

A Snapshot of Child Care Center Quality and Child Development in Mexico for Children Under Three: Methodology and Results

Marta Rubio-Codina
Julieth Parra
Daniel Jensen
Ana Mylena Aguilar

Social Protection and Health
Division

TECHNICAL
NOTE N°
IDB-TN-2391

A Snapshot of Child Care Center Quality and Child Development in Mexico for Children Under Three: Methodology and Results

Marta Rubio-Codina
Julieth Parra
Daniel Jensen
Ana Mylena Aguilar

December 2021



Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

A snapshot of child care center quality and child development in Mexico for children under three: methodology and results / Marta Rubio-Codina, Julieth Parra, Daniel, Jensen, Ana Mylena Aguilar.

p. cm. — (IDB Technical Note ; 2391)

Includes bibliographic references.

1. Child development-Government policy-Mexico. 2. Child care services-Government Policy-Mexico. 3. Home visits (Education)-Mexico. I. Codina-Rubio, Marta. II. Parra, Julieth. III. Jensen, Daniel. IV. Aguilar, Ana Mylena. V. Inter-American Development Bank. Social Protection and Health Division. VI. Series.

IDB-TN-2391

<http://www.iadb.org>

Copyright © 2021 Inter-American Development Bank. This work is licensed under a Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) and may be reproduced with attribution to the IDB and for any non-commercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.



scl-sph@iadb.org

www.iadb.org/en/social-protection/social-protection

A snapshot of child care center quality and child development in Mexico for children under three: methodology and results^a

Marta Rubio-Codina,^b Julieth Parra,^b Daniel Jensen,^c Ana Mylena Aguilar^b

Abstract

From 2018 to 2020, we conducted a groundbreaking, nationwide study in Mexico to assess service quality at public child care centers for children under three. The assessment is representative at the national level of the main child care models. We gathered evidence from 245 centers, 426 classrooms, and 2,115 children attending the centers throughout the country, as well as from their households, in order to identify specific interventions to promote child development at the centers. In addition to investigating structural quality, the study focused on process quality (or quality of interactions), measured using the CLASS-Toddler (Classroom Assessment Scoring System). The study found that the structural quality of child care centers in Mexico is high, with safe physical spaces in good condition where children have access to generally age-appropriate play materials and activities. The centers have medium-to-low levels of process quality, especially in the Engaged Support for Learning area, which offers a unique opportunity to continue improving children's experiences at the centers and promote their development. There are also significant challenges related to targeting vulnerable populations.

JEL classification codes: I00, I20, I25, I38, J13.

Keywords: Early childhood, child care center quality, quality of interactions, CLASS, child development, Mexico.

^a This study was led and funded by the Inter-American Development Bank (IDB) as part of two technical cooperations (ME-T1337; ME-T1384), in collaboration with the Executive Secretariat (SE) of the National System for the Integral Protection of Children and Adolescents (*Sistema Nacional de Protección Integral de Niñas, Niños y Adolescentes*—SE-SIPINNA). It was financially supported by the [Early Childhood Development Innovation Fund](#), which is coordinated and managed by the IDB in partnership with Fundación FEMSA, Fundação Maria Cecília Souto Vidigal (FMCSV), Open Society Foundation, and Porticus. It also benefited from the invaluable collaboration of the country's main child care service providers (SNDIF, IMSS, ISSSTE, and SEP). The authors appreciate the openness and willingness of the provider agency authorities to facilitate the study's activities at their child care centers. We are also grateful to Brenda Gonzalez and Sesangari Ortiz from the SE-SIPINNA for their coordination efforts, excellent inputs and constant support. We especially thank all the directors, caregivers, and families participating in the study, as well as all who participated in data collection and data processing activities, including field team trainers. We would also like to thank Collin Stewart and Jana Bundy for translating this document from the original Spanish version. The views in this document reflect those of its authors and not those of the IDB, its board of directors, or the countries they represent, nor do they represent the views of SE-SIPINNA or the provider agencies.

^b Social Protection and Health Division, Inter-American Development Bank.

^c Senior Quantitative Specialist at Compassion International.

1. Introduction

Over the last decade, Mexico has taken steps to promote a comprehensive early childhood protection system that ensures care and development opportunities at child care centers for the youngest children.¹ To this end, the 2011 General Law on Child Care Services and Comprehensive Child Development (*Ley General de Prestación de Servicios para la Atención, Cuidado y Desarrollo Integral Infantil*) established the legal basis for transforming day care and child care centers into environments of equality, quality, warmth, security, and protection in which children are able to develop physically, cognitively, emotionally, and socially. Pursuant to this law, in 2013 the National Council on Child Care Services and Comprehensive Child Development (*Consejo Nacional de Prestación de Servicios para la Atención, Cuidado y Desarrollo Integral Infantil—COPSADII*) was formed, and the National Program for Child Care Services and Comprehensive Child Development (*Programa Nacional de Prestación de Servicios para la Atención, Cuidado, y Desarrollo Integral Infantil*) was established.²

Subsequently, the constitutional educational reform of May 15, 2019 ordered the creation of the National Early Child Care Strategy (*Estrategia Nacional de Atención a la Primera Infancia—ENAPI*), which aims to ensure the rights of children under six, addressing social, geographic, and gender inequalities, as well as the diversity inherent to every individual.³ The ENAPI stems from a high-level agreement and from the efforts of the National System for the Integral Protection of Children and Adolescents (*Sistema Nacional de Protección Integral de Niñas, Niños y Adolescentes—SIPINNA*) to join its national policy with the constitutional mandate in a single public policy instrument. This constitutional reform also established initial education (for children ages 0–3 years) as part of elementary schooling. To this end, the Secretariat of Public Education (*Secretaría de Educación Pública—SEP*) spearheaded a large effort to develop the National Initial Education Policy (*Política Nacional de Educación Inicial—PNEI*), in coordination with other public, civil society, academic, and international organizations.⁴ The PNEI acknowledges all the modalities of this service and sets common guidelines to foster the comprehensive development of children from birth to age three by providing quality services.

According to data from the 2018-2019 National Health and Nutrition Survey (*Encuesta Nacional de Salud y Nutrición—ENSANUT*), 75% of children under three attending a child care center were served by the network of public providers. This is an important achievement in the delivery of social services in Mexico. However, expanding the coverage of equal and high-quality services continues to be a challenge. Only 10% of children under three attend centers, although this percentage rises to 14.7% for children aged two to three. Meanwhile, children from the richest quintile are three times more likely to attend a center than children from households in the poorest quintile (2018–2019 ENSANUT). Moreover, public child care services are fragmented across

¹ “Child care center” is defined as “Any establishment, regardless of name, modality, or type that provides services for comprehensive child care and development, with full respect for children’s rights, from the age of 43 days to six years or when the academic year ends, as may be defined by the care model or by the specific offering of services in question.” (COPSADII Agreement 06/ORD.03/2019).

² Other mechanisms added include the National Register of Child Care Centers (*Registro Nacional de Centros de Atención Infantil—RENCAl*) and the Comprehensive Child Care Center Monitoring and Assessment Program (*Programa Integral de Supervisión, Acompañamiento, Monitoreo y Evaluación del Funcionamiento de los Centros de Atención Infantil—PROISAMEF*).

³ Notice of the ENAPI was published in the Mexican Official Gazette (*Diario Oficial de la Federación—DOF*).

⁴ The PNEI is in its final stages prior to official publication.

various institutional agencies serving different population groups—largely children of formal workers, with different care modalities and different regulatory frameworks. These many factors make coordination in the child care sector very challenging and exacerbate inequalities in access.

In addition to the challenge of broadening service coverage, especially for the most vulnerable populations, there is also the challenge of providing quality services in a context of highly fragmented and heterogeneous public services. The notion of quality of services at child care centers is generally tied to elements of *structural quality*, such as the physical conditions of the centers, equipment, and routines, as well as classroom organization or the child-to-caregiver ratio, for example. These aspects are relatively simple to measure and modify. However, one especially important aspect of service quality is the nature of interactions between caregivers⁵ and children. This is known as *process quality* (or quality of interactions) and is a key element in promoting child development in child care centers, although it is also far more difficult to measure (López Boo et al., 2016). Relationships and bonds formed through warm, receptive, language-rich, and consistent interactions between child and adult caregiver are associated with greater degrees of cognitive, language, and socio-emotional development. They also help cultivate the child's autonomy and self-esteem. The lack of an environment that is stimulating—rich in effective interactions and activities that encourage play and learning—emotionally positive, and stable, may cause developmental problems that are difficult to reverse later.

A first step to understanding how to enhance child development at child care centers is to measure service quality. Thus, in 2019 we visited a nationally representative sample of 246 centers, which is also representative of the main public child care service models, to collect information about service quality to fill in the gaps and offer an assessment focused on measuring process quality.

To this end, we filmed interactions between caregivers and children for four consecutive hours in up to two classrooms at the centers sampled. The recordings were scored according to the CLASS-Toddler (Class Assessment Scoring System, La Paro et al., 2012; Thomason & La Paro, 2009) protocols, the gold standard for observing interactions between caregivers and toddlers. We also gathered information on structural quality through surveys at centers and classrooms, and visited households of up to six children in each sampled classroom to learn about their socioeconomic conditions, the quality of their living environments, and their vocabulary levels. The household survey used the same questions as the 2018-2019 ENSANUT to compare users of public child care services with the Mexican population overall. The study was supplemented by semi-structured interviews of officials from the participating providers and by administrative information from providers regarding their operating costs.

This assessment aims to compile evidence on service quality at child care centers that can be used to make policy recommendations to improve children's experiences at child care centers and promote their overall development. This evidence can inform and further the legislative and institutional efforts made in recent years, translating them into specific policy actions.

⁵ "Caregivers" are "all those working in Early Childhood Education—regardless of institution, modality, role on the organizational chart, level of education, geographical area, or population characteristics—who have a decisive role in creating a service that recognizes human's state of vulnerability at birth and who focus their teaching efforts on developing and transforming children's intellectual, emotional, and physical abilities through a set of learning experiences" (SEP, 2013).

This document comprises five sections, including this introduction as the first. The second section explores what quality of care at centers is and how it is measured. The third section describes the study's methods: target population, sample design, the information collected, and how that information was processed. We then describe the main results and, finally, the last section considers the study's implications for enhancing child care center quality in Mexico and draws conclusions.

2. What is care quality and how is it measured?

Child care service quality is a multifaceted concept that comprises both environmental aspects and children's experiences that promote their well-being and development (Layzer & Goodson, 2006). The literature differentiates between structural and process aspects in measuring child care center quality.

Structural characteristics fall into four main categories: (i) health and safety; (ii) group size and child-to-caregiver ratios—meaning the number of children per caregiver or teacher; (iii) profile of the center's caregivers—for example, their degree of schooling, child development training, experience, or pay; and (iv) infrastructure and equipment—meaning the physical space, availability of suitable educational material, and condition of the center's facilities, etc. (López Boo et al., 2016). Meanwhile, process characteristics concern the nature and frequency of interactions between children and caregivers during the school day as well as the extent to which these interactions aid learning.

Structural quality is important for improving children's experiences because it establishes the conditions of the environment in which they interact with their caregivers. For example, if the child-to-caregiver ratio is high, warm and affectionate interactions may occur less frequently, or if insufficient educational material is available to children, there is less chance for them to engage in activities that improve their classroom experience and promote learning. Traditionally, greater attention has been paid to structural quality because it can be easier to measure using checklists, interviews, or questionnaires. Structural quality is also easier to change. In contrast, process variables are far more difficult and costly to measure because they require hours of observation on location or filming a school day and training people to objectively assess the interactions between adults and children (Layzer & Goodson, 2006). There is a growing consensus regarding the importance of children's interactions with their caregivers as the means by which children's experiences at centers translate to changes in their development. This consensus is based on evidence that shows positive associations between process quality and learning gains (Pianta et al., 2016).

Various instruments have been used to measure process quality in classrooms with children under age three. The most widely used are the Infant and Toddlers Environment Rating Scale, Revised Edition (ITERS-R) for children under 30 months, and the CLASS-Toddler for children aged 15 to 36 months, which is the instrument used in this study. While the ITERS-R focuses on overall aspects of classroom quality—both structural and process-related—the CLASS-Toddler exclusively measures process quality across eight dimensions that fall into two key domains: Emotional and Behavioral Support and Engaged Support for Learning. Box 1 describes this instrument in detail.

Box 1. CLASS-Toddler

CLASS-Toddler (Class Assessment Scoring System) is the international gold standard for measuring the quality of interactions with children ages 15 to 36 months (La Paro, Hamre y Pianta, 2012; Thomason y La Paro, 2009). It measures interaction quality between caregivers and children in the classroom through live or recorded observation. It evaluates two key domains for **effective interactions**: **Emotional and Behavioral Support** and **Engaged Support for Learning**.

Emotional and Behavioral Support

This domain covers caregivers' sensitivity to children's concerns, needs, desires, and emotions, as well as the strategies used to encourage their autonomy and redirect their behavior. **It includes five dimensions:**

- ✓ **Positive Climate** reflects the degree of quality and respect in relationships between caregivers and children, whether there is frequent smiling, and enjoyment of interactions.
- ✓ **Low Negative Climate** evaluates the absence of anger, yelling, threats, or negativity in interactions.
- ✓ **Teacher Sensitivity** shows how caregivers respond to children's needs and emotions and whether they offer children assurance, calm, and stimuli.
- ✓ **Regard for Child Perspectives** measures whether caregivers pay attention to children's needs and interests and promote their responsibility and independence.
- ✓ **Behavior Guidance** reflects whether caregivers help children regulate their behavior, cultivating redirection strategies that minimize negative behavior.

Engaged Support for Learning

This domain addresses whether caregivers use their interactions with children as an opportunity to promote their learning and development, for example, through conversations; whether they encourage their participation in classroom activities; and offer them feedback. **It includes three dimensions:**

- ✓ **Facilitation of Learning and Development** reflects whether play activities foster development, as well as the extent to which caregivers tie and integrate classroom activities into children's day-to-day routines.
- ✓ **Quality of Feedback** shows whether caregivers give clues or explanations or help children express their interpretations; or whether caregivers recognize each child's individual achievements and efforts, letting them know that they did a good job and encouraging their classroom participation and learning.
- ✓ **Language Modeling** reflects whether caregivers use techniques that stimulate children's language use by repeating and responding to their babbling, vocalizations, or verbalizations; expanding on their language; or describing their surroundings and the activities being done, among other techniques.

Each dimension contains various indicators that are scored from **1 to 7**, providing an average score for each dimension and domain, as well as an overall average score. Scores of 1 to 3 indicate low process quality. Scores between 3 and 5 indicate medium levels, and scores over 5 reflect high quality.



Internationally, the CLASS has been used increasingly as a tool for monitoring child care program quality. For example, the United States uses it to monitor quality and as a tool for allocating public child care funding for the Early Head Start (EHS) program. In addition, a U.S. study called “Baby FACES” follows 2 cohorts of children under age three registered with the EHS program. In 2010, when these cohorts of children were 2 years old, the study measured EHS center quality using the CLASS-Toddler. It found medium quality levels, with higher scores in Emotional and Behavioral Support than in Engaged Support for Learning (Vogel et al., 2015).

In Latin America, countries have shown growing interest in measuring process quality at public child care centers using the CLASS-Toddler. Ecuador, Peru, and Argentina are recent examples. Specifically, Araujo et al. (2015) evaluated the quality of Living Well Child Care Centers (*Centros Infantiles del Buen Vivir*) in Ecuador, which are the primary vehicle of public child care services for children under three in the country; Araujo, Dormal and Schady (2017) evaluated the quality of centers within Peru's national *Cuna Más* day care service program, the largest public care program in the country; and López Bóo and Ferro Venegas (2019) assessed the quality of Early Childhood Spaces (*Espacios de Primera Infancia—EPI*) of Greater Buenos Aires in Argentina.⁶ The Ecuador and Peru studies found medium quality levels in Emotional and Behavioral Support and low levels in Engaged Support for Learning. Researchers found slightly higher scores in both domains in Buenos Aires, with Emotional and Behavioral Support scores that were again higher than those in the Engaged Support for Learning domain.

