

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

Florencia Lopez Boo
María de la Paz Ferro Venegas



Catalogación en la fuente proporcionada por la
Biblioteca Felipe Herrera del Banco Interamericano de Desarrollo

Lopez Boo, Florencia.

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 / Florencia Lopez Boo, María de la Paz
Ferro.

p. cm. — (Monografía del BID ; 733)

Incluye referencias bibliográficas.

1. Child care services-Argentina. 2. Day care centers-Argentina. 3. Child development-Argentina.

I. Ferro, María de la Paz. II. Banco Interamericano de Desarrollo. División de Protección Social y Salud.

III. Título. IV. Serie.

IDB-MG-733

<http://www.iadb.org>

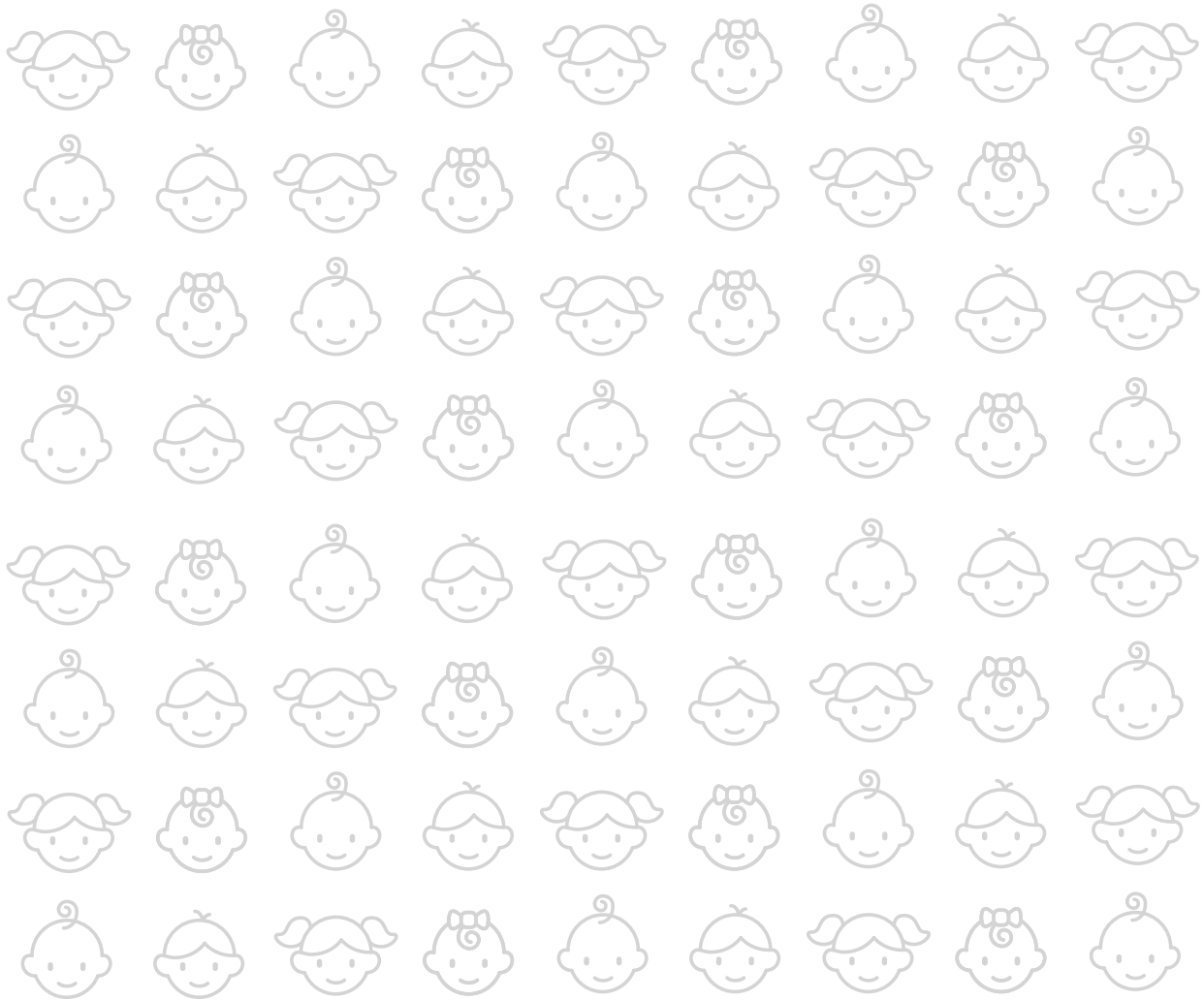
Copyright © 2019 Banco Interamericano de Desarrollo. Esta obra se encuentra sujeta a una licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas (CC-IGO 3.0 BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando el reconocimiento respectivo al BID. No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI (UNCITRAL). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID, no están autorizados por esta licencia CC-IGO y requieren de un acuerdo de licencia adicional.

