.

37. Bernal, R., C. Fernández, C.E., Flórez, et al. 2009. Evaluación de impacto del Programa Hogares Comunitarios de Bienestar del ICBF. *Universidad de los Andes, Bogotá, Colombia. Documento CEDE* 16, 1–29. Accessed February 8, 2018. https://ideas.repec.org/p/col/000089/005854.html.

38. Kroenke, K., R.L. Spitzer & J.B. Williams. 2001. The phq-9. *J. Gen. Intern. Med.* 16: 606–613.

39. Gerber, E.B., M. Whitebook & R.S. Weinstein. 2007. At the heart of child care: predictors of teacher sensitivity in center-based child care. *Early Child. Res. Q.* 22: 327–346.

40. Peisner-Feinberg, E. & M. Burchinal. 1997. Concurrent relations between child care quality and child outcomes: the study of cost, quality, and outcomes in child care centers. *Merrill-Palmer Q.* 43: 451–477.

41. Burchinal, M.R., E. Peisner-Feinberg, D.M. Bryant & R. Clifford. 2000. Children’s social and cognitive development and child-care quality: testing for differential associations related to poverty, gender, or ethnicity. *Appl. Dev. Sci.* 4: 149–165.

42. Auger, A., G. Farkas, M.R. Burchinal, et al. 2014. Preschool center care quality effects on academic achievement: an instrumental variables analysis. *Dev. Psychol.* 50: 2559.

43. Sylva, K., B. Taggart, I. Siraj-Blatchford, et al. 2007. Curricular quality and day-to-day learning activities in pre-school. *Int. J. Early Years Educ.* 15: 49–65.

44. Sammons, P., K. Sylva, J. Hall, et al. 2017. Establishing the effects of quality in early childhood: comparing evidence from England. Early Education: The British Association for Early Childhood Education. http://hdl.handle.net/10871/26994.

45. Araujo, M.C., F. López-Boo, R. Novella, et al. 2015. The quality of Centros Infantiles del Buen Vivir in Ecuador. Inter-American Development Bank, Washington, DC. Accessed February 8, 2018. https://publications.iadb.org/handle/11319/7251.

46. Verdisco, A. & M. Pérez Alfaro. 2010. Measuring education quality in Brazil. *Inter-American Development Bank, Washington, DC*. Accessed February 8, 2018. https://publications.iadb.org/handle/11319/3100.

47. Hallam, R., B. Rous, S. Riley-Ayers & D. Epstein. 2011. Teacher survey of early education quality. National Institute for Early Education Research, New Brunswick, NJ. Accessed February 8, 2018.

48. Sammons, P., K. Sylva, E. Melhuish, et al. 2003. Measuring the impact of pre-school on children’s social/behavioural development over the pre-school period. The Effective Provision of Pre-School Education (EPPE) Project. Technical Paper 8b. University of London, Institute of Education, London, UK.

49. Brunsek, A., M. Perlman, O. Falenchuk, et al. 2017. The relationship between the Early Childhood Environment Rating Scale and its revised form and child outcomes: a systematic review and meta-analysis. *PLoS One* 12: e0178512.

50. Kocalevent, R.D., A. Hinz & E. Brähler. 2013. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *Gen. Hosp. Psychiatry* 35: 551–555.

51. Zaslow, M., R. Anderson, Z. Redd, et al. 2016. I. Quality thresholds, features, and dosage in early care and education: introduction and literature review. *Monogr. Soc. Res. Child Dev.* 81: 7–26.

52. Dickinson, D.K. & M.V. Porche. 2011. Relation between language experiences in preschool classrooms and children’s kindergarten and fourth-grade language and reading abilities. *Child Dev.* 82: 870–886.

53. HighScope. 2012. *HighScope Preschool Curriculum*. Ypsilanti, MI: HighScope Press.

54. Bodrova, E. & D.L. Leong. 2007. *Tools of the Mind*. Columbus, OH: Pearson.

55. Frede, E., K. Jung, W.S. Barnett & A. Figueras-Daniel. 2009. The APPLES Blossom: Abbott Preschool Program Longitudinal Effects Study (APPLES) Preliminary Results through 2nd Grade. New Brunswick, NJ: National Institute for Early Education Research. Accessed February 8, 2018. http://nieer.org/policy-issue/the-apples-blossom-abbott-preschool-program-longitudinal-effects-study-apples-preliminary-results-through-2nd-grade.

56. Barnett, W.S. & E.C. Frede. 2017. Long-term effects of a system of high-quality universal preschool education in the United States. In *Childcare, Early Education and Social Inequality: An International Perspective*. H.P. Blossfeld, N. Kulis, J. Skopec & M. Triventi, Eds.: 152–172. Northampton, MA: Edward Elgar.

57. Barnett, W.S., A.H. Friedman-Krauss, G.G. Weisenfeld, et al. 2017. The State of Preschool 2016: *State Preschool Yearbook*. New Brunswick, NJ: National Institute for Early Education Research. Accessed February 8, 2018. http://nieer.org/state-preschool-yearbooks/yearbook2016.

58. Yoshiokawa, H., D. Leyva, C.E. Snow, et al. 2015. Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes. *Dev. Psychol.* 51: 309–322.

59. Harris, T., R.M. Clifford & D. Cryer. 1998. *Early Childhood Environment Rating Scale Revised Edition*. New York: Teachers’ College, Columbia University.