3. Methods

3.1. Target population

This study's target population comprises Mexican child care centers currently run by federal government institutions that serve children under age three; their classrooms; and the children who attend them. The main child care service providers are the National System for Integral Family Development (*Sistema Nacional para el Desarrollo Integral de la Familia—SNDIF*), the Mexican Social Security Institute (*Instituto Mexicano del Seguro Social—IMSS*), the Government Workers' Social Security and Services Institute (*Instituto de Seguridad y Servicios Sociales para los Trabajadores del Estado—ISSSTE*), and the Secretariat of Public Education (*Secretaría de Educación Pública—SEP*).

Each of these agencies has different modalities for child care, targeting its services, management, and caregiver training. The IMSS offers its child care services to the children of formal workers as part of their employment benefits, provided either directly (“Ordinary” scheme and IMSS Mothers [*Madres IMSS*]) or indirectly by outsourcing the service through a provider. In this aspect, it is similar to the ISSSTE, which serves children of government workers at its own or outsourced centers. The SEP's centers also serve its workers' children.⁷ The SNDIF offers child care services through the Community Child Assistance Centers (*Centros de Asistencia Infantil Comunitario—*

⁶ EPIs are care centers reliant upon municipal government, civil society organizations, provinces, and, in some cases, the national government.

⁷ When this study was designed and conducted, the SEP was offering this service to its workers. However, it now offers child care to the entire population, subject to available space. The way SEP centers operate changed again in 2018, while this study was underway, when it launched a new infant and toddler education program known as “A Good Start” (*Un Buen Comienzo*).

CAIC) and Child Development Assistance Centers (*Centros Asistenciales de Desarrollo Infantil—CADI*) modalities. These two modalities are very different from each other in terms of population served, mode of operation, supervision, level of decentralization, sources of financing, and whether they are operated at the state or municipal levels of the System for Integral Family Development (DIF). Table 1 gives an overview of the main characteristics of the child care provider agencies included in this study.

Table 1. Principal institutions providing child care services and their modalities

Institutional agencies	Care modality	Target population	Age range of children
National System for Integral Family Development (SNDIF)	Community Child Assistance Centers (CAICs)	Vulnerable groups in marginalized urban areas	2 years to 5 years, 11 months
	Child Development Assistance Centers (CADIs)	Children of low-income working mothers without social benefits	45 days to 5 years, 11 months
Mexican Social Security Institute (IMSS)	Day cares, direct provision (“Ordinary” scheme and “IMSS Mothers”) and indirect provision	Children of working mothers covered by this social security plan	45 days to 4 years (or 6 years for “IMSS Mothers”)
Government Workers’ Social Security and Services Institute (ISSSTE)	Centers for child development and welfare, whether run by the ISSSTE or outsourced	Children of workers who are ISSSTE beneficiaries	60 days to 5 years, 11 months
Secretariat of Public Education (SEP)	Child Care Centers (CAI-SEP)	Children of Secretariat of Education workers	45 days to 5 years, 11 months ¹

Notes: Prepared based on the Official Gazette of December 28, 2020.

¹Currently serve ages up to three.

3.2. Sample design

During this study’s design phase, we lacked comprehensive information about the centers in operation, such as their location, the number of classrooms by age range, or the total population served. Therefore, between April 2017 and March 2018, we created a census of Mexico’s public centers with administrative information from the public agencies providing child care at that time. This information allowed us to build the study’s sampling frame and identify the most widespread center-based child care modalities in operation at the beginning of 2019: SNDIF CAICs and CADIs, ISSSTE centers, IMSS day cares, SEP centers, and child care centers operated by SEDESOL (now the Secretariat of Welfare or *Secretaría de Bienestar*, formerly the Secretariat of Social Development or *Secretaría de Desarrollo Social*). These modalities are collectively referred

to as “providers” in the rest of this study. Table 2 shows that the statistics compiled are very similar to official totals up to December 2018.⁸

Table 2. The study’s sampling frame

Provider	Administrative data ¹				Study census ²			
	Centers		Children		Centers		Children	
	No.	%	No.	%	No.	%	No.	%
SEDESOL (Secretariat of Welfare)³	9,583	70.3	320,309	42.6	9,126	69.4	305,244	47.2
SNDIF CAICs	1,702	12.5	70,542	9.4	1,699	12.9	69,681	10.8
SNDIFCADIs	476	3.5	34,484	4.6	486	3.7	33,594	5.2
IMSS	1,364	10	197,588	26.3	1,348	10.3	193,711	30
ISSSTE	243	1.8	27,385	3.6	243	1.9	23,936	3.7
Federal SEP	222	1.6	96,741	12.9	204	1.6	15,339	2.4
PEMEX³	20	0.1	3,095	0.4	20	0.2	3,095	0.5
SEMAR³	22	0.2	1,697	0.2	22	0.2	1,697	0.3
Total	13,632	100	751,841	100	13,148	100	646,297	100

Notes: Prepared based on the [Official Gazette](#) of December 28, 2020.

¹ Statistics up to December 2018, also available from the [Official Gazette](#) of December 28, 2020.

² Statistics from the census of centers furnished by providers between March 2017 and April 2018 (study’s sampling frame).

³ These providers were not included in this study because, in the case of SEDESOL (now the Secretariat of Welfare), child care centers were only directly funded until 2018, while PEMEX and SEMAR had fewer than the minimum expected number of centers in the sampling.

In February 2019, various government decisions amended the Operational Rules of the Child Care Facilities Program of the Secretariat of Welfare, and the program as such ended, which affected the sampling design.⁹ We tried to keep these centers in the study, but coordinating with them directly entailed many logistical challenges. We therefore ultimately decided to exclude the centers previously funded by SEDESOL from the study.

We selected a target sample of 60 centers for each of the five providers included in the study, for a total of 300 centers to be visited.¹⁰ We also defined an oversample of 20 centers (backup centers) per provider in case of sample loss.¹¹ We did not include PEMEX (Mexican Petroleum or *Petróleos Mexicanos*) and SEMAR (Secretariat of the Navy or *Secretaría de Marina*) as providers in the study because they had fewer centers than the minimum required for the target sample. Likewise, since some of the IMSS and ISSSTE care modalities have few centers, these centers were sampled as part of the same child care “modality.” However, given their size, the SNDIF’s CAICs and CADIs were treated as two separate “providers” in the sample. For study inclusion, centers had to: (i) receive funds from provider agencies for all or part of their services; and (ii) have at least one classroom serving children under age three. At each center, up to two

⁸ Information on overall enrollment was gathered from each provider. It was not possible to disaggregate the information by classroom, by children under the age of three, or distinguish between infant, toddler and preschool enrollment, for example.

⁹ The direct subsidy that initially went to centers now goes directly to parents.

¹⁰ As described in Appendix I, power calculations indicated that at least 50 centers per provider were required to identify statistical differences between providers.

¹¹ A high non-response rate was expected because this is a highly changeable sector: centers may have closed, changed modality or provider, be located in areas that are inaccessible for safety reasons, or have incorrect or irretrievable contact information, for example.

classrooms were randomly selected with equal probability and, in each classroom, up to six children were randomly selected.

3.3. Field operation

3.3.1. Field team

Seventeen teams were trained, each consisting of four people in different roles.¹² The team *facilitator* and leader was responsible for arranging access to centers, making lists of classrooms and households, presenting the project to directors and parents, recruiting them to participate in the study, and obtaining their informed consent. The *supervisor* was responsible for overseeing and coordinating all activities to collect information at the centers, as well as administering the survey to directors and caregivers and assisting with household surveys. The *videographer* was in charge of making video recordings, and the *survey administrator* was responsible for conducting household surveys, with the supervisor's assistance.

Facilitators, supervisors, and survey administrators received one and a half weeks of initial training, followed by a second, four-day training session.¹³ Those who did not attend both of these in-person training sessions in their entirety received a supplementary online training. Meanwhile, videographers received one week of training on handling cameras, CLASS recording protocols, and data protection and backup. During training, all participants practiced using the questionnaire or performing their specific, assigned tasks. Refresher training specific to each role was also given as needed throughout the field operation.

3.3.2. Phases of the operation

Facilitation. The facilitator made initial contact with the directors of the sample's centers to update contact details and arrange access to the center.¹⁴ Since the census information did not indicate the number of classrooms and age ranges of children served, facilitators used these calls to make a list of classrooms for children under the age of 36 months.

First visit to the center. Using the consolidated list of eligible classrooms, we selected 497 classrooms to create the target sample.¹⁵ In a second phone call with the center director, the facilitator notified the director of the selected classrooms and asked that he or she call an in-person meeting with the parents or guardians of the children in those classrooms, as well as the classrooms' caregivers. This meeting was generally held during the field team's first visit to the center. At this meeting, the field team presented the study to parents or guardians and requested their informed consent for filming and at-home surveys. Parents could agree to participate in one or both of the activities. This meeting was also an opportunity to compile a list of children in each classroom, from which up to 12 children at the center were selected (ideally six per classroom included in the sample).

¹² Of these, six teams were dismissed during operations because they failed to follow study protocols, and another eight left the study. These changes required us to reorganize field operations (hiring and training new teams), which caused substantial delays.

¹³ This second round was necessary due to insufficient opportunities to practice using the instruments during initial training.

¹⁴ Initial contact was made by phone in most cases.

¹⁵ Appendix I provides more details about adjustments made to the sample during field operations.

Second visit to the center and home visit. Between May and December 2019, teams gathered information during simultaneous visits to the center and homes in the sample. During the second visit to centers, the supervisor administered a survey to directors and the caregivers of selected classrooms,¹⁶ and the videographer filmed customary activities, routines, and child-caregiver interactions for four hours, according to CLASS protocols. Generally speaking, this second visit to the center required two consecutive days: one day to film each selected classroom. Meanwhile, the survey administrator, with assistance from the supervisor, visited the homes of children in the sample to administer the household survey to the children's primary parent or guardian.

3.3.3. Data collected

Teams gathered information from the following sources:

- *Center director surveys.* Directors were surveyed to gather information about their professional profiles, terms of employment (type of contract, professional membership, working hours, etc.), knowledge of child development, and training received; center size and class size; personnel employed; availability of play materials in the classrooms and routines; equipment at the center; and condition of physical infrastructure. Survey administrators confirmed physical infrastructure condition and equipment availability and condition by direct observation. This survey lasted an average of 100 minutes (28 to 210 minutes).
- *Caregiver surveys.* Caregivers were surveyed to gather information about their professional profiles, terms of employment, hours worked per week, knowledge of child development, in-class teaching activities, disciplinary practices, etc. This survey lasted an average of 43 minutes (20 to 90 minutes).
- *Video footage of classrooms.* Four consecutive hours of normal classroom activities, routines, and interactions between caregivers and children were filmed at the same time of the day at all centers, according to CLASS protocols.
- *Household surveys.* Survey administrators asked the mother or primary parent/guardian of targeted children about household assets, parental characteristics (age, education, employment status), quality of home environment (play materials and play activities involving an adult caregiver), and perceptions about the center and the quality of care. The questionnaire was based on questions from the Early Childhood Development module of the 2018-2019 ENSANUT (de Castro et al., 2019; Shamah-Levy et al., 2020), which is nationally representative of all children under five, in order to compare the situation of children served by public child care services with Mexico's overall population under age three. Therefore, the questionnaire included questions from the **Family Care Indicators (FCI)** form (Kariger et al., 2012) to measure home environment quality, and the 50-word versions of the **MacArthur-Bates Communicative Development Inventories I, II and III** to measure children's

¹⁶ Only one caregiver per selected classroom was interviewed. When there was more than one caregiver, interviewers asked who spent the most time caring for the children and interviewed that person.

vocabulary level (Jackson-Maldonado et al., 2003, 2012).¹⁷ Inventories I, II and III ask whether children say each of the 50 words in the inventory. These inventories are administered to children aged 12 to 18 months, 19 to 30 months, and 31 to 42 months, respectively.¹⁸ On average, the survey lasted 47 minutes (17 to 87 minutes).

- *Semi-structured interviews.* Interviews with provider agency officials used open-ended, qualitative questions to collect information on the provider's administrative structure; coordination between federal, state, and local authorities; strengths and weaknesses in their care modality and targeting strategy; and how their budget was set and allocated. The project coordinator conducted these interviews by telephone according to a structured guide. On average, these interviews lasted 98 minutes (38 to 125 minutes).¹⁹

3.4. Processing the data

3.4.1 Editing and coding videos

The CLASS protocols were followed for editing and coding videos: starting from the first minute of valid filming, four 20-minute segments were obtained for coding.²⁰ Each segment was independently analyzed and scored by two coding experts, and the two scores were then compared. If the compared scores did not meet the reliability criteria, the segment was coded a third time.²¹ In this process, 5.9% of segments were sent for a third coding. Their final score was an average of the two most similar scores.

The coding team had five members, including a supervisor, all with prior experience using CLASS protocols.²² This team received four days of retraining from a CLASS expert with whom they had worked previously. During retraining, they reviewed definitions and practiced coding. During the seven months of coding, the CLASS expert led weekly group calibration sessions to review problematic segments and answer questions.

Two significant coding challenges occurred when: (i) children under the age of 15 months, who are outside the CLASS-Toddler framework, were present in classrooms; and (ii) footage contained insufficient codable material to create segments (2.1% of classrooms).²³ To address the first challenge, we developed a guide to help coders analyze segments with children under

¹⁷ The 50-word versions were developed by the author of the original inventories (which consist of 100 words) and were piloted in Mexico for use in the 2018 ENSANUT (De Castro et al., 2019).

¹⁸ Inventory I also asks whether the child understands the word. However, only Inventories II and III were analyzed because these were the ones included in the 2018-2019 ENSANUT and, therefore, were the only ones for which it was possible to compare children in the study with children of the same ages in the country.

¹⁹ We also requested information from provider agencies about the service's average monthly per-capita cost, broken down by state and operation model. However, the information gathered was highly disparate between providers and noisy, so it was not used in this study.

²⁰ Each segment may have a maximum of five minutes of non-codable material, when: (i) the camera is covered up because diapers or clothing are being changed, or because someone who has not agreed to participate is involved; (ii) the footage shows interactions with a single child, instead of a group; (iii) the children are sleeping and, therefore, there are no interactions.

²¹ The reliability criteria required differences of one or fewer points in the dimensions of negative climate, quality of feedback, and language modeling or of two or fewer points for the remaining dimensions.

²² The coding team and its supervisor were located in Ecuador. We worked with this team because of its knowledge and prior experience with CLASS protocols. The materials were sent over the cloud, following strict encryption protocols to protect the contents of the video recordings.

²³ This occurred because the school day lasted less than four hours, because the children spent most of the period sleeping, or if someone who had not agreed to participate in the study entered the classroom.

the age of 15 months, based on sample interactions in mixed-age classrooms observed during group coding sessions. For the second challenge, we generated the largest number of segments possible using the available material and used available scores to create the classroom's final score.

3.4.2 Index construction

Using the household survey information, we devised indexes to classify the households served by the centers. The household wealth index covers several characteristics of the house itself and the availability of various assets and serves as a proxy for the household's socioeconomic status. The index was created using data from this study and from the 2018-2019 ENSANUT, which allowed us to compare the wealth distribution of households served by the centers with that of households with children under age three in Mexico. With this information, we also calculated the percentage of households served with wealth levels comparable to the richest quintile for wealth distribution of all Mexican households with children of the same age. We calculated two indexes to compare the distribution of varieties of toys and varieties of play activities in the households served by the centers with that of Mexican households with children of comparable ages. In addition, following instructions from the FCI and De Castro et al. (2019), we used the household survey and the 2018-2019 ENSANUT to calculate the percentage of children who have at least three picture books, at least three varieties of toys, and who do at least four varieties of play activities. Finally, we created two expressive language indexes with the words spoken by the child: one for children aged 19 to 30 months and another for those aged 31 to 36 months. Then we compared the language-level distribution of children served by centers with that of Mexican children in the same age ranges and calculated the percentage of children served by centers with scores higher than the national average for each of these two age groups. Appendix II.1 explores the components of each index in greater detail and the methodology used to guarantee comparability.

Furthermore, we created the following seven indexes to include different aspects of structural quality, based on information gathered from director and caregiver surveys: (i) elements of infrastructure in good condition and without hazards; (ii) characteristics of emergency exits; (iii) emergency equipment; (iv) characteristics of bathrooms; (v) variety of training received by directors and caregivers; (vi) availability of play materials within children's reach; and (vii) variety of activities in the classroom. Each index was calculated as the sum of the characteristics (or elements) present at a center or classrooms, based on the index. Appendix II.2 presents the index's structure and components in greater detail.