Note que el enlace URL incluye términos y condiciones adicionales de esta licencia.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.





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

Florencia Lopez Boo
María de la Paz Ferro Venegas



La calidad de los procesos y el desarrollo infantil en los Espacios de Primera Infancia del Gran Buenos Aires: validación de una lista corta de monitoreo de centros infantiles *

Resumen

Este estudio analiza información recogida en el año 2017 en una muestra de Espacios de Primera Infancia (EPI) del Gran Buenos Aires en Argentina. Su propósito es caracterizar diferentes aspectos de la calidad de los servicios que los EPI prestan a los niños menores de 3 años por medio de dos instrumentos de amplio uso internacional que miden la calidad de los servicios de cuidado (ITERS y CLASS). Asimismo, pretende validar una lista corta de monitoreo de la calidad de procesos para su uso frecuente y a escala. El principal hallazgo es que, si bien los EPI estudiados presentan márgenes de mejora en lo que se refiere a algunos aspectos estructurales y de procesos, los centros muestran también fortalezas, en particular en los dominios de *Apoyo emocional y conductual* (CLASS) y de *Interacciones* (ITERS). La calidad de proceso de los EPI es, en general, superior en comparación con otros programas de la región; resalta la magnitud de la diferencia en *Clima positivo*, *Orientación de la conducta* y *Modelaje lingüístico*. No obstante, el puntaje de los EPI en *Modelaje lingüístico* permanece en el rango medio bajo, lo que también permite mejoras. Los cambios más complejos -pero también los más necesarios- tienen que ver con aspectos de la calidad de procesos y, en particular, con aspectos relacionados con la formación de la fuerza laboral encargada de los niños y con la construcción de competencias que permitan ofrecer experiencias que promuevan el aprendizaje en todas las áreas del desarrollo infantil. Todos los instrumentos presentaron excelentes propiedades psicométricas en su primera aplicación en Argentina. A medida que la atención en los jardines maternos continúe expandiéndose, es crucial desarrollar métodos para monitorear su calidad. Por ello la principal contribución de este estudio fue identificar empíricamente un subconjunto de 32 ítems que se centran en aspectos críticos de la calidad de procesos y que se pueden recopilar de forma rutinaria. Encontramos, además, que estos 32 ítems no solo se correlacionan con ITERS (alto: 0,88) y CLASS (moderado: 0,45); sino que logran discriminar aulas de diferentes niveles de calidad.

Clasificación JEL: I00, I10, I20, I25, I30, I38, J13

Palabras clave: Servicios de cuidado infantil, guarderías, salas cuna, calidad, Argentina, Espacios de Primera Infancia, desarrollo infantil, CLASS, ITERS, Bayley, PRUNAPE, herramientas de supervisión de la calidad.

*Este trabajo fue realizado por un equipo del Banco Interamericano de Desarrollo, liderado por Florencia Lopez Boo (florencia@iadb.org) y que consta de: Mayaris Cubides, Sarah Schodt, Cynthia van der Werf, María de la Paz Ferro y Claudia Viñuela Ortiz. Este trabajo fue posible gracias al apoyo incondicional de las autoridades del SENAF, el Ministerio de Desarrollo Social de la Nación y el Ministerio de Desarrollo Social de la Provincia de Buenos Aires. En particular agradecemos a Santiago López Medrano, Analía Corbalán y Lorena Ríos por facilitar el ingreso a los centros, y a Javier Quesada por la motivación para lograr llevar a escala esta herramienta. Los autores reconocen también el excelente trabajo realizado por el equipo de MORI (Argentina) y de Kantar Millward Brown (Ecuador) durante el levantamiento de la información. Los autores agradecemos por sus valiosos aportes a las autoridades del Gobierno argentino ante quienes se presentó un avance y la versión final de estos resultados en noviembre de 2017 y en marzo de 2018. En particular, agradecemos a las directoras y educadoras de los EPI que nos abrieron sus salas y con quienes compartimos estos resultados en Agosto 3 de 2018. También reconocemos los valiosos comentarios compartidos por Analía Corbalán, Milagros Nores y un revisor anónimo, como parte del protocolo de publicaciones del BID.