3.4.3. Validating children's ages

Twenty-six percent (26%) of children's dates of birth in the household survey did not match the dates that their guardians reported on the informed consent form during the meeting with parents. To understand this discrepancy and identify the correct date, we reviewed the hard copies of all questionnaires and informed consent forms to retrieve both dates of birth and the Unique Population Registry Code (*Clave Única de Registro de Población—CURP*) for the children. The CURP can be used to search for a child's official date of birth on the

National Registry of Population and Personal Identification (*Registro Nacional de Población e Identidad—RENAPO*) website. Using this process, we were able to retrieve the CURP for 78% of those surveyed.²⁴ For these children, the dates of birth recorded in the RENAPO matched those from the household survey and the informed consent form in 82.8% of cases. They only matched the survey in 3% of cases and only matched the informed consent report in 10.1% of cases. They matched neither the dates of birth in the survey nor those on the informed consent form in 4.2% of cases. Therefore, since this is administrative data, we chose to use the RENAPO date of birth whenever there was a discrepancy among sources.

By reviewing dates of birth on the physical documents, we were also able to correct some dates that had been incorrectly transcribed, even for those whose information was not retrieved from the RENAPO. Although this allowed us to correct many of the remaining discrepancies, there were still 130 cases in which the date of birth on the survey did not match the date on the informed consent (5.6% of the survey sample). In these cases, we had a protocol for cleaning up the date of birth: the survey date was used if the difference between it and the consent form date was less than one month. If the date of birth differed by more than one month, we compared the remaining dates with the classroom age group and used the date that matched up to that age range. If there was still a discrepancy, we assessed whether the vocabulary test score was atypical, meaning well above or well below the typical score for the age calculated using the dates of birth from the survey and from the informed consent. If the score was deemed reasonable for the age calculated using the survey's date of birth and was atypical for the informed consent form's age, we used the survey's date of birth. If, on the other hand, the score was reasonable for the age stated on the informed consent form, but atypical for the survey's date of birth, we used the date of birth from the informed consent. There remained 29 cases in which the discrepancy could not be resolved (1.3%). In these cases, the date of birth on the survey was used.

3.4.4. Analysis

We calculated averages for the national total and per provider. However, we anonymized and randomized the order of the results for each provider to maintain impartiality. Each observation was adjusted for its respective sampling weight, corresponding to the number of centers, classrooms or households each observation represents within the study's target population, which is equivalent to the inverse of the probability of being selected for study inclusion. Thus, the values presented in section 4 are a true estimate of the indicator value for the study's target population. Although sample sizes among providers are relatively similar, adjusting each observation in the sample for its respective sampling weight makes it so information from providers with more centers in the country contributes more to the average than that of providers with fewer centers.

We processed the information using the software Stata. Wherever possible, we compared results with findings from similar studies in Latin America and the United States. Specifically, we used the results obtained for Living Well Child Care Centers in Ecuador (Araujo et al., 2015); Peru's

²⁴ The CURP consists of various elements, including the first letter and first vowel of the first surname, the first letter of the second surname, first letter of the first name, date of birth, two letters corresponding to the Mexican state, etc. The main obstacles in reconstructing the CURP from the records were: (i) incorrect registration or correction of the child's date of birth; (ii) incorrect registration of the child's first and last names; (iii) lack of information about the child's state or place of birth; and (iv) illegibility or difficulty in deciphering information reported on the printed documents.

national *Cuna Más* day care service program (Araujo et al. 2017); Early Childhood Spaces of Greater Buenos Aires in Argentina (López Bóo & Ferro Venegas, 2019); and Early Head Start centers in the United States (Vogel et al., 2015).

Finally, we performed a correlation analysis to explore which characteristics of the center's structural quality and of the director and caregiver's profiles were linked to process quality. We also examined whether there is a correlation between children's vocabulary levels and CLASS scores. We calculated partial correlations using the sampling weights and adjusting for "interviewer effects" to correct for any bias survey administrators may have introduced when recording the information, using the "svy_corr" command in Stata.²⁵

3.5. Final sample for analysis

As we processed information, we eliminated observations that, *a posteriori*, did not meet criteria for inclusion in the study. One classroom in which the children were asleep in all video footage was eliminated, as were another three in which all children surveyed turned out to be older than age three once their date of birth was confirmed in the RENAPO. Since two of these classrooms were the only ones for that center, these centers were excluded from our analysis. We also excluded 46 households that were not selected as part of the sample and were surveyed by mistake (2% of the survey sample). Finally, 33 children who were not given the vocabulary test when they were, in reality, older than 8 months were eliminated (1.4% of the survey sample), as were another 117 who were given a vocabulary test section that did not align with their actual ages (5.1%). These discrepancies were identified after checking the children's dates of birth in the RENAPO.

The final sample analyzed included a total of 245 centers and 426 classrooms. Table 3 shows the number of centers and classrooms, by provider. As indicated above, it does not show their names, and their order has been randomized in the results in order to maintain the study's impartiality. In total, we visited 45 to 54 centers per provider and 51 to 104 classrooms per provider. Of these centers, 26.1% have a single classroom in the sample, while the remaining 73.9% have two. For Provider 3, a single classroom was included in the study for 86.7% of centers because the centers' other classrooms were usually at the preschool level. In contrast, 96.2% of Provider 2 centers have two classrooms in the sample, which is why this provider has the highest number of classrooms analyzed.

The classrooms studied were sorted by predominant age group, as follows: (i) infant classrooms, which primarily serve children under the age of 18 months, and (ii) toddler classrooms, serving children aged 18 to 36 months. To sort the classrooms, we used the ages for children recorded in the household survey. Of the classrooms analyzed, 28.2% were infant classrooms and 71.8% were toddler classrooms. Provider 2 has the most infant classrooms in the sample (47 out of 104), while Providers 3 and 4 have the fewest infant classrooms. The lower numbers are because Provider 3 has few classrooms serving this age range while Provider 4 asked that we only include children older than 12 months in the study.

²⁵ This command is the same as the "pcorr" command, but it allows probability weights instead of analytic weights.

We analyzed information for 2,115 children and their households—between 263 and 550 per provider. The children in the study sample had an average age of 24 months; 24.1% of these children were under the age of 18 months (infants), while 75.9% were aged 18 to 36 months (toddlers). The average age of Provider 2 children is the lowest, as could be expected since it has the most infant classrooms in the study. The study analyzes an average of 8.6 children per center and five per classroom.

Table 3. Sample of centers, classrooms, and children analyzed

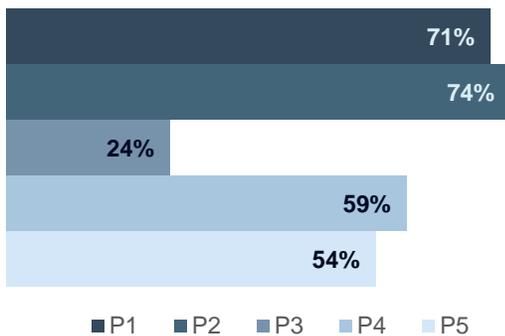
	Provider					Total
	P1	P2	P3	P4	P5	
I. Centers analyzed						
Total centers	47	53	45	54	46	245
II. Classrooms analyzed						
Total classrooms	89	104	51	98	84	426
Total infant classrooms	25	47	5	24	19	120
Total toddler classrooms	64	57	46	74	65	306
III. Classrooms per center						
Centers with 1 classroom, %	10.6	3.8	86.7	18.5	17.4	26.1
Centers with 2 classrooms, %	89.4	96.2	13.3	81.5	82.6	73.9
IV. Children analyzed						
Total children	414	516	263	550	372	2,115
Average age, months	23.1	20.8	28.9	24.9	24.7	24.0
Infant age, %	23.2	41.3	8.4	17.3	22.6	24.1
Toddler age, %	76.8	58.7	91.6	82.7	77.4	75.9
Number of children per center	8.8	9.7	5.8	10.2	8.1	8.6
Number of children per classroom	4.7	5.0	5.2	5.6	4.4	5.0

4. Results

4.1. Where are the child care centers located, and what are the characteristics of the children attending them?

Almost all public child care centers are located in urban areas with low or very low economic marginalization levels. Only 2.2% of the centers are located in rural areas, and 1.5% are in a municipality with high or very high economic marginalization levels, according to the 2015 Municipality Marginalization Index of the National Population Council (*Consejo Nacional de Población—CONAPO*).

Figure 1. Percentage of households in the country's richest quintile of households, by provider



Consistent with the centers' urban location and the fact that three out of the five providers serve children of formal employees, the households served by the centers have a higher socioeconomic status than the average population, with more educated parents. On average, approximately 60% of children who receive public child care services come from households in the country's richest quintile, as identified in the 2018-2019 ENSANUT. More than 70% of households served by the centers of Providers 1 and 2 belong to the country's richest quintile of households (Figure 1). Provider 3 serves households of lower socioeconomic status than those of the other providers, with a lower percentage of households in the richest quintile

(24%). Figure A.III.1 of Appendix III shows that the socioeconomic status of the households served by each provider varies greatly.

In keeping with the socioeconomic status and higher educational level of their parents, the children attending the centers have better play and learning opportunities at home than the average Mexican child under age three. Between 74% and 77% of the children attending centers of Providers 1 and 2 have at least three picture books. This percentage is lower for those attending Provider 3 centers. Even so, 55% of the children attending these centers have at least three picture books, compared to 22% of children under age three nationwide (Figure 2, Chart 1). More than 86% of the children attending the centers of the different providers play with at least three varieties of toys at home.²⁶ For all Mexican households with children of these ages, this percentage is 39% (Figure 2, Chart 2). In more than 86% of households served by Providers 1, 2, and 5, at least four varieties of play activities are carried out with the children. This value is lower for Provider 3 households (74%), although it remains 20 percentage points higher than the national average for households with children in this age group (Figure 2, Chart 3).²⁷ The quality

²⁶ Five varieties of toys were considered: (i) at least three picture books; (ii) toys for assembling; (iii) objects for role play; (iv) toys for learning shapes; and (v) electronic toys.

²⁷ Varieties of play activities refers to whether anyone older than 15 in the household had carried out any of the following activities with the child in the three days leading up to the survey: (i) reading books to the child or looking at pictures in a book with him or her;

of the home environment for each provider varies widely, as shown in Figure A.III.2 of Appendix III.

Figure 2. Quality of play and learning environment in households served by the centers, by provider and national total

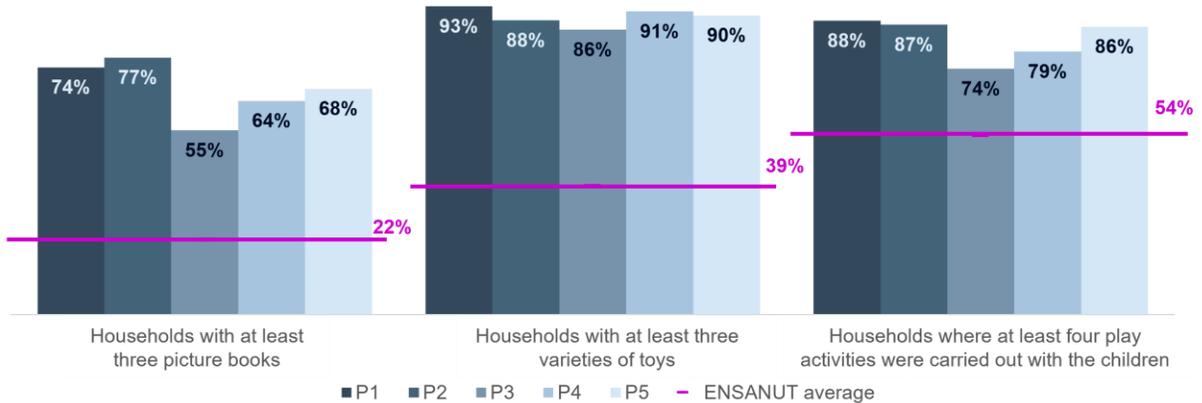
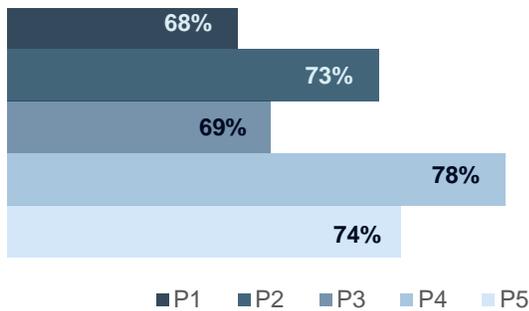


Figure 3. Percentage of children served by the centers with vocabulary scores higher than the national average, by provider



Finally, children aged 19 to 36 months served by the centers have higher levels of language development than children of similar ages nationwide (Figure 3). Of the children served by Provider 4, 78% have vocabulary levels higher than the average child aged 19 to 36 months in Mexico. These percentages range from 73% to 74% for Providers 2 and 5 and from 68% to 69% for Providers 1 and 3. As Figure A.III.3 of Appendix III shows, the vocabulary scores of children aged 19 to 36 months who attend the centers of each provider vary greatly.

(ii) telling the child stories; (iii) singing songs to the child; (iv) taking the child on an outing; (v) playing with the child with the child's toys; (vi) drawing, painting, writing or scribbling; and (vii) naming objects or colors.

4.2. Center size and service characteristics

The centers are open 8-12 hours a day, 10-12 months a year (Table 4). There is little difference in the number of months of child care each provider offers. However, centers' hours vary greatly. For example, the centers of Providers 2 and 3 are open an average of eight hours a day, while Provider 4 centers offer care for 10.2 hours a day, on average (Appendix IV).

The average center has capacity for 138 children and between six and seven classrooms. There are no infant classrooms at 10.3% of the centers, while 43.1% have three or more. Per provider, 80.2% of Provider 3 centers have no infant classrooms, while more than 50% of the centers of Providers 2 and 4 have three or more classrooms of this type. Meanwhile, 14% of the centers have one toddler classroom, and 59% have three or more. There is at least one preschool classroom at 92.8% of centers. These classrooms were not included in the study.

On average, 12.9 children attend each infant classroom and 16.8 attend each toddler classroom. The average attendance is higher at Provider 4 centers, where the average infant classroom has 15.1 children and the average toddler classroom 19.8 (Appendix IV).

Table 4. Center location and size

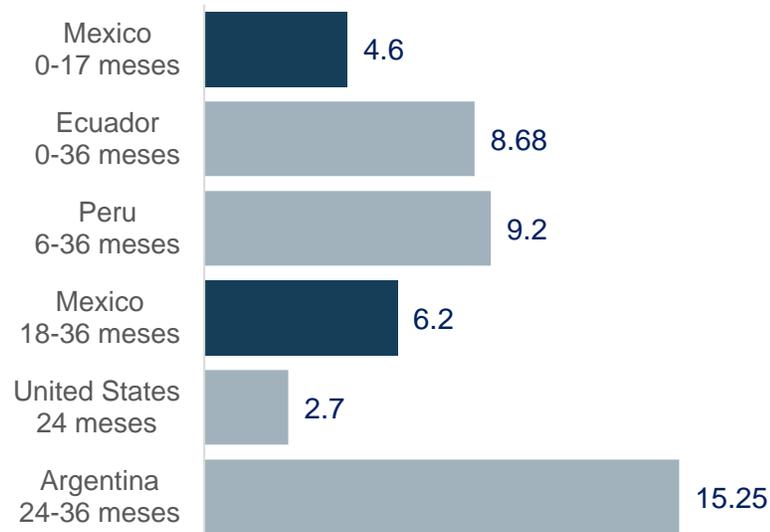
	N	Mean/ Percentage	Standard deviation	P10	P90
Service characteristics					
Months of child care offered per year	242	11.5	0.7	10	12
Hours a day that the center is open	243	9.6	1.4	8	12
Size of the center					
Center's total capacity	245	138.0	60.8	52	214
Number of classrooms	245	6.9	2.1	4	9
Has no infant classrooms, %	245	10.3			
Has 1 infant classroom, %	245	19.6			
Has 2 infant classrooms, %	245	27.0			
Has 3 or more infant classrooms, %	245	43.1			
Has no toddler classrooms, %	245	1.9			
Has 1 toddler classroom, %	245	14.0			
Has 2 toddler classrooms, %	245	25.1			
Has 3 or more toddler classrooms, %	245	59.0			
Has at least 1 pre-school classroom, %	245	92.8			
Average attendance per classroom					
Total	245	16.0	4.9	9.7	22.4
Infants	193	12.9	5.6	6	20
Toddlers	237	16.8	6.4	8	25

4.3. Structural quality

4.3.1. Child-to-caregiver ratios

The number of children per caregiver, or child-to-caregiver ratio, is often used as a proxy for process quality, since having fewer children to take care of facilitates interactions and promotes development. On average, the child care center classrooms have 5.8 children under the age of 36 months per caregiver: 4.6 in infant classrooms, and 6.2 in toddler classrooms. These ratios are lower than those observed in other Latin American countries in which similar studies have been conducted, although still higher than those observed in the United States (Figure 4).

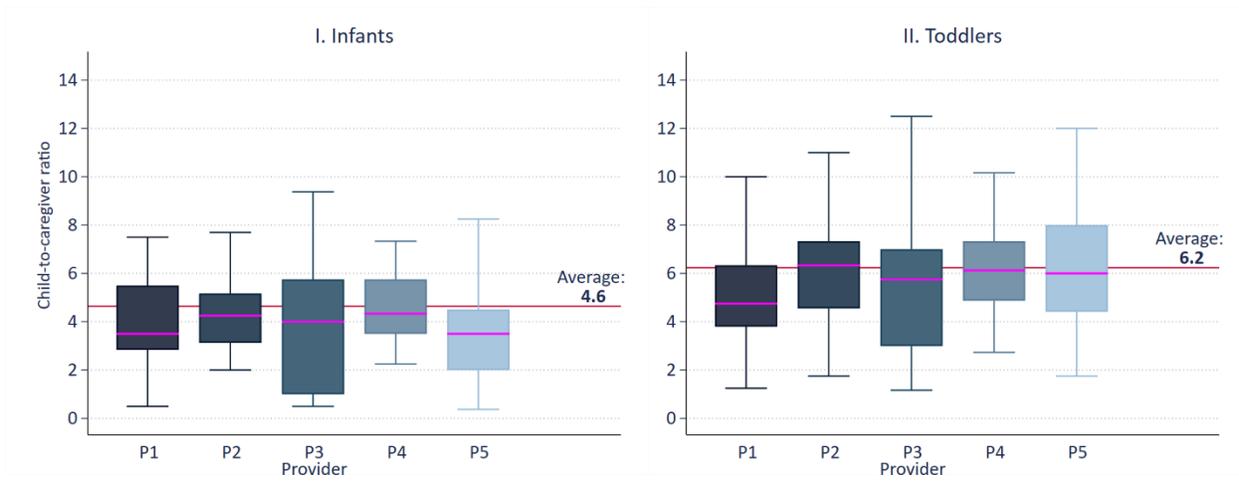
Figure 4. Child-to-caregiver ratios in child care programs in Latin America and the United States



Source: Argentina (López Bóo & Ferro Venegas, 2019); Ecuador (Araujo et al. 2015); Mexico: Director Survey and Caregiver Survey; Peru (Araujo et al., 2017); the United States (Vogel et al., 2015).

Child-to-caregiver ratios vary greatly by provider. Figure 5 shows the distribution of child-to-caregiver ratios per provider in a box-and-whisker plot. The diagram's box indicates the interquartile range for child-to-caregiver ratios, meaning the values observed that are between the 25th and 75th percentiles (the second and third quartiles). The lines extending out from the boxes are the whiskers and indicate the distribution of child-to-caregiver ratios, or of whatever variable is being plotted, in the first and fourth quartiles. In other words, the whiskers show the child-to-caregiver ratios at the bottom 25% of the range (bottom whisker) and top 25% (top whisker). The magenta line indicates the median (50%) child-to-caregiver ratio for each provider. For infant classrooms, this value ranges from 3.5 to 4.3, and for toddler classrooms from 4.8 to 6.3. The child-to-caregiver ratios for Provider 3 range between 0.5 and 9.4 for infant classrooms and between 1.2 and 12.5 for toddler classrooms. Provider 4 centers have less variable child-to-caregiver ratios, while Provider 1 tends to have lower median levels.

Figure 5. Child-to-caregiver ratio in infant and toddler classrooms, by provider



4.3.2. Profile of the director and caregiver

Directors and caregivers have more schooling (14 to 16 years) than the average working-age person in the country (10.1 years).²⁸ Among directors, 83.2% have an undergraduate degree, mostly in preschool education, pedagogy, or psychology. This percentage ranges from 47.8% for Provider 3 to 95.5% for Provider 2. On the other hand, 31.2% of caregivers have an undergraduate degree, primarily in preschool education, child care, or pedagogy. On average, the directors have 14.2 years of experience in their position and have been at the center for 8.7 years, while caregivers have an average of 9 years of experience in their position and have been at the center for 6.4 years (Table 5). Directors and caregivers have received a similar variety of trainings,²⁹ although the average number of training hours reported was 46.4 and 37.8 hours, respectively.³⁰

Table 5. Profile of directors and caregivers

	N	Mean/ Percentage	Standard deviation	P10	P90
I. Director					
Years of education	244	15.9	1.8	15	18
Has undergraduate degree, %	244	83.2			
Years of experience	244	14.2	8.8	4	27
Length of service at the center	245	8.7	8.2	1	20
Hours worked per week	245	46.7	14.2	30	65
Monthly income (including bonuses), Mex\$	238	13,724	7,654	6,800	25,749
Training variety index	245	4.5	1.5	3	6
Hours of training	245	46.4	52.6	8.0	110.0

²⁸ Based on the 2019 National Occupation and Employment Survey.

²⁹ The training variety index has a scale of 0 to 6 and covers 6 topics: (i) concepts, milestones, and fostering child development; (ii) pedagogy, aspects of play, and activities to promote learning; (iii) music and other artistic activities; (iv) how to treat children in a gender- and age-appropriate manner; (v) discipline and behavior management; and (vi) curriculum planning and organization.

³⁰ We found that 91.5% of directors and 87.8% of caregivers had received training on concepts, milestones, and fostering child development, as the most common training topic, followed by training on curriculum planning and organization (84.7% and 85.1%, respectively) and training in pedagogy (84.6% and 81.4%, respectively).

	N	Mean/ Percentage	Standard deviation	P10	P90
II. Caregiver					
Years of education	425	14.0	2.2	11	16
Has undergraduate degree, %	425	31.2			
Years of experience	426	9.0	7.8	1.2	22
Length of service at the center	426	6.4	6.9	1	16
Hours worked per week	426	52.6	16.0	30	74
Monthly income (including bonuses), Mex\$	413	6,001	3,076	3,600	10,000
Training variety index	426	4.5	1.6	2	6
Hours of training	426	37.8	49.4	6.0	94.0

Staff hiring, retention, and training practices vary widely between providers, and even between states and municipalities for more decentralized modalities. This explains the variability in staff profiles (Appendix V). Provider 2's directors and caregivers have more years of education, more years of experience, and a longer length of service at the center, as well as higher pay. Conversely, a lower percentage of Provider 4's caregivers hold an undergraduate degree, and they also have fewer years of experience and a shorter length of service at centers, on average. For this particular provider, more classroom assistants than caregivers were given the survey because assistants spend more time with the children. That said, during the semi-structured interviews, Provider 4's officials emphasized the high turnover rate in their classroom assistants as one of their main challenges, since it means the agency has to frequently retrain new assistants and loses the knowledge and skills generated during initial training.

In Mexico, the educational level of child care center directors is on par with that of other programs' directors and coordinators in the region, although Mexican directors have more experience working at centers (Table 6). The average Mexican caregiver has more years of education and has served longer at centers than caregivers at public child care centers in Argentina, Ecuador, and Peru. The higher educational level and lower turnover of caregivers in Mexico could translate to greater process quality, since there is greater consistency in the people spending the most time with children.

Table 6. Years of education and experience of coordinators, directors, and caregivers at public child care centers in Latin America

Years of:	Coordinator/Director		Caregiver	
	Education	Experience	Education	Experience
Peru	15.8	2.1	10.0	2.1
Ecuador	16.3	3.3	11.1	3.0
Mexico	15.9	8.7	14.0	6.4
Argentina	14.2		13.9	

Source: Director Survey and Caregiver Survey; Peru (Araujo et al., 2017); Ecuador (Araujo et al., 2015); Argentina (López Bóo & Ferro Venegas, 2019)

4.3.3. The centers' physical infrastructure quality and safety

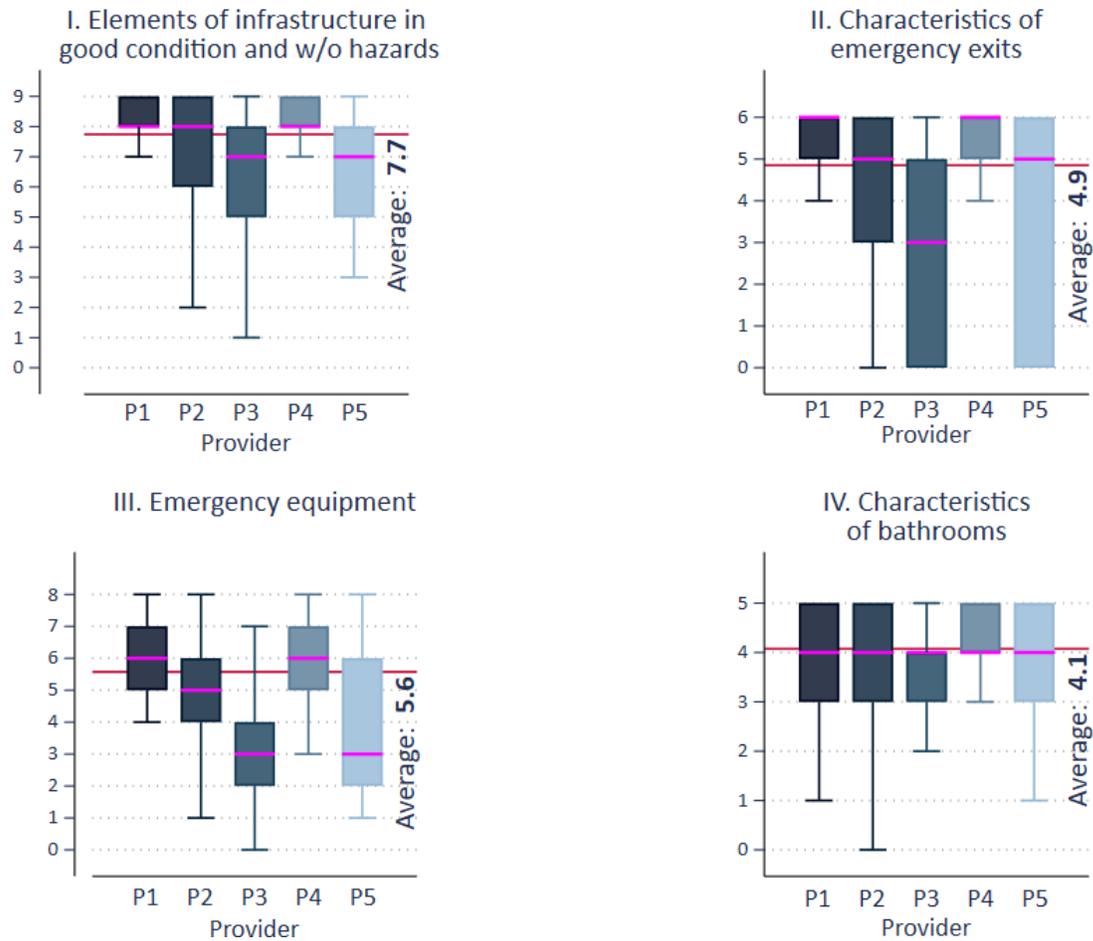
The child care centers have high levels of safety and physical infrastructure quality, although there is room for improvement. Four indexes were constructed to measure the quality and safety of the centers' physical infrastructure using the methodology and components presented in Appendix II.2: (i) elements of infrastructure in good condition and without hazards; (ii) characteristics of emergency exits; (iii) emergency equipment; and (iv) characteristics of bathrooms. The results for each index are described below.

Elements of infrastructure in good condition and without hazards. The centers have an average of 7.7 of the 9 components in the index of elements in good condition and without hazards. This index's value is generally high: 40.3% of the centers have all the characteristics, while 4.7% have four or fewer. More than 80% of the centers have every single characteristic indexed, except for protection where the floor level changes, which was observed in 67.9% of the centers. The index varies somewhat depending on the provider. Providers 1 and 4 have eight or more characteristics at 75% of their centers; Providers 3 and 5 have seven or more characteristics at 75% of their centers, while 25% of the remaining centers had five or fewer (Figure 6, Chart I). Table A.VI.1 of Appendix VI presents the results for each index component and provider.

Characteristics of emergency exits. The centers have an average of 4.9 out of six emergency exit characteristics included in this index, while 50.8% of the centers have emergency exits with all six characteristics and 10.2% have two or fewer. The index varied widely among the centers of Providers 2, 3, and 5, with scores across the entire possible range. Seventy-five percent of the centers of Providers 1 and 4 have five or six emergency exit characteristics, which is above the national average (Figure 6, Chart II). Table A.VI.1 of Appendix VI sets out the results for each component. For example, this tables shows that more than 85% of the centers have emergency exits that are correctly marked with signs or can be opened by an adult, but not by a child.

Emergency equipment. The centers have an average of 5.6 types of emergency equipment of the eight listed. Of the centers studied, 19.8% had eight types of equipment, 58.3% had six or more, and 5.1% had one or none. The index score ranged from four to eight for Provider 1 and from three to eight for Provider 4. These two providers had the highest average scores (Figure 6, Chart III). Twenty-five percent of Provider 2 centers had between one and four types of emergency equipment, while the other centers had more than four. Half of Provider 3 and 5 centers had three or fewer of the different types of equipment. Table A.VI.1 of Appendix VI breaks down the index components. Notably, 87.5% of centers had smoke detectors, 79.8% had flashlights, 77.6% had doorbells, and 73.5% had sirens.

Figure 6. Quality and safety of centers' physical infrastructure, by provider



Characteristics of bathrooms. The centers had an average of 4.1 of the 5 features of bathroom facilities listed. Of the centers studied, 80.1% had four or more of these features while 2.6% had one or none. All providers had the same median index score, although a higher percentage of Provider 1, 2 and 5 centers have two or fewer attributes (Figure 6, Panel IV). Table A.VI.1 of Appendix VI lists the index's components, but more than 87% of bathrooms had soap, toilet paper, or items for drying hands.

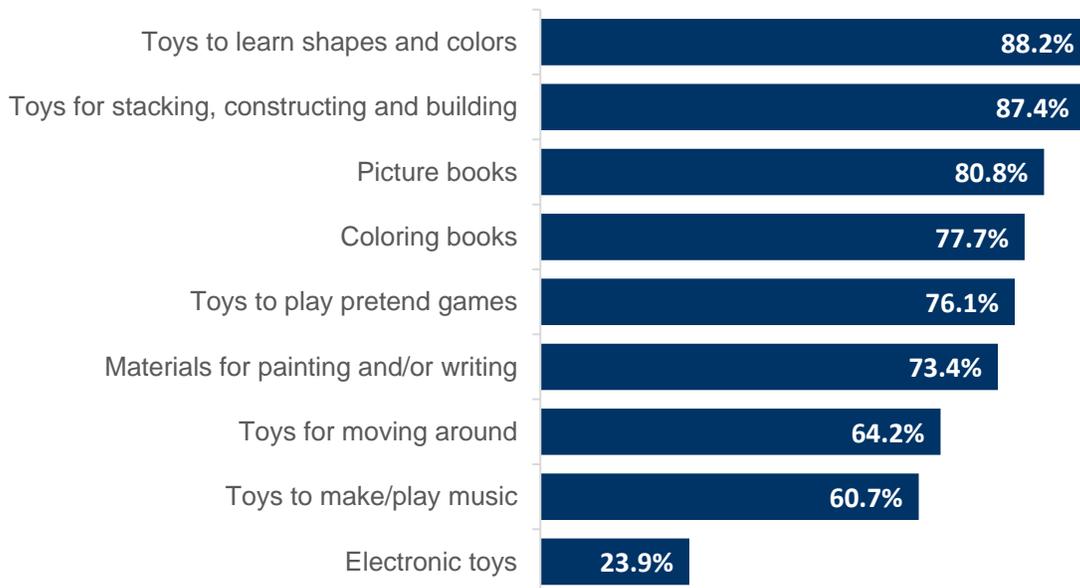
Other elements of physical infrastructure quality. We considered other infrastructure elements, but these were not included in the indexes because they varied little among child care centers. A notably high percentage of centers had safety features like all power outlets located inside walls and protected with covers, safe stairways, wall-mounted switches, protected wiring and stairs, safe outdoor playground equipment, or restricted access to hazardous areas (see Table A.VI.1, Appendix VI).