Resumen Ejecutivo

En este informe hemos documentado, por un lado, aspectos claves del desarrollo infantil utilizando dos instrumentos rigurosos: la escala de lenguaje receptivo del Bayley (considerado el instrumento regla de oro de la medición del desarrollo) y la PRUNAPE (el único instrumento con normas argentinas). Por otro lado, hemos analizado la calidad de los EPI del Gran Buenos Aires en Argentina, evaluándolos con la ayuda de dos instrumentos validados internacionalmente: el CLASS y dos subescalas del ITERS, así como una Lista de Verificación de uso más sencillo y costo-efectivo. Esta lista fue específicamente diseñada por el equipo de investigación del BID y recopila información sobre la calidad de los procesos, en particular los que se dan entre educadora-niños pero también niño-niño y entre educadora-padres. Ésta pretende ser una herramienta de fácil capacitación y administración, que permita al personal de supervisión obtener datos confiables y frecuentes sobre la calidad del EPI.

Los instrumentos mencionados fueron aplicados por investigadoras universitarias capacitadas en una muestra representativa de 95 EPI. En la última sección del documento, tras documentar los niveles de calidad y desarrollo infantil, analizamos en detalle los elementos de la Lista de Verificación con el fin de seleccionar aquellos que podrían ser adecuados para constituir un instrumento más corto y menos costoso. La versión reducida de la Lista de Verificación se ofrece como una herramienta alternativa no solo para caracterizar y monitorear la calidad en la prestación del servicio sino, en la medida de lo posible, para proporcionar retroalimentación frecuente a los programas y a las educadoras en el ejercicio de sus funciones.

Todos los instrumentos presentaron excelentes propiedades psicométricas en su primera aplicación en Argentina. Igualmente, la Lista de Verificación se correlacionó bien con los instrumentos internacionales más complejos administrados en paralelo (alto con ITERS y moderado con CLASS) y con el salario de la coordinadora del centro en particular.

Un tercio de las salas de los EPI incluidos en este estudio opera a tiempo completo y el resto son de jornada simple. Las educadoras observadas tienen, en promedio, 14 años de educación (el 80% posee educación terciaria incompleta o superior), 6 años de experiencia y la mitad ha trabajado con niños en su empleo anterior. Cada una está a cargo de grupos de 15 niños en promedio, lo cual puede dificultar interacciones de calidad ya que se superan las recomendaciones internacionales que sugieren un máximo de 6 niños por educadora en este grupo etario.

Por otro lado, los niños de la muestra tenían alrededor de dos años y medio, pero se encontraban en salas donde cerca del 40% de los niños estaban fuera del rango etario que deberían atender. Esto es indicativo de una estructura mixta por edades y podría tener implicaciones en la calidad del servicio. Más de la mitad de estos niños presentaron rezagos en su desarrollo.

En cuanto a la infraestructura de los EPI, se observa que solo la mitad tiene juegos infantiles en buen estado y 4 de cada 10 EPI tienen una sala de bebés, a pesar de que la normativa de estos espacios es atender a niños desde los 45 días de edad. Uno de cada cinco EPI tiene buena calefacción y aire acondicionado, lo que presenta uno de los desafíos estructurales más importantes para el bienestar de los niños en el centro. Las estadísticas tampoco son muy alentadoras en cuanto a la protección contra incendios, ya que solo el 3% tiene un detector de gas, el 14% tiene detector de humo, un tercio ha hecho simulacros de evacuación y menos de la mitad tiene señales de evacuación.

Los resultados CLASS son alentadores porque sugieren que muchos niños reciben una atención cálida y sensible. Sin embargo, los resultados de *Apoyo comprometido con el aprendizaje* sugieren que la facilitación de actividades para promover el desarrollo del niño se encuentra en un nivel medio bajo. Por ejemplo, se ha encontrado que los educadores podrían guiar más el aprendizaje y el desarrollo de los niños durante las actividades o rutinas con su participación activa; existen, asimismo, pocas ocasiones donde los educadores provean andamiaje verbal o físico para que un niño logre hacer algo que no puede hacer por sí solo o donde alienten la metacognición para ayudar a los



niños a pensar sobre sus procesos de aprendizaje.

Por otro lado, si bien el puntaje de la subescala de *Interacciones* de ITERS indica un nivel de calidad mínimo, casi la mitad de los centros tienen un puntaje de 5 o mayor (nivel de calidad bueno o superior) en los ítems *Supervisión del juego y aprendizaje*, *Interacción entre los niños*, *Interacción entre el personal y niños*. Por último, la subescala de *Escuchar y hablar* demuestra que hay desafíos en relación al desarrollo de habilidades lingüísticas y pre lectura y, en particular, el ítem *Uso de libros* demuestra una ausencia generalizada de libros y actividades de lectura con los niños.

Finalmente, este estudio identifica un subconjunto de 32 ítems que se centran en aspectos críticos de la calidad de procesos y que, además de presentar buenas propiedades psicométricas, pueden ser recopilados de forma rutinaria como parte de los esfuerzos de monitoreo de los servicios de cuidado infantil. La observación en aula es corta (dos horas) y se puede hacer con personal no profesional, siempre que la capacitación sea la adecuada. Es importante destacar que este instrumento logra discriminar aulas de diferentes niveles de calidad, aun en el rango alto de la distribución.

Este estudio provee una base concreta para los programas de la región que busquen abordar un sistema de mejora continua de la calidad que sea costo-efectivo, frecuente, sistemático y que emplee instrumentos rigurosos. En situaciones ideales, las listas de monitoreo cortas y simples a nivel censal deberán ser complementadas con el uso de instrumentos globales validados y aplicados con una periodicidad menor y a nivel muestral. Este conjunto de información permitirá identificar áreas críticas de mejora y destinar los recursos necesarios para que se lleven a cabo.



1. Introducción

En años recientes, muchos países en la región han realizado inversiones importantes con el objetivo de expandir la cobertura de sus servicios de desarrollo infantil, en particular de aquellos que brindan servicios de cuidado (guarderías, salas cuna, jardines maternos, etc.) a los niños pequeños. El objetivo principal de estas reformas ha sido facilitar la incorporación de las mujeres al mercado laboral. Lamentablemente, muchas veces se ha pasado por alto la importancia de asegurar ambientes de alta calidad para los niños, usuarios de estos servicios, durante un período crítico de su desarrollo. La evidencia internacional sugiere que servicios de cuidado de baja calidad tienen un impacto nulo o, incluso, negativo sobre los niños (Baker, Gruber y Milligan, 2008, 2015; Howes et al., 1992; Bernal, Attanasio, Peña y Vera-Hernández, 2017; Araujo, Dormal y Schady, 2018; Rosero y Osterbeek, 2011).

Desde el lanzamiento por parte del presidente argentino de un Plan Nacional de Primera Infancia (PNPI) en abril de 2016, el Ministerio de Desarrollo Social de la Nación Argentina (MDS) está llevando a cabo reformas importantes que afectarán la organización y la operación de los Espacios de Primera Infancia (EPI en adelante). Los EPI son centros que ofrecen atención integral, contención y estimulación a niños desde los 45 días hasta los 5 años de edad. Asimismo, si bien el Ministerio de Educación no forma parte del PNPI, también aporta a las políticas prioritarias de primera infancia. Una primera reforma liderada por el Ministerio de Educación tiene que ver con la universalización de las salas de 3 años.¹ Una segunda reforma liderada por el MDS y la Secretaría de Niñez, Adolescencia y Familia (SENAF), en el marco del PNPI, tiene que ver con mejorar la calidad del servicio que prestan los EPI. Estos esfuerzos de mejora de la calidad se concentran en: infraestructura, coeficiente educador-niño, metros cuadrados por niño, mobiliario, materiales, capacitación, supervisión de los recursos humanos y de la calidad de procesos en cada aula, entre otros.²

Entre el 10 de julio y el 20 de diciembre de 2017 se recogió información sobre una muestra de 95 EPI con subsidios de distintas instancias públicas, pero gestionados por organizaciones de la sociedad civil del Gran Buenos Aires (GBA), en Argentina. Este estudio presenta un análisis de dicha información con el objeto de caracterizar, por primera vez para Argentina, diferentes aspectos de la calidad de los servicios que estos prestan a niños menores de tres años.³ Asimismo, se proponen 32 ítems identificados empíricamente para el monitoreo frecuente a escala. Dada la evidencia para la Argentina de heterogeneidad e irregularidad en los sistemas de monitoreo (FLACSO/UNICEF/CIDECC, 2019), el instrumento propuesto se vuelve particularmente importante. El estudio fue elaborado por el Banco Interamericano de Desarrollo, por solicitud del Ministerio de Desarrollo Social de Argentina. Se espera que estos hallazgos informen el proceso de mejoras que se encuentran atravesando los EPI.