4.3.4. Variety of play materials within children's reach

Having play materials within children's reach fosters free play, exploration, and curiosity and boosts children's autonomy. As shown in Figure 7, the centers have a wide variety of toys within

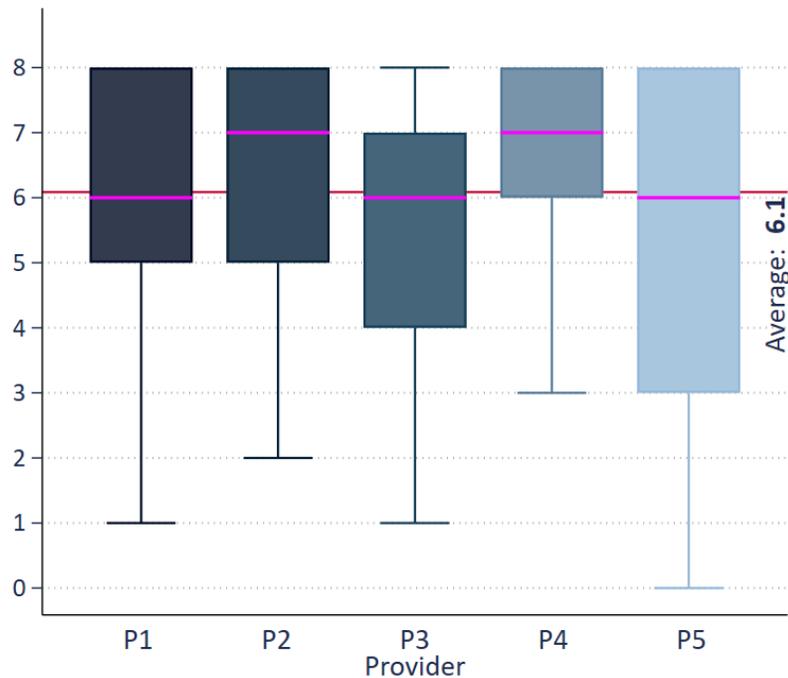
children’s reach. The most prevalent were toys for learning shapes and colors (88.2% of centers) and toys for assembling or building (87.4%). Although a high percentage of centers do have picture books, 19.2% do not have picture books within children’s reach (between 13.5% and 26%, depending on the provider), and 4.1% of centers have no picture books at all (between 1.1% and 8.4%, depending on the provider). Picture books are key to developing attention focus and concentration, building vocabulary, and forming bonds, among other developmental aspects, for children of all ages. Toys used for role play were observed within children’s reach at 76.1% of the centers. These objects are also very important for developing imagination and creativity in children. Other play materials within children’s reach at the centers were: materials for painting and/or writing (73.4% of centers), toys requiring physical movement (64.2%), and toys that make music (60.7%). Finally, 23.9% of centers have electronic toys within children’s reach.

Figure 7. Variety of play materials within children’s reach



On average, child care centers have 6.1 of the 8 varieties of play materials, excluding electronic toys (Figure 8). The types of play materials within children’s reach varied widely between providers, although more than half of each provider’s child care centers have six or more varieties. Of Provider 5 classrooms, 6.6% have no variety of play materials within children’s reach and 25% have three or fewer varieties (Appendix VII).

Figure 8. Variety of play materials within children’s reach, by provider



4.3.5. Variety of play activities

The children engage in a wide variety of play and learning activities with their caregivers during the week. The most frequent activities are singing, listening to music, or dancing; doing physical activities that involve gross motor skills; learning the parts of the body; and building with blocks and cubes. Activities less suitable for smaller children, like cutting and pasting pre-drawn shapes, learning the letters of the alphabet, or counting and learning numbers are done less frequently overall, but they are more common in toddler classrooms. For example, 84% of the toddler classrooms count or learn numbers once a week, compared to 57% of the infant classrooms. Likewise, 54% of the toddler classrooms cut and paste shapes, compared to 33% of the infant classrooms (Figure 9).

On average, 13.3 of the 16 educational activities take place at least once a week in the classrooms: 12.1 activities in infant classrooms and 13.8 in toddler classrooms. The number of activities carried out in each provider’s classrooms varies somewhat (Figure 10). Half of Provider 2 classrooms do 12 or fewer activities, which has to do with the fact that nearly half of these classrooms are for infants. Appendix VIII shows that activities like counting or learning numbers, learning the letters of the alphabet, and cutting and pasting pre-drawn shapes occur less frequently in Provider 2 classrooms. These activities are more common in Provider 3 classrooms, of which only 9% are for infants.

Figure 9. Play activities carried out at least once a week in infant and toddler classrooms

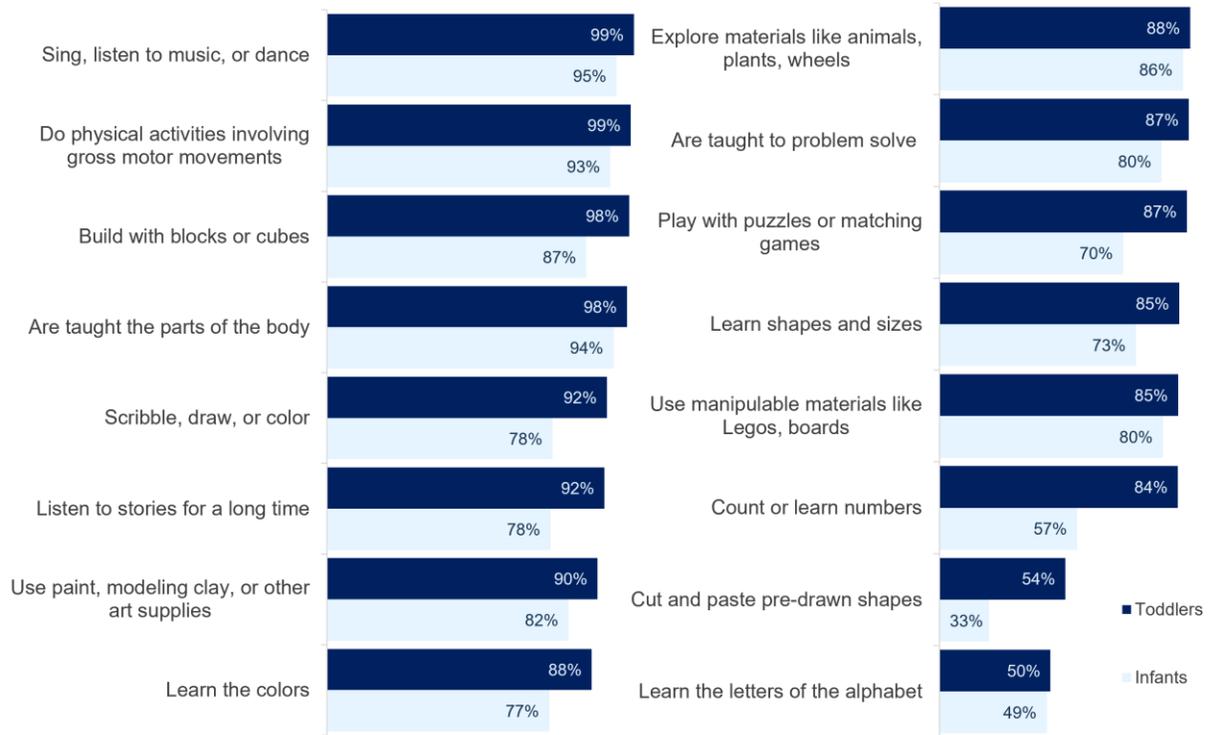
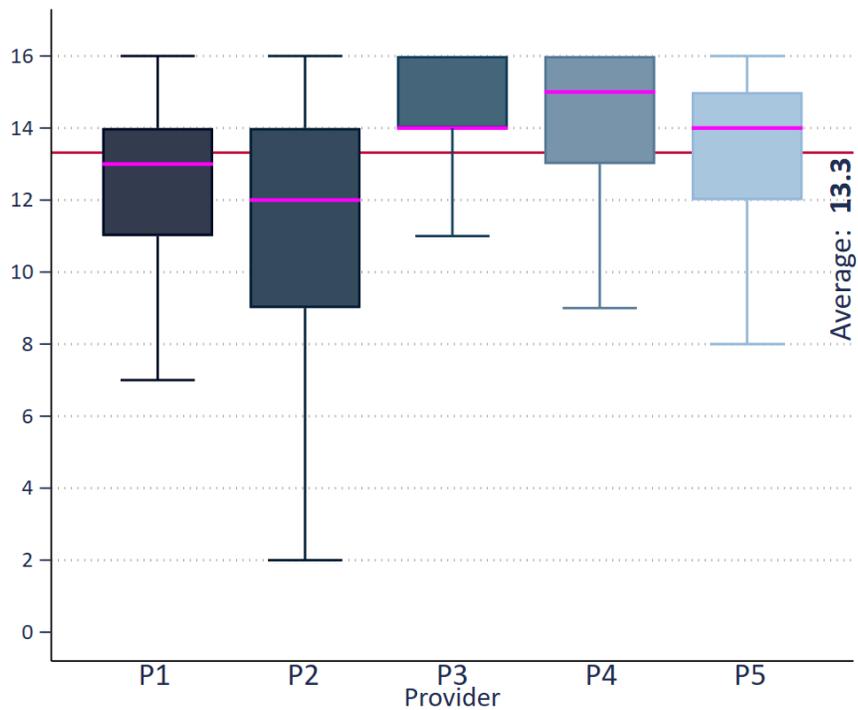


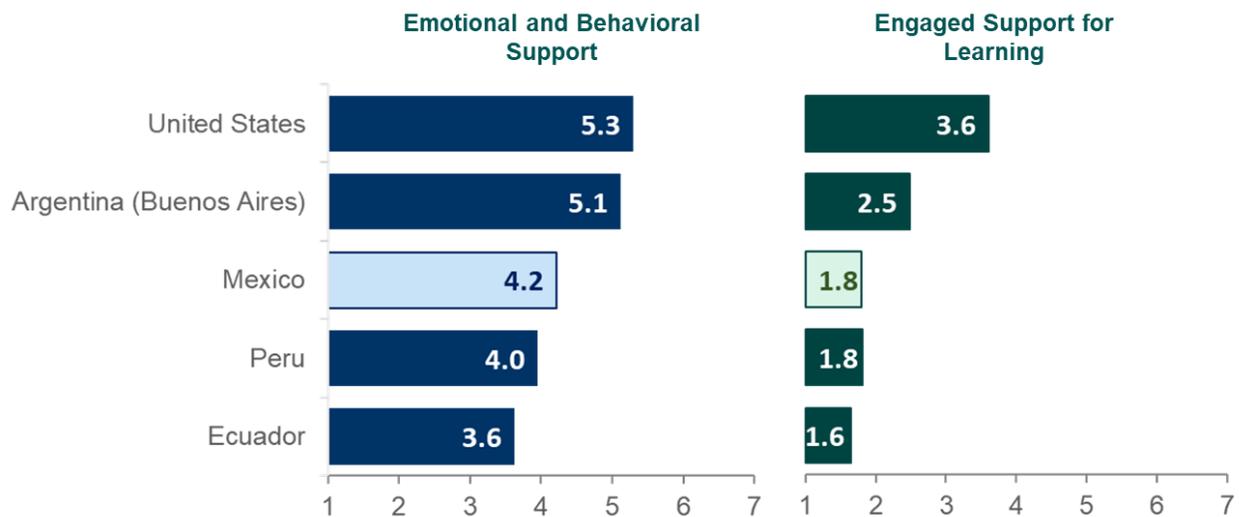
Figure 10. Variety of play activities in classrooms, by provider



4.4. Process quality

The centers have medium to low process quality levels. The average CLASS score for all providers is 3.3, with a score of 4.2 in Emotional and Behavioral Support and 1.8 in Engaged Support for Learning.³¹ These scores are comparable to those observed at other public child care services in Latin America and lower than those seen in the United States, a country with a stronger tradition of using process quality monitoring tools as a strategy for improving children’s classroom experience. In all countries with similar measurements, lower scores were found for Engaged Support for Learning than for Emotional and Behavioral Support (Figure 11).

Figure 11. Process quality in Mexico, the United States, and other Latin American countries



Source: Argentina (López Bóo & Ferro Venegas, 2019); Ecuador (Araujo et al., 2015); Mexico; Coded Videos: Peru (*Videos codificados*; Peru) (Araujo et al., 2017); United States (Vogel et al., 2015).

Scores are similarly distributed among providers, but with a certain amount of internal variability. The median Emotional and Behavioral Support score was between 4.1 and 4.5 (Figure 12, Chart I), while the median Engaged Support for Learning score was between 1.7 and 1.9 (Figure 12, Chart II). Provider 5 centers showed greater variability, with scores ranging from 3.1 to 5.1 in Emotional and Behavioral Support and 1.3 to 2.7 for Engaged Support for Learning.

³¹ See Box 1 for definitions of the two CLASS domains and their constituent dimensions.

Figure 12. Process quality by domain and dimension, by provider

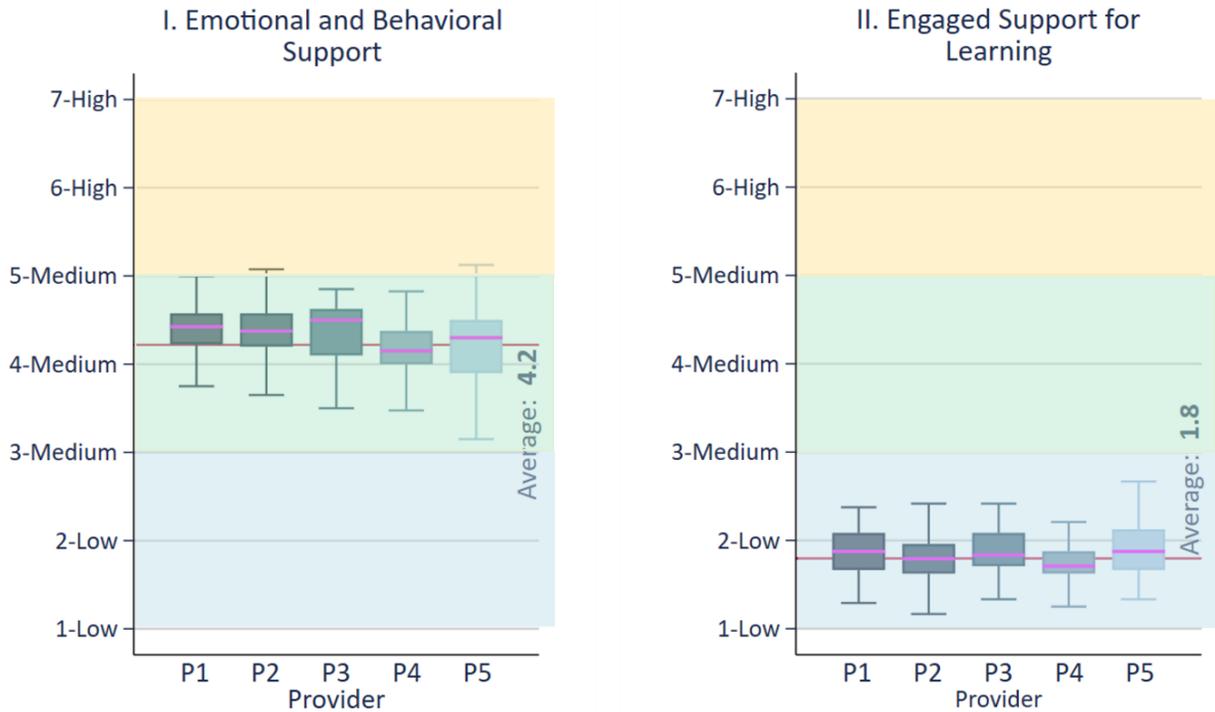


Table 7. CLASS scores for each domain and dimension

	N	Mean/ Percentage	Standard deviation	P10	P90
Total	426	3.3	0.3	3.0	3.6
Emotional and Behavioral Support	426	4.2	0.3	3.8	4.6
Positive climate	426	4.1	0.5	3.5	4.8
Low negative climate	426	6.9	0.2	6.6	7
Caregiver sensitivity	426	4.0	0.5	3.3	4.6
Regard for child perspectives	426	2.4	0.3	2	2.9
Behavior guidance	426	3.7	0.4	3.3	4.3
Engaged Support for Learning	426	1.8	0.2	1.5	2.1
Facilitation of learning	426	2.5	0.3	2	2.9
Quality of feedback	426	1.3	0.2	1	1.6
Language modeling	426	1.6	0.3	1.3	2

By dimension, we found the best performance in Low negative climate, with an average score of 6.9, which is in the high quality range. The other dimensions fell in the middle or low quality ranges (Table 7), with very similar average scores among providers (Figures A.IX.1 and A.IX.2, Appendix IX).