¹Diario Clarín, consultado 11 de Abril 2016: https://www.clarin.com/politica/Mauricio-Macri-Nacional-Primera-Infancia_3_1556874306.html

²El Banco Interamericano de Desarrollo está acompañando al Gobierno argentino en estas reformas por medio de una operación de Crédito Condicional para proyectos de inversión (CCLIP AR-O0003) incluida en el Decreto 822/2017, aprobado en julio del 2017. En particular, este crédito busca, entre otras cosas, fortalecer EPI existentes (no construir nuevos ni expandirlos) en dos ámbitos: i) fortalecimiento estructural de 800 EPI e ii) iniciar un proceso de monitoreo y mejora de las interacciones entre educadoras y niños en el universo de EPI. Se trabajará, además, en 10 nuevos EPI para probar una metodología de estimación de demanda y generar aprendizajes operativos.

³Un estudio reciente de los EPI que atienden niños en la Ciudad Autónoma de Buenos Aires -llamados CPI- estudió parte de estos aspectos de manera cualitativa y para una muestra no representativa (FLACSO/UNICEF/CIDECC, 2019). Encontraron que existe una alta heterogeneidad en las formas de monitoreo de la calidad de los servicios y acompañamiento técnico. En particular, en la mayoría de los centros encuestados, los mecanismos de monitoreo eran irregulares o poco estandarizados, mediante devolución a docentes, una vez al año, de informes sobre las fortalezas y las debilidades identificadas por la directora o intercambios informales durante el día.



2. Los Espacios de Primera Infancia

Los Espacios de Primera Infancia⁴ son centros de cuidado de niños de 45 días a 5 años de edad que, en mayor o menor medida, están abocados a la estimulación del desarrollo. Su personal consta de una coordinadora, educadores y personal de apoyo (limpieza, cocina, portería y, a veces, nutricionista y trabajador social). La coordinadora es la persona encargada de las tareas administrativas y de mantenimiento de los centros, que incluye la programación de las actividades pedagógicas y el fortalecimiento de las capacidades de las educadoras. Por otro lado, las educadoras están en contacto directo con los niños, encargadas de su cuidado y educación.

Estos centros surgieron masivamente a fines de los 90 y durante la crisis post-2001 y dependen de gobiernos municipales, organizaciones de la sociedad civil, provincias y, algunos, del gobierno nacional. Por lo general, no están regulados como centros de cuidado y justamente el PNPI busca iniciar una agenda de trabajo en esta dirección. Existen en todo el país al menos 3.000 centros con, por lo menos, 180 mil niños en condiciones muy vulnerables, de acuerdo con el Registro Nacional de Espacios de Primera Infancia (RENEPI, en adelante). De estos 3.000 centros, un tercio se ubican en la Provincia de Buenos Aires⁵ y aproximadamente 630 centros en el Conurbano de dicha provincia.⁶ La base de datos RENEPI (2015) muestra que un tercio de los centros del Conurbano eran públicos y el resto comunitarios, es decir, centros cofinanciados entre el sector público y la sociedad civil; y a ellos asistían un total de 65.000 niños entre 45 días y 5 años. De estos, 135 conforman el universo de este estudio ya que son los centros que se adhieren al PNPI, se concentran en los 24 partidos de los cordones 1 y 2 del GBA y tienen, al menos, una sala de 2.^{7,8}

3. Métodos

3.1. Descripción de la muestra y comité de ética.

La población de la cual se seleccionó la muestra para este estudio consiste en los 135 centros elegibles arriba descriptos. De estos, 32 rechazaron participar en el estudio, 2 se utilizaron en las prácticas y 6 en el piloto (que se excluyen de la muestra aquí utilizada), con lo que la muestra final resultó en 95 centros. Estos centros operan en promedio con 8,73 educadoras y 15,25 niños por educador (ver Tabla 4.1). De acuerdo con los cálculos de poder (Apéndice A) con 90 centros y 5 niños evaluados por centro, el efecto mínimo detectable de correlación entre calidad y desarrollo infantil será de 0,15 desvíos estándar (DE) (CIC⁹=0,1), 0,17 DE (CIC=0,2) y 0,19 DE (CIC= 0,3). La unidad de análisis para este estudio fue un grupo de niños y su(s) educadora(s) –en adelante grupo–

⁴En la Provincia de Buenos Aires, muchos de estos espacios de primera infancia se encuentran enmarcados en el Programa UDI (Unidades de Desarrollo Infantil) dependiente del Ministerio de Desarrollo Social de la Provincia de Buenos Aires (MDS PBA) bajo la modalidad Jardín Maternal Comunitario. Históricamente, se les conoce también como jardines maternos o jardines sociales, entre otras denominaciones.