We describe the implications of each dimension's scores below:

Low negative climate. We observed no displays of hostility, negativity, or irritability from caregivers or moments of anger or intense frustration by children. The average score for Low negative climate is high and very similar for all providers (Figure A.IX.1, Appendix IX).

Positive climate. We observed mid-range quality levels, with an average score of 4.1 (Figure A.IX.1, Appendix IX). This score means that although there are demonstrations of warmth and affection, they are not as frequent as desired or do not involve the majority of the children. It is important for expressions like laughter, smiles, and hugs to be more frequent and include all children so they are more enthusiastic about being together and interacting throughout the day. Caregivers could more consistently use respectful language and a calm, friendly voice, using expressions like "please" and "thank you" more often and constantly encouraging children to do the same. Likewise, it is important for the caregiver to be closer to children, maintaining eye contact while talking to them and turning her body toward them to show she is paying attention.

Caregiver sensitivity. This dimension fell within the medium range of quality, with an average score of 4.0 and a range of 3.9 to 4.2, depending on the provider (Figure A.IX.1, Appendix IX). We observed that caregivers are not always aware of children's needs, concerns, desires, and emotions. Caregivers receive high scores when they are consistently in tune with children's interests and difficulties or with verbal and non-verbal cues that indicate a need for support and attention. It is also important for caregivers to offer children a sense of calm and safety when offering support, although this may be difficult with large class sizes, when all of the children require individual support simultaneously.

Behavior guidance. We observed medium levels in this dimension, with an average score of 3.7 and little variability among providers (Figure A.IX.1, Appendix IX). This indicates that caregivers do not always consistently monitor behavior and communicate behavior expectations. We suggest strengthening the use of strategies like recognizing and reinforcing desired positive behaviors, with simple reminders accompanied by explanations when children's behavior is problematic or inappropriate so they can understand why this is the case.

Regard for child perspectives. We found low quality in this dimension, with an average score of 2.4 and little variation among providers (Figure A.IX.1, Appendix IX). We found that caregivers tended to direct and control children's activities and movements during the day. It is important to pay greater attention to children's perspectives and views, and to be flexible in planning and implementing educational activities, so children are able to pursue and complete activities at their own pace. Ideally, caregivers should also give children different options to choose from, like activities, songs, or colors for painting, to foster their independence and autonomy.

Facilitation of learning. We found low quality in this dimension, with an average score of 2.5 and little variation among providers (Figure A.IX.2, Appendix IX). We observed that caregivers did not offer sufficient learning opportunities to the children, whose participation in daily activities was rather passive. Children could be given learning opportunities using strategies like relating educational activities to their day-to-day lives, experience, and what they learned previously;

drawing connections to what happens in their environment; or introducing notions of cause and effect or sequentiality. It is also key to encourage children to participate in activities, offering them new challenges that stimulate and motivate them to continue playing and learning.

Quality of feedback. Child care centers scored an average of 1.3 in this dimension. All providers had a similar, low-quality score (Figure A.IX.2, Appendix IX). In general, caregivers tend to give children brief, superficial, and non-specific feedback. Effective feedback must be specific, letting children know what they did well and why. It should also be based on each child's individual achievements and efforts. In this regard, children who are frustrated by an activity should also be offered support and encouraged to persist in their tasks. This is obviously hard to do in very large groups.

Language modeling. The average score for this dimension was 1.6 and was very similar among providers (Figure A.IX.2, Appendix IX). We found that caregivers engaged in very few conversations with the children, and their communication tended to be limited to explaining activities. It is important for the caregiver to engage in two-way conversations with the children, repeating, responding, and expanding on their attempts to communicate. For example, when a child says "woof" or "dog," a caregiver might say: "Yes, it's a big dog and it has four paws!" The caregiver may also introduce new, progressively more complex words and phrases to help build children's vocabularies.

4.5. Correlations with process quality

International literature shows two tendencies: a negative correlation between process quality and group size and child-to-caregiver ratios; and a positive correlation between process quality and caregivers' qualifications, experience at the centers, and training received during their time of service (Slot, 2018). At public child care centers in Mexico, the study found process quality to be negatively correlated with center size (measured by total capacity), child-to-caregiver ratios in toddler classrooms, and infrastructure quality. There is also a positive correlation between the CLASS-Toddler score and the variety of classroom activities that foster development. However, these correlations are not consistent for both CLASS-Toddler domains (Table 8). The CLASS scores are also consistent and correlate positively with the director's income and years of experience and correlate negatively with the number of hours worked per week. There is also a positive association between scores and the caregiver's years of experience, length of service at the center, monthly pay, and knowledge of child development, although this last characteristic only correlates with Emotional and Behavioral Support.

We did not find statistically significant associations between children's levels of language development and process quality in the classroom. It is important to note that the correlation between child development and CLASS-Toddler scores is not consistent in the literature. For example, in Argentina, researchers found a positive correlation ($r = 0.28$) between Emotional and Behavioral Support and receptive language measured using the Bayley-III, but they found a negative correlation ($r = -0.29$) with Engaged Support for Learning (López Bóo & Ferro Venegas, 2019). In Peru, a positive association was found between both CLASS-Toddler scores and children's development level, measured using the ASQ-3 (Araujo et al., 2017).

Generally speaking, the vocabulary level of the children attending the centers correlates positively, as one would expect, with household wealth and variety of play materials and play activities (Appendix X), but not with the structural characteristics of the centers or process quality in the classroom.

Table 8. Process quality correlations

	Emotional and Behavioral Support	Engaged Support for Learning
I. Characteristics of center		
Total capacity of center	-0.140*	-0.216**
Children per caregiver in infant classrooms	-0.156	-0.066
Children per caregiver in toddler classrooms	-0.201**	-0.138
Elements of infrastructure in good condition and without hazards	-0.045	-0.151*
Variety of play materials within children's reach	-0.077	-0.102
Variety of play activities in classrooms	-0.027	0.204**
II. Characteristics of the director		
Years of education	-0.067	-0.091
Years of experience	0.149*	0.123*
Length of service at the center (years)	0.089	0.148*
Monthly pay	0.200**	0.142*
Total hours worked	-0.181*	-0.167*
Knowledge about child development	0.058	-0.053
III. Characteristics of the caregiver		
Years of education	0.075	0.051
Years of experience	0.206***	0.246***
Length of service at the center (years)	0.148**	0.135*
Monthly pay	0.231***	0.152*
Total hours worked	-0.083	-0.097
Knowledge about child development	0.149*	0.085
IV. Language development in children aged 19 to 36 months		
Vocabulary level	-0.051	-0.060

Note: The unit of analysis is the 426 classrooms, with the exception of Panel IV, where correlations are drawn at the child level. Vocabulary tests were standardized by subtracting the mean, then dividing the result by the standard deviation in each month of age. This eliminates the effect of age from the two inventories and allows the results to be combined to produce a single correlation with vocabulary using a larger sample size.

Statistically significant at 5% (*), 1% (**), and 0.1% (***).

5. Conclusions and recommendations

This study provides a preliminary assessment of the quality of child care at public child care centers in Mexico, breaking new ground in several ways. First, it contains information from a sampling of centers that is nationally representative and also representative of the five main providers: SNDIF's CADIs and CAICs, IMSS day cares, ISSSTE centers, and SEP child care centers. Second is the richness of information gathered by surveying the centers' directors and caregivers, who provided information on: the condition of centers' physical infrastructure; their equipment; the availability of play materials; staff profiles; and routines and play activities carried out with the children, in one or two classrooms. We also filmed four hours of daily activities in one or two classrooms at each center and visited up to 12 households per center to learn about their socioeconomic status, the quality of the learning and play environment in the home, and the children's vocabulary levels. The household survey used the same questions as the 2018-2019 ENSANUT, which allowed us to compare the distribution of various indicators for children who attend the centers and their households with that of all Mexican children under age three. We also conducted phone interviews with provider agency officials to better understand how centers are managed and organized. Finally, the study's most innovative aspect was its focus on measuring process quality, to which end we coded video segments according to CLASS-Toddler protocols.

One of the study's findings is that public child care centers serve children in households with better socioeconomic conditions, with parents with more years of formal education, and with a larger variety of picture books, toys, and play activities. These facts are not surprising given that three of the five child care modalities focus on the children of formal employees. Although one of the modalities in the study serves more vulnerable groups, most of the children served by this provider live in better socioeconomic conditions than the average Mexican household with children under age three. In keeping with this reality, most children served by the centers have a better vocabulary level than Mexican children in general in the same age range. This finding confirms existing challenges related to targeting and care service coverage for the center-based modality, in addition to limited coverage in marginalized and rural areas.

The study finds that child care centers in Mexico have high levels of structural quality, with safe physical spaces that are in good condition where children have access to generally age-appropriate play materials and activities. These elements are indispensable to high-quality care. However, providers have a unique opportunity to continue improving children's experience at the centers by strengthening both emotional support and engaged support for learning in the children's interactions with their caregivers.

International evidence indicates how complex it is to achieve high levels of process quality and suggests that it depends less on rules and regulations and more on what actually occurs in classrooms. This challenge is more difficult the younger the children at centers are (Helburn et al., 1995; Howes & Smith, 1995; National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network, 2000). However, various strategies may enhance children's experiences at centers. These strategies include those that focus on positive reinforcement; recognizing achievements and efforts; redirecting problematic behaviors with positive disciplinary strategies; including play and educational activities that match children's

interests and developmental needs; recognizing and expanding on children's attempts to communicate using language modeling; or reducing the child-to-caregiver ratio in each classroom and reorganizing spaces and routines to facilitate dedicated care.

To promote the adoption of these strategies, we propose developing a continuous process quality assurance system that includes training and professional development programs for caregivers that combine theoretical content and practical and experiential learning—for example, individual or group mentoring or strategies for ongoing support. These programs will equip caregivers to develop some of the skills and techniques needed to organize classroom routines and activities more effectively, maximizing the potential of classroom dynamics and interactions, and offering children more and more effective learning opportunities. It is also important to develop mechanisms to reduce caregiver turnover and maximize the returns from any training initiative.

Lastly, it is essential to develop and provide simple tools for continually measuring and monitoring process quality. These tools, together with the providers' existing structural quality supervision instruments, will be key to identifying concrete aspects to address in mentoring experiences and for following up on the results of a continuous quality assurance system. Several provider agencies that participated in the study are already undertaking similar efforts. To ensure they are scalable and sustainable, strategies should be designed to easily integrate into and enhance existing service structures, taking into account the needs of the centers' staff.

References

- Araujo, M. C., Dormal, M., & Schady, N. (2017). Child Care and Child Development. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, February, 391–394. <https://doi.org/10.1016/B978-0-08-097086-8.23208-X>
- Araujo, M. C., Lopez-Boo, F., Novella, R., Schodt, S., & Tomé, R. (2015). La calidad de los Centros Infantiles del Buen Vivir en Ecuador. *Resumen de Políticas IDB-PB*, 248. <https://publications.iadb.org/es/publicacion/17031/la-calidad-de-los-centros-infantiles-del-buen-vivir-en-ecuador>
- de Castro, F., Vázquez-Salas, R. A., Villalobos, A., Rubio-Codina, M., Prado, E., Sánchez-Ferrer, J. C., Romero, M., & Shamah-Levy, T. (2019). Contexto y resultados del desarrollo infantil temprano en niños y niñas de 12 a 59 meses en México. *Salud Pública de México*, 61(6, nov-dic), 775. <https://doi.org/10.21149/10560>
- Helburn, S. W., Culkin, M. L., Morris, J. R., Mocan, H. N., Howes, C., Phillipsen, L. C., Bryant, D. M., Clifford, R. M., Cryer, D., Peisner-Feinberg, E. S., Burchinal, M., Kagan, S. L., & Rustici, J. (1995). *Cost, Quality and Child Outcomes in Child Care Centers. Technical Report*. <https://files.eric.ed.gov/fulltext/ED386297.pdf>
- Howes, C., & Smith, E. W. (1995). Relations among child care quality, teacher behavior, children's play activities, emotional security, and cognitive activity in child care. *Early Childhood Research Quarterly*, 10(4), 381–404. [https://doi.org/10.1016/0885-2006\(95\)90013-6](https://doi.org/10.1016/0885-2006(95)90013-6)
- Jackson-Maldonado, D., Marchman, V. A., & Fernald, L. C. H. (2012). Short-form versions of the Spanish MacArthur–Bates Communicative Development Inventories. *Applied Psycholinguistics*, 34(14), 837–868.
- Jackson-Maldonado, D., Thal, D., Marchman, V., Newton, T., Fenson, L., & Conboy, B. (2003). *Mac Arthur Inventarios del Desarrollo de Habilidades Comunicativas. User's guide and technical manual*. Paul H. Brookes Publishing Co.
- Kariger, P., Frongillo, E. A., Engle, P., Britto, P. M. R., Sywulka, S. M., & Menon, P. (2012). Indicators of Family Care for Development for Use in Multicountry Surveys. *Journal of Health, Population and Nutrition*, 30(4), 472–486. <https://doi.org/10.3329/jhpn.v30i4.13417>
- Laparo, K. M., Hamre, B. K., & Pianta, R. C. (2012). *Classroom Assessment Scoring System™ (CLASS™) Manual, Toddler*.
- Layzer, J. I., & Goodson, B. D. (2006). The “quality” of early care and education settings: Definitional and measurement issues. *Evaluation Review*, 30(5), 556–576.
- López Boo, F., Araujo, M. C., & Tomé, R. (2016). How is Child Care Quality Measured?: A toolkit. <https://doi.org/10.18235/0000242>
- López Bóo, F., & Ferro Venegas, M. de la P. (2019). Calidad de procesos y desarrollo infantil en los Espacios de Primera Infancia del Gran Buenos Aires: Validación de una lista corta de monitoreo de centros infantiles. *Monografía del BID 733*. <https://doi.org/10.18235/0001898>

- National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network. (2000). Characteristics and Quality of Child Care for Toddlers and Preschoolers. *Applied Developmental Science*, 4(3), 116–135. https://doi.org/10.1207/S1532480XADS0403_2
- Pianta, R., Downer, J., & Hamre, B. (2016). Quality in early education classrooms: Definitions, gaps, and systems. *The Future of Children*, 26(2), 119–138.
- Shamah-Levy T, Vielma-Orozco E, Heredia-Hernández O, Romero-Martínez M, Mojica-Cuevas J, Cuevas-Nasu L, Santaella-Castell JA, Rivera-Dommarco J. Encuesta Nacional de Salud y Nutrición 2018-19: Resultados Nacionales. Cuernavaca, México: Instituto Nacional de Salud Pública, 2020. <https://www.insp.mx/produccion-editorial/novedades-editoriales/ensanut-2018-nacionales>
- Slot, P. (2018). Structural characteristics and process quality in early childhood education and care: A literature review. *OECD Education Working Paper*, 176. <https://doi.org/10.1787/edaf3793-en>
- Thomason, A. C., & la Paro, K. M. (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>
- Vogel, C. A., Caronongan, P., Xue, Y., Thomas, J., Bandel, E., Aikens, N., Boller, K., & Murphy, L. (2015). *Toddlers in Early Head Start: A Portrait of 3-Year-Olds, Their Families, and the Programs Serving Them: Volume II: Technical Appendices*. https://www.acf.hhs.gov/sites/default/files/documents/opre/bfaces_age_3_vol_ii_techapp_3_18_15_b508.pdf

Appendixes

Appendix I. From sampling frame to sample for analysis

We designed the study to measure the quality of the centers operated by six providers: CAICs, CADIs, IMSS, ISSSTE, SEP and SEDESOL, which were the providers with the broadest center-based child care coverage in early 2019. At that time, power calculations indicated that at least 50 centers per provider were required to identify statistically significant differences between providers. To this end, we defined a target sample of 300 centers, with 50 centers per provider. After SEDESOL was excluded from the study, we kept a target sample of 300 centers by increasing the sample per provider from 50 to 60 centers, plus an additional reserve of 20 centers per provider.