⁵Un estudio del INDEC de 2010 identificaba 577 centros en toda la provincia, con lo que se nota un crecimiento de 60 centros en 6 años, o 10 al año.

⁶El Conurbano Bonaerense se divide territorialmente en 229 localidades y barrios, dentro de los 24 municipios que lo componen, denominados constitucionalmente partidos. (SIEMPRO/INDEC, 2010)

⁷El primer cordón está constituido por los partidos de: Avellaneda, Lanús, Lomas de Zamora, La Matanza (parte este), Morón, Tres de Febrero, San Martín, Vicente López y San Isidro. El segundo cordón se constituye de: Quilmes, Berazategui, Florencio Varela, Almirante Brown, Esteban Echeverría, Ezeiza, Moreno, Merlo, Hurlingham, Ituzaingó, La Matanza (parte oeste), Tigre, San Fernando, José C. Paz, San Miguel y Malvinas Argentinas.

⁸De los 194 EPI que inicialmente conformaron el universo se llegaron a los 135 elegibles porque al verificar la información telefónicamente, varios no tenían sala de 2 (49 casos), otros no pertenecían al GBA o estaban repetidos dentro de la misma base de datos y en 2 casos no fue posible entablar contacto.

⁹Correlación intra-cluster.



por cada EPI. La selección de cada grupo se hizo siguiendo un protocolo de selección aleatoria establecido por el equipo del BID. La selección y anuncio del grupo que se iba a estudiar se realizó el mismo día en que se llevó a cabo la visita y no con anticipación, para evitar que el EPI cambiara sus rutinas normales por nuestra presencia. A su vez, dentro de cada grupo se seleccionó de manera aleatoria a los niños a los que se les aplicarían los tests de desarrollo. El número mínimo de niños se definió previamente por los cálculos de poder. La muestra final resultó en 930 niños.

Antes de salir a campo, el Consejo Nacional de Políticas Sociales presidido por Roberto Candiano, y oficiando como comité de ética, aprobó el protocolo de este estudio (incluyendo el consentimiento de los padres para evaluar y observar a los niños).¹⁰

3.2. Instrumentos

Instrumentos de medición de calidad en la sala

Es importante notar que, a lo largo de este análisis, la calidad se concibe como global, es decir, como el conjunto de variables estructurales (que identifican la ausencia o no de determinados recursos y, por lo general, son fáciles de regular) y de procesos (aquellas que se focalizan en aspectos más dinámicos; por ejemplo, las interacciones educador-niños) que facilitan y promueven el desarrollo infantil. Se decidió administrar múltiples instrumentos que miden la calidad de los servicios de cuidado porque cada instrumento tiene un enfoque diferente. Por ejemplo, algunos instrumentos cubren aspectos estructurales y de procesos, mientras que otros se enfocan con mayor profundidad únicamente en uno u otro aspecto. Además, se tenía el interés de explorar la correlación entre instrumentos (o su validez concurrente) en un contexto de países en desarrollo, en particular el argentino, y de validar un instrumento de monitoreo frecuente de la calidad nunca antes usado en la región.

Los instrumentos que se usaron en este estudio para caracterizar la calidad de los servicios de cuidado de los EPI son los siguientes (ver Tabla B.1 para mayor detalle sobre cada instrumento):