During the field operation, especially the facilitation phase, we found that many of the centers in the sample did not serve children under the age of three, so we had to use the reserve centers and even expand the sample more than once. We contacted 1,391 centers, of which 1,054 were excluded and another 90 declined to participate.³² All told, we visited and collected information on 247 centers (Table A.I.1). For these centers, we made a list of 958 eligible classrooms (for children under the age of 36 months).³³ From this list, we randomly selected 497 classrooms for the target sample. Of these classrooms, we excluded 28—usually because more than 25% of the children were older than three—and we did not obtain the required parental consent for another 39. We also excluded classrooms in which more than 25% of the households declined to participate. Finally, at the request of two providers, we also excluded classrooms in which at least one of the families did not agree to their children being filmed.

During meetings with parents at the centers, we compiled a list of 5,531 households, of which 24% did not agree to participate, 14% did not attend the meeting with parents, and 6% had children over the age of 36 months. Therefore, the list was reduced to 3,117 eligible children, from whom we randomly selected 2,345 for inclusion in the study.

Information was therefore collected for a total of 247 centers, 430 classrooms, and 2,311 households distributed among the providers (P1 to P5), as shown in Table A.I.1.

Table A.I.1. Sample of centers, classrooms, and households

	P1	P2	P3	P4	P5	Total
I. Centers						
Target sample	60	60	60	60	60	300
Expanded target sample + replacements	89	65	1,092	70	75	1,391
Obligatory exclusions	6	2	1,034	1	11	1,054
Consent declined	36	10	12	15	17	90
Centers visited and surveyed	47	53	46	54	47	247
II. Classrooms						

³² These centers were excluded because they did not serve children under the age of three (98.1% of exclusions), because they were closed (1.3%), or for safety reasons (0.6%). Of the decisions not to participate, 40% were made by directors, 28% by parents, and 7% by school unions.

³³ For one provider, classrooms with children under the age of 12 months were not included, at that provider's request.

	P1	P2	P3	P4	P5	Total
Eligible classrooms	179	277	62	270	170	958
Classrooms sought	107	115	55	119	101	497
Obligatory exclusions	8	7	2	1	10	28
Consent declined	10	2	1	20	6	39
Classrooms visited, surveyed, and filmed	89	106	52	98	85	430
III. Households						
Eligible households	529	831	294	1,030	433	3,117
Households sought	456	575	282	611	421	2,345
Obligatory exclusions	0	4	2	2	9	17
Consent declined	3	9	0	0	5	17
Households visited and surveyed	453	562	280	609	407	2,311

Appendix II. Index construction

II.1 Indexes for characterizing households

We constructed the following indexes using principal components analysis:

- (i) A wealth index using information on household assets and housing characteristics;
- (ii) Two indexes measuring the quality of the home environment for fostering child development: one for varieties of toys and the other for play activities based on information gathered using the FCI; and
- (iii) Two indexes for expressive language with the words the child “says,” one for children aged 19 to 30 months and another for those aged 31 to 36 months, based on information gathered using the 50-word versions of the MacArthur-Bates Communicative Development Inventories II and III.

We also constructed these indexes for all Mexican households with children under age three using data from the 2018-2019 ENSANUT to compare the socioeconomic status of the households served by the child care centers with the overall distribution of Mexican households. Although we used variables common to both surveys, they do not use the same population or sampling frame (the distribution support). Therefore, to make the distribution supports comparable, we calculated the first principal component using data from the 2018-2019 ENSANUT, which is equivalent to the linear combination of variables from the index, as follows:

$$y_i = \alpha_{i1} x_1 + \alpha_{i2} x_2 + \dots + \alpha_{in} x_n$$

where y_i is the principal component i , the α_{ik} coefficients are eigenvectors, and the x coefficients are the subset of n variables included in each index. Next, we constructed a linear combination of the same variables in the surveyed households but using the α coefficients estimated with the 2018-2019 ENSANUT data.

Table A.II.1 sets out the 10 variables of the household wealth index, and Figure A.III.1 of Appendix III shows the distribution by provider and for all Mexican households with children under age three. We also calculated the percentage of households served by the centers whose wealth index values were comparable to those of the richest quintile of households in the 2018-2019 ENSANUT.

Table A.II.1 also lists the five variables that compose the index for varieties of toys and the seven variables in the index for varieties of play activities. Figure A.III.2 presents the distributions of these two indexes, together with the distribution of the number of books. To characterize the quality of the home environment, we calculated the percentage of households with at least three picture books, the percentage of households with at least three varieties of toys, and the percentage of households where at least four play activities were carried out with the children. We used the household survey we collected and the 2018-2019 ENSANUT data for children under three for these calculations.

For the vocabulary indexes, we used the 50 words spoken by the children in each age group: 19 to 30 months and 31 to 36 months. Figure A.III.3 presents indexes' distributions. We also calculated the percentage of children served by child care centers with index values above the average for their respective age group.

Table A.II.1. Components of the indexes for characterizing households

Index	Components
Household wealth	<ol style="list-style-type: none"> 1. Modern flooring materials 2. Number of rooms in the house 3. The household has a landline telephone 4. The household has an automobile 5. The household has a microwave 6. The household has a washing machine 7. The household has a water heater 8. The household has a computer 9. The household has a television 10. The household has a radio
Varieties of toys	<ol style="list-style-type: none"> 1. There are at least three picture books 2. The child plays with toys for assembling 3. The child plays with objects or dolls for role play 4. The child plays with toys for learning shapes and colors 5. The child plays with electronic toys
Varieties of play activities	<p>This index reflects whether anyone over age 15 did any of these activities in the last three days:</p> <ol style="list-style-type: none"> 1. Read a book to the child 2. Told the child a story 3. Sang with the child 4. Took the child on an outing 5. Played with the child with the child's toys 6. Made drawings or scribbles with the child 7. Play with the child that involved naming or counting objects

II.2. Indexes for measuring centers' structural quality

We constructed various indexes to measure structural quality. These indexes are the sum of various characteristics that were first converted into dichotomous variables (with values of 0 or 1). This information was reported by directors and caregivers on their respective surveys, although the field team supervisor verified various characteristics of the physical infrastructure and equipment by direct observation when visiting the center while the survey was being conducted. Table A.II.2 lists the indexes and their components. We used the same method to construct all structural quality indexes and confirmed that they included items that were conceptually related to the dimension being measured; to do so, we considered the providers' feedback. We also constructed the same indexes using factor analysis. However, since the results are qualitatively identical, for ease of interpretation, we present the results of the sum of component indexes.

Table A.II.2. Components of the structural quality indexes

Index	Components
Elements of infrastructure in good	<ol style="list-style-type: none"> 1. All child-accessible spaces¹ have roofing in good condition 2. All child-accessible spaces¹ have walls in good condition 3. All child-accessible spaces¹ have flooring in good condition

Index	Components
condition and without hazards	<ol style="list-style-type: none"> 4. The center has locks and barriers in the doors 5. There is protection where the floor level changes 6. There is a recreational area in good condition 7. The drainage system is suitable 8. All furniture is in good condition 9. There is an emergency exit
Characteristics of emergency exits	<ol style="list-style-type: none"> 1. Emergency exit is made of nonflammable material or is treated with a fire retardant 2. Emergency exit has no padlocks, locks with the bolt drawn, or other obstacles 3. Emergency exit swings open on a vertical access and opens outward 4. Emergency exit has appropriate signage 5. Emergency exit has mechanisms allowing it to be opened from inside by simply pushing on it 6. Emergency exit can be opened by an adult, but not by a child
Emergency equipment	<ol style="list-style-type: none"> 1. Gas detector 2. Smoke detector 3. Flashlight with batteries 4. Emergency stairway 5. Megaphone 6. Bell 7. Siren 8. Doorbell
Characteristics of bathrooms	<ol style="list-style-type: none"> 1. Appropriately sized toilets 2. All bathrooms have toilet paper 3. All bathrooms have soap 4. All bathrooms have items for drying hands 5. Toilets in good condition
Variety de play materials within children's reach	<ol style="list-style-type: none"> 1. Picture books 2. Toys for assembling or building 3. Objects for role play 4. Toys for learning shapes and colors 5. Toys requiring movement 6. Toys that make music 7. Materials for painting, drawing and/or writing 8. Coloring books
Variety of play activities	<ol style="list-style-type: none"> 1. Build with blocks or cubes 2. Scribble, draw, or color 3. Play with puzzles or matching games 4. Explore materials like animals, plants, wheels, levers 5. Sing, listen to music, or dance 6. Do physical activities involving gross motor movements 7. Use manipulable materials like Legos, boards 8. Count, learn numbers 9. Cut and paste pre-drawn shapes 10. Learn the colors 11. Learn the letters of the alphabet 12. Learn shapes and sizes 13. Listen to stories for a long time 14. Use paint, modeling clay, or other art supplies 15. Are taught to problem solve 16. Are taught the parts of the body
Variety of director of caregiver's training	<ol style="list-style-type: none"> 1. Concepts, milestones, and how to foster child development 2. Pedagogy, aspects of play, and activities that promote learning 3. Music and other artistic activities 4. How to treat children in a gender- and age-appropriate manner 5. Discipline and behavior management 6. Curriculum planning and organization

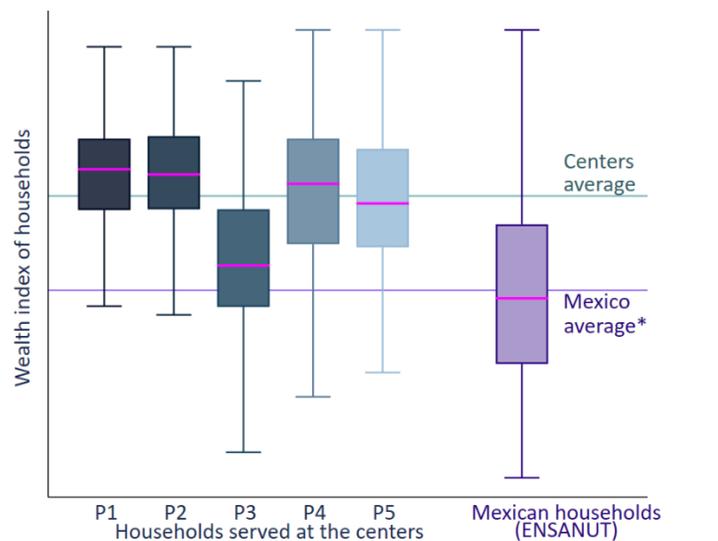
¹Child-accessible areas: multipurpose room, kitchen, nursery room, bathrooms.

Appendix III. Additional results from characterizing households served by the centers in comparison with all Mexican households with children under age three

Although we found that the public child care centers serve households of a higher socioeconomic status than the average Mexican household with children under age three, we also found a high level of variability in the households served by the providers. Using the methodology reported in Appendix II.1, we constructed indexes to characterize the children and households served by the child care centers using information from the household survey and from the 2018-2019 ENSANUT. Next, we plotted the distribution of these indexes in box-and-whisker plots to compare the households served by each provider with all Mexican households with children under age three.

Using Figure A.III.1 as an example, the box-and-whisker plots can be interpreted as follows: the box shows the interquartile range, meaning the values in the second and third quartiles of the distribution, or 50% of the data distributed between the 25th and 75th percentiles. The whiskers are the lines coming out of the boxes. They show how households are distributed in the lowest or highest quartiles. The longer the boxes and/or whiskers, the more variability there is in the data. For example, the box-and-whisker plots of Providers 1 and 2 are shorter than that of Provider 3, which serves households with a larger spread (greater variability) in the wealth index. The magenta line indicates the index's median. The upper box and whisker of Providers 1 and 2 would fit in the upper whisker of the household distribution plot in the 2018-2019 ENSANUT. This means that 75% or more of the households of these providers have wealth levels comparable to the richest 25% of all Mexican households with children under age three. Although Provider 3 serves households that have a lower socioeconomic level, the wealth level of nearly 75% of its households is above the median for all Mexican households with children of comparable ages.

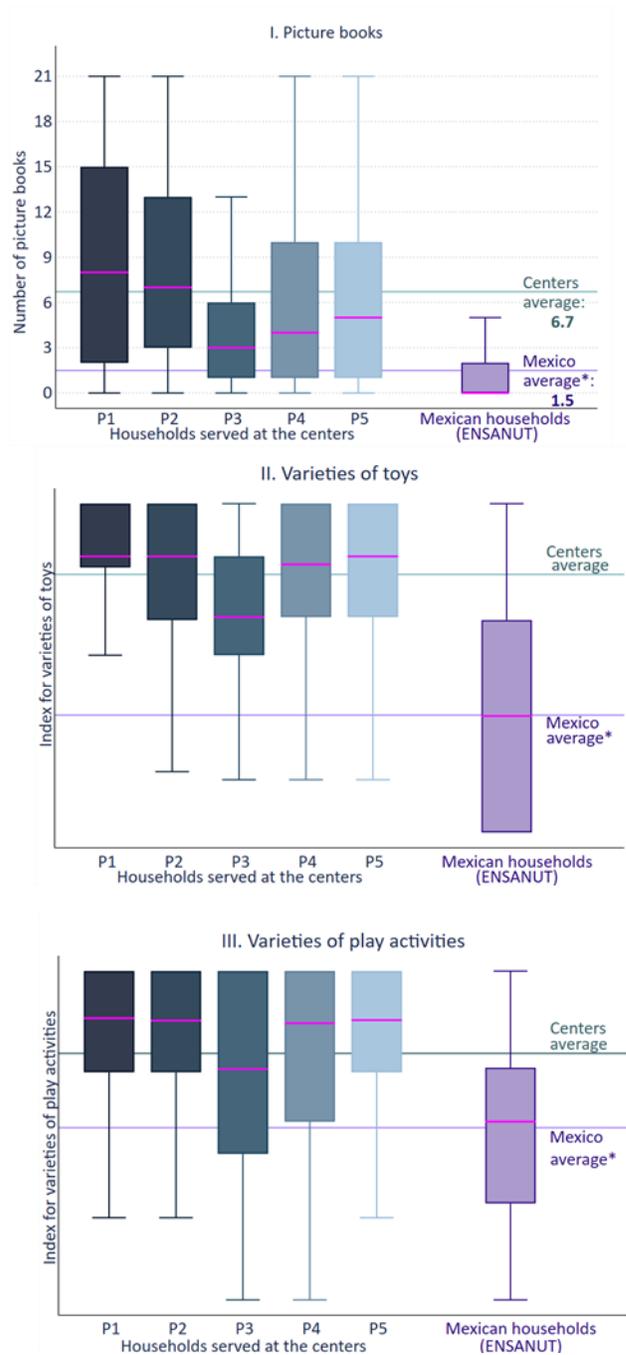
Figure A.III.1. Wealth index of households served at the centers, by provider, and comparison with households with children under age three in Mexico



Note: *National average for households with children under age three.