1. El Sistema de Observación y Calificación en el Aula para niños párvulos, CLASS párvulos (CLASS de ahora en adelante) (Pianta, La Paro and Hamre, 2008). El CLASS párvulos es un instrumento de tipo observacional desarrollado para evaluar la interacción entre el niño y su educador. Se enfoca exclusivamente en variables de proceso. Su medición está sujeta a una observación presencial o filmación de una jornada de cuatro horas, la cual se divide en cuatro segmentos de veinte minutos de observación cada uno. En este estudio se utilizó la primera forma de aplicación, observación presencial. El CLASS párvulos explora ocho dimensiones de la calidad de procesos: a) *Clima positivo*, b) *Clima negativo*¹¹, c) *Sensibilidad del educador*, d) *Consideración por la perspectiva del niño*, e) *Orientación de la conducta*, f) *Facilitación del aprendizaje y el desarrollo*, g) *Calidad de la retroalimentación* y h) *Modelaje lingüístico*. Estas dimensiones se agrupan a su vez en dos dominios: a) *Apoyo emocional y conductual* (que comprende las cinco primeras) y b) *Apoyo comprometido con el aprendizaje* (que comprende las tres últimas). Cada dimensión recibe una calificación del 1 al 7, donde 1 representa un nivel inadecuado y 7 un nivel excelente. El CLASS ha sido utilizado en varios países en Latinoamérica, incluyendo Chile, Ecuador y Perú. (Duración para el estudio/en campo: 2,5 horas

¹⁰El comité estaba conformado por Fabiana Rubinstein, Leandro Gielis y María Eugenia Perez Ponsa. Se llevó a cabo en la ciudad de Buenos Aires. Del comité de ética salieron recomendaciones sobre el contenido del consentimiento informado y de los cuestionarios, así como la necesidad de adaptarlo al contexto local. El equipo incorporó estos cambios y los documentos fueron aprobados en su totalidad el día 12 de junio de 2017.

¹¹En esta dimensión se reversa el puntaje, de tal forma que entre mayor es el puntaje menor es el clima negativo en el grupo.



por grupo/aula- ver sección 3.3.2 para detalles del periodo de observación en este estudio).

2. Las subescalas *Escuchar y hablar* e *Interacción* de la Escala de calificación del ambiente para bebés y niños pequeños, edición revisada (ITERS-R, por sus siglas en inglés, o *Infant and Toddlers Environment Rating Scale, Revised edition*) (Harms, Cryer y Clifford, 2006). El ITERS-R (en adelante, ITERS) es un instrumento que combina la observación con una corta encuesta. Probablemente es el instrumento que más se ha aplicado en los ejercicios de medición de la calidad de los servicios de cuidado para infantes y párvulos tanto en América Latina como en otros países del mundo (Berlinski y Schady, 2015). En América Latina se ha aplicado en Bolivia, Ecuador, Perú, Chile y Colombia, entre otros. Requiere de un mínimo de tres horas de observación de un aula independientemente del número de subescalas aplicadas (Harms, Cryer y Clifford, 2006). Este instrumento cuenta con siete subescalas, de las cuales como se mencionó anteriormente se seleccionaron dos, *Escuchar y hablar* e *Interacción*. Estas dos subescalas se destacan por su enfoque en las interacciones entre los educadores y los niños (en vez de elementos estructurales) y son las que mejor predicen el desarrollo de los niños (Berlinski y Schady, 2015). Además, para los efectos de nuestro estudio, son las más comparables al instrumento CLASS, proveyendo información que se alinea para profundizar nuestro entendimiento de los procesos que contribuyen al desarrollo de los niños en los EPI. Los Ítems de la subescala *Escuchar y hablar* son *Ayuda a los niños a entender el lenguaje*; *Ayuda a los niños a utilizar el lenguaje*, y *Uso de libros*. Los Ítems de la subescala de *Interacción* son *Supervisión del juego y aprendizaje*; *Interacción entre los niños*; *Interacción entre el personal y niños*, y *Disciplina*.¹² (Duración en campo: 3 horas por grupo/aula).
3. Una lista de verificación tipo *checklist*, más simple que CLASS e ITERS, que busca recoger información de la calidad de procesos para ser incorporada a un sistema de monitoreo frecuente. Incorpora elementos de los instrumentos ITERS y el CC-IT-HOME¹³ que aparecían como los más promisorios en términos de rigor psicométrico y factibilidad en otros estudios de investigación (Lopez Bóo, Dormal y Weber, 2019). Así, la Lista de Verificación (LV de ahora en adelante) consiste en 66 ítems agrupados en 12 constructos: *Receptividad*, *Organización*, *Materiales de aprendizaje*, *Involucramiento*, *Variedad*, *Rutinas de cuidado personal*, *Prácticas de salud*, *Escuchar y hablar*, *Actividades*, *Interacción*, *Estructura del programa* y *Siesta*. La LV identifica si los ítems contenidos en cada uno de los constructos se encuentran o no presentes en el centro, es decir, la codificación es 0/1 para facilitar la observación y codificación de parte de supervisores con un perfil no necesariamente profesional. El puntaje total corresponde al porcentaje de ítems presentes en la sala. La LV se pilotea por primera vez en la región en este estudio. (Duración estimada: 2 horas por grupo/aula).

La Lista de Verificación fue construida por el equipo del BID con el objetivo de proveer a los gobiernos de la región una herramienta simple y costo-efectiva para el monitoreo frecuente, la supervisión y el seguimiento de la calidad de proceso de los proveedores de servicios, con un foco particular en las interacciones educador-niño. En primer lugar, se mapeó cada una de las subescalas de los instrumentos de medición de la calidad más usadas (CLASS, ITERS y CC-IT-HOME) a trece constructos de interacción educador-niño considerados fundamentales para

¹²De esta forma, el puntaje total ITERS computado en el presente estudio corresponde únicamente al promedio de los Ítems de las subescalas recolectadas (*Escuchar y hablar* e *Interacción*).

¹³El CC-IT-HOME es una adaptación del conocido IT-HOME, que evalúa la calidad del ambiente en el hogar. Por su parte, el CC-IT-HOME se centra en la calidad de los procesos, aunque también evalúa la presencia de variables estructurales, en entornos de cuidado (hogar o institucional). Es un instrumento principalmente de observación y permite que algunos de los ítems sean evaluados por entrevista. Cada ítem tiene dos opciones de respuesta (presente o ausente), y se agrupan en seis subescalas: *Materiales de aprendizaje*, *Receptividad*, *Aceptación*, *Organización*, *Involucramiento* y *Variedad*. (Bradley, Cadwell y Corwyn, 2003)



el desarrollo de los niños (Halle, Anderson, Blasberg, Chrisler, y Simkin, 2011). Se encontró que los constructos positivos y neutros que miden estos tres instrumentos son muy similares, sin embargo 3 de los constructos negativos (desapego, consideración negativa y sentimientos negativos) no se medían explícitamente en estas tres herramientas por lo que quedaron fuera de los constructos factibles de ser medidos. Así es que la LV se construye a partir de 8 constructos positivos, 1 neutro y 1 negativo, factibles de medición, que se enumeran en la Tabla C.1. En segundo lugar, y con el objetivo de seleccionar un número menor de subescalas que fueran las más relevantes, se usaron datos secundarios de los tres instrumentos mencionados recolectados en 400 centros de Ecuador (Araujo et al, 2015) y se realizó un análisis factorial confirmatorio para estos 10 constructos latentes (Tabla C.1). Se encuentra que CLASS tiene las cargas factoriales más altas con los constructos analizados, en particular con Sensibilidad/Receptividad, Consideración positiva/calidez y Atención conjunta. Por otro lado, cinco subescalas de ITERS (*Escuchar y hablar, Interacción, Rutinas de cuidado, Actividades y Estructura del programa*) y cinco del CC-IT-HOME (*Receptividad, Involucramiento, Organización, Variedad y Materiales de aprendizaje*) presentan cargas factoriales altas. En conclusión, como varias subescalas de ITERS y CC-IT-HOME muestran resultados alentadores con varios constructos que no presentaron una asociación tan alta con el CLASS, así como por las altas cargas factoriales de los primeros dos instrumentos, y la mayor complejidad del ajuste de un puntaje continuo (1-7) a uno binario (0-1) para CLASS¹⁴, se decidió excluir CLASS y utilizar únicamente ITERS y CC-IT-HOME para la construcción de la LV. Se eliminan subescalas con carga factorial por debajo de 0,30 y se preseleccionan las 10 subescalas mencionadas del CC-IT-HOME y del ITERS. En tercer lugar, y en base a un análisis de teoría de respuesta al ítem (TRI), se seleccionaron 66 ítems que resultaron de filtrar aquellos ítems con mayor carga factorial (mayor a 0,4) de las 5 subescalas ITERS y las 5 de CC-IT-HOME. La lista final de los 66 ítems se encuentra en la Tabla C.2. Estos 66 ítems se organizaron finalmente en 12 constructos, provenientes de un análisis factorial exploratorio adicional donde una estructura de 12 variables latentes (10 de ellas correspondientes a las 10 subescalas de los dos instrumentos originales de donde provienen los ítems y 2 adicionales *Siesta y Prácticas de Salud*) proporcionó un mejor ajuste que un modelo con solo las 10 subescalas de los instrumentos originales.

4. Un cuestionario detallado sobre