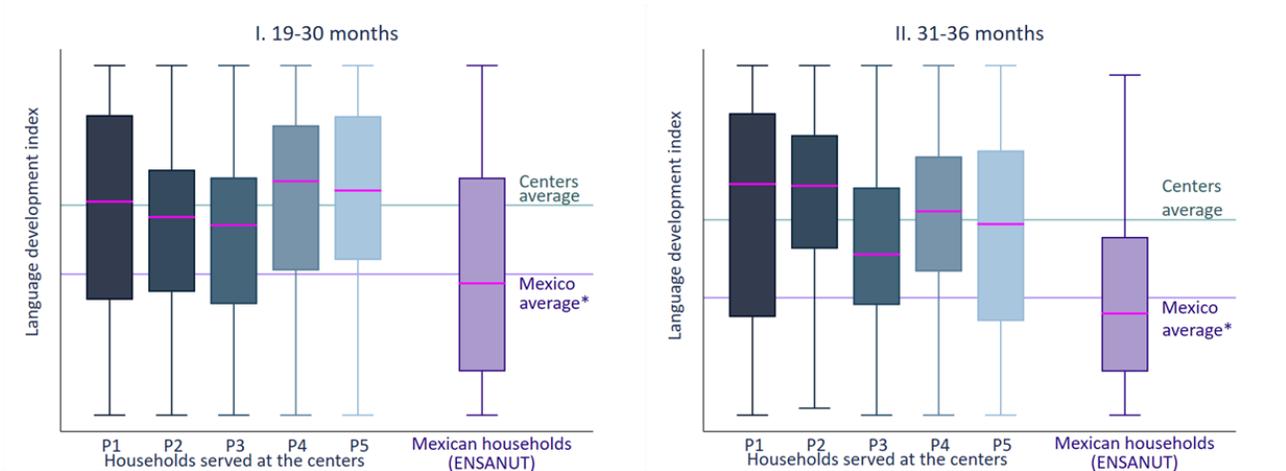
Figure A.III.2 shows the distribution of the number of picture books, the index for varieties of toys, and the index for the variety of play activities of the households served by each provider compared to all Mexican households with children under age three. On average, the children who attend the centers have 6.7 picture books, while the average for all Mexican households with children under age three is 1.5 picture books (Chart I). The varieties of toys available at slightly over 75% households served by the providers, with the exception of Provider 3, is similar to that of the top 25% of all Mexican households (Chart II). Likewise, approximately 75% of the households served by the centers of Providers 1, 2 and 5 carry out the same variety in play activities as the top 25% of all Mexican households (Chart III). Although Provider 3 households have less variety in terms of books, activities, and play materials, around 75% of these households have a better play and learning environment than the average child in Mexico.

Figure A.III. 2. Varieties of books, toys, and play activities in the households served by the centers, by provider, and comparison with all households with children under age three in Mexico



Finally, the children served by the centers have higher levels of language development than children of similar ages nationwide. As Figure A.III.3 shows, language development levels vary widely among providers, although scores tend to be lower for children aged 19 to 30 months at the centers of Providers 2 and 3 and for children aged 31 to 36 months who attend Provider 3 centers.

Figure A.III.3. Vocabulary level of children served by the centers, by provider, and comparison with all children under age three in Mexico



Note: *National average for households with children under age three.

Appendix IV. Size of center, by provider

Table A.IV.1. Size of center, by provider

	Provider				
	P1	P2	P3	P4	P5
Service characteristics					
Months of child care offered per year	11.1	10.6	11.1	11.8	11.2
Hours per day the center is open	9.3	8.2	7.9	10.2	9.1
Size of center					
Center's total capacity	132.5	174.8	40.6	157.6	88.4
Number of classrooms	7.5	8.5	2.9	7.4	5.2
Has no infant classrooms, %	15.1	2.3	80.2	1.7	17.5
Has 1 infant classroom, %	22.1	14.6	4.8	11.0	55.3
Has 2 infant classrooms, %	30.6	32.6	14.0	32.2	8.6
Has 3 or more infant classrooms, %	32.2	50.5	1.0	55.1	18.7
Has no toddler classrooms, %	4.6	3.3	9.6	0.0	3.7
Has 1 toddler classroom, %	23.7	18.6	74.6	0.0	33.7
Has 2 toddler classrooms, %	46.1	45.7	15.7	18.3	29.5
Has 3 or more toddler classrooms, %	25.7	32.4	0	81.7	33.2
Has at least 1 pre-school classroom, %	94.7	87.0	90.0	100	71.1
Average attendance per classroom					
Total	13.9	16.5	9.8	18.0	12.5
Infants	9.2	12.0	5.8	15.1	7.1
Toddlers	12.0	17.2	8.1	19.8	11.3

Appendix V. Profiles

Table A.V.1 Director and caregiver profiles, by provider

	P1	P2	P3	P4	P5
I. Director					
Years of education	16.1	17.2	14.0	16.0	15.6
Has undergraduate degree, %	91.5	95.5	47.8	84.6	79.7
Years of experience	15.7	17.1	10.9	13.4	15.6
Length of service at the center	8.7	12.8	6.6	8.9	6.8
Hours worked per week	47.3	47.0	42.9	47.4	44.8
Monthly income (including bonuses), Mex\$	16,682	23,549	8,321	12,942	12,021
Training variety index	3.9	4.1	4.2	4.9	4.0
Hours of training	47.2	67.6	24.5	46.1	43.8
II. Caregiver					
Years of education	14.3	15.1	14.2	13.9	13.5
Has undergraduate degree, %	43.1	61.4	46.2	25.3	29.9
Years of experience	14.6	15.7	9.4	6.5	13.5
Length of service at the center	11.6	11.8	6.5	4.2	10.4
Hours worked per week	49.0	55.3	46.7	53.6	48.8
Monthly income (including bonuses), Mex\$	8,381	9,626	5,302	5,066	7,125
Training variety index	3.9	3.6	4.0	4.8	4.3
Hours of training	42.5	39.9	26.0	38.5	31.3

Appendix VI. Physical infrastructure quality and safety at centers

Table A.VI.1. Physical infrastructure quality and safety at centers, by provider

	Provider					Average
	P1	P2	P3	P4	P5	
Elements of infrastructure in good condition and without hazards						
There is an emergency exit	95.4	88.2	74.4	100.0	71.6	92.2
All spaces* have roofing in good condition	90.6	72.5	79.5	97.2	87.1	91.7
Doors at the center have protection (locks and barriers)	87.7	88.9	82.3	95.1	83.4	91.1
All spaces* have walls in good condition	87.0	78.0	76.9	95.6	82.9	90.0
All spaces* have floors in good condition	94.7	77.8	83.5	95.2	73.9	89.3
The drainage system is suitable	89.5	82.2	77.2	88.3	81.4	86.1
There is a recreational area in good condition	92.4	78.4	66.2	90.4	70.5	84.8
All furniture is in good condition	87.6	68.5	62.8	87.1	68.5	81.0
There is protection where the floor level changes	78.1	86.6	52.6	72.5	42.3	67.9
Characteristics of emergency exits						
Has appropriate signage	95.4	70.4	69.4	98.3	66.1	88.5
Can be opened by an adult, but not by a child	93.0	85.3	66.2	90.6	69.5	85.3
Has no padlocks, locks with the bolt drawn, or other obstacles	86.6	62.5	54.5	95.7	52.6	82.2
Can be opened from inside by simply pushing on it	88.6	63.4	48.6	92.3	58.5	81.2
Swings open on a vertical access and opens outward	84.4	77.4	46.7	86.1	61.3	78.7
Is made of nonflammable material or is treated with a fire retardant	84.5	62.0	22.4	76.9	53.2	69.2
Emergency equipment						
Smoke detector	93.7	68.4	77.1	96.4	66.4	87.5
Flashlight	81.9	63.4	33.4	95.8	47.4	79.8
Doorbell	80.7	68.8	61.3	77.7	85.3	77.6
Siren	71.9	55.2	16.0	90.0	46.0	73.5
Emergency stairway	57.7	59.1	34.1	74.8	50.8	65.3
Gas detector	67.3	41.4	16.8	81.4	35.5	65.1
Megaphone	75.6	66.0	33.7	62.9	38.9	58.6
Bell	52.2	58.9	20.2	52.1	46.1	49.8
Characteristics of bathrooms						
All bathrooms have soap	84.2	88.7	85.5	100	90.1	95.0
All bathrooms have toilet paper	76.8	92.2	92.8	90.7	89.1	89.3
All bathrooms have items for drying hands	79.2	83.1	71.5	95.6	73.3	87.7
All toilets appropriately sized	78.5	84.2	61.2	85.3	71.2	80.8
All toilets in good condition	53.3	52.5	44.5	57.5	50.5	54.7
Other aspects of physical infrastructure quality and safety						
All power outlets are located inside walls*	97.3	100	95.8	98.7	97.8	98.3
Any chemicals present are stored in a secure area	98.8	91.9	100	98.4	100	98.3
Any chemicals present are identified	100	97.7	100	96.6	98.9	97.6
The stairways are safe	96.9	96.4	99.0	96.0	96.0	96.3
All switches are wall mounted*	95.3	98.0	97.9	97.4	90.6	96.1
All wiring is protected	94.3	96.7	97.6	95.2	90.8	94.6
Outdoor playground equipment is safe	98.3	98.1	94.6	94.7	90.5	94.6
Stairways are protected	91.6	87.0	95.3	93.6	95.1	93.2

	Provider					Average
	P1	P2	P3	P4	P5	
Access to hazardous areas is blocked	94.8	86.8	92.5	96.3	84.9	93.2
All power outlets are protected *	87.9	92.2	94.3	95.8	70.8	90.3
Fencing is in good condition or unnecessary	91.9	86.6	92.5	91.8	79.4	89.3
Open spaces are fenced in	93.8	81.1	79.8	87.1	69.1	83.7
Access to stairways is blocked	75.5	41.8	88.4	70.8	83.7	72.0

Appendix VII Varieties of play materials

Table A.VII.1. Percentage of centers with each variety of toy within children's reach, by provider

	Provider				
	P1	P2	P3	P4	P5
Toys for learning shapes and colors	85.4	90.5	79.3	91.4	80.8
Toys for assembling or building	84.5	83.6	77.3	92.3	77.3
Picture books	84.2	86.5	74.8	82.1	74.0
Coloring books	74.2	71.9	51.1	83.3	72.1
Objects for role play	72.8	77.3	66.4	82.3	59.5
Materials for painting and/or writing	62.0	67.9	54.8	80.8	63.7
Toys requiring movement	72.7	70.4	54.1	65.3	55.8
Toys that make music	66.6	72.5	63.1	61.4	48.2
Electronic toys	24.4	24.6	26.9	24.7	19.7

Appendix VIII. Varieties of play activities

Table A.VIII.1. Percentage de classrooms where each play activity is carried out at least weekly, by provider

	Provider					Average
	P1	P2	P3	P4	P5	
Sing, listen to music, or dance	96.2	95.4	97.5	98.7	98.7	98.2
Do physical activities involving gross motor movements	94.6	96.8	89.5	98.0	95.2	97.1
Learn parts of the body	94.7	93.8	94.4	97.3	96.8	96.6
Build with blocks or cubes	90.1	77.8	97.5	98.7	91.9	95.1
Scribble, draw, or color	82.4	58.8	89.8	92.9	91.4	88.5
Listen to stories for a long time	81.8	80.6	89.5	90.5	83.4	87.9
Use paint, modeling clay, or other art supplies	79.6	80.5	96.7	88.6	93.4	87.9
Explore materials like animals, plants, wheels	80.3	78.1	89.9	88.4	92.4	87.3
Learn to problem solve	84.4	91.2	88.7	83.8	90.0	85.4
Learn the colors	79.6	60.4	93.8	90.4	80.7	85.4
Use manipulable materials like Legos, boards	72.6	73.9	87.8	85.0	88.1	83.4
Play with puzzles or matching games	71.2	57.3	89.1	86.3	85.2	82.2
Learn shapes and sizes	74.0	65.9	90.9	82.9	90.0	81.6
Count or learn numbers	59.3	44.4	89.5	84.3	72.0	77.0
Learn the letters of the alphabet	37.5	23.4	68.9	56.6	38.8	49.7
Cut and paste pre-drawn shapes	32.7	16.7	58.8	55.2	44.2	48.3

Appendix IX. Process quality, by dimension and provider

Figure IX.1. Process quality by provider: average value for the dimensions of Emotional and Behavioral Support

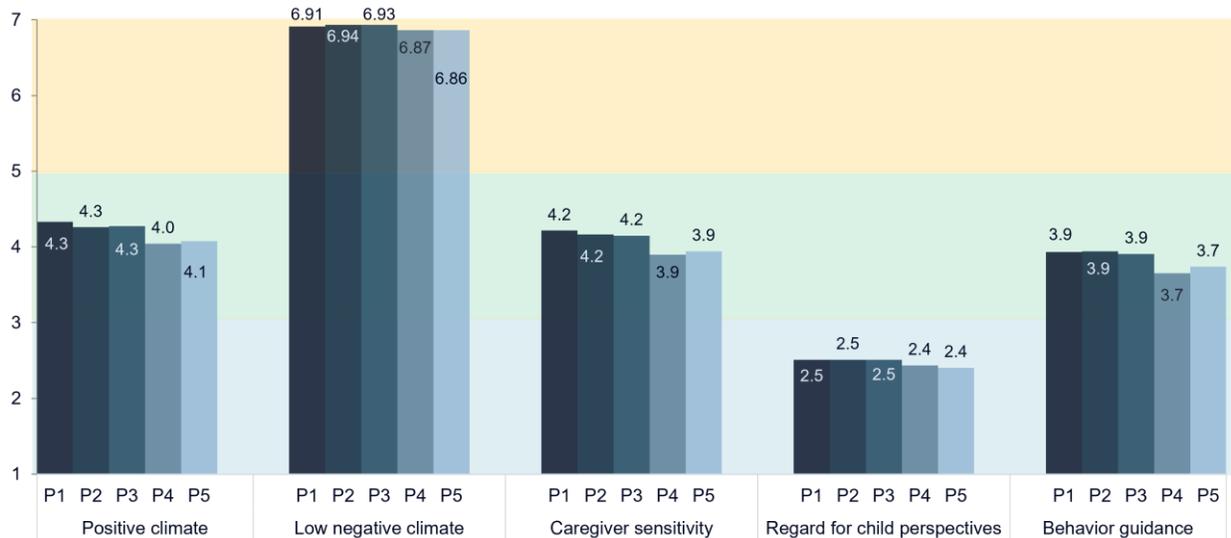
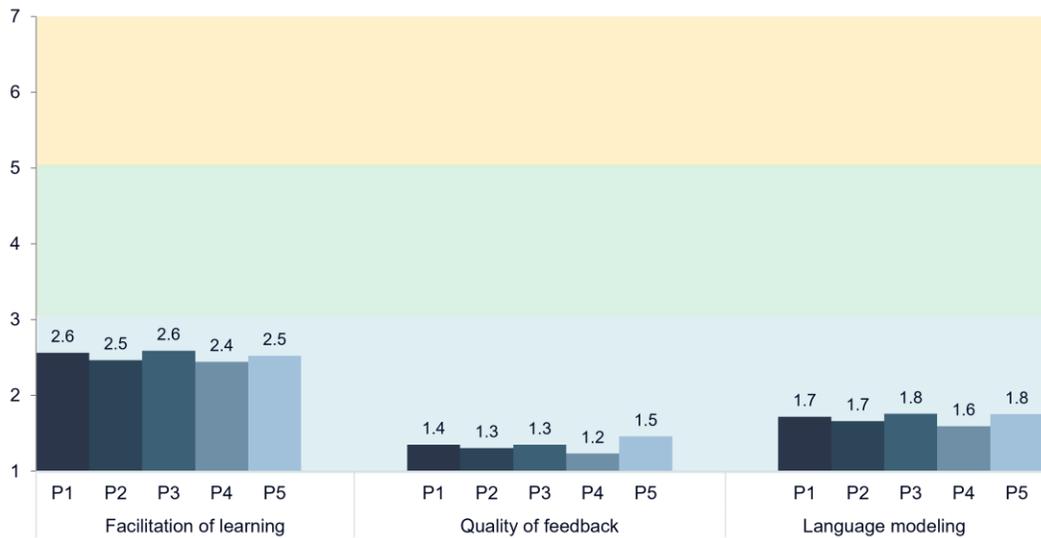


Figure A.IX.2. Process quality by provider: average value for the dimensions of Engaged Support for Learning



Appendix X. Correlations between the vocabulary level of the children served by the child care centers and the socioeconomic characteristics of their households

Table A.X.1. Correlations between the children’s vocabulary level and the characteristics of their households

	Vocabulary ¹
I. Household characteristics	
Level of wealth	0.073*
Mother’s age (years)	0.035
Mother’s level of education (years)	0.014
II. Quality of home environment	
Number of picture books in the household ²	0.120**
Varieties of toys in the household ²	0.176***
Varieties of play activities ²	0.124***

¹ Vocabulary tests were standardized by subtracting the mean, then dividing the result by the standard deviation in each month of age. This eliminates the effect of age from the two inventories and allows the results to be combined to produce a single correlation with vocabulary using a larger sample size. ²The number of books and the varieties of toys and play activities were also standardized to remove the effect of age because these increase as the children grow. Statistically significant at 5% (*), 1% (**), and 0.1% (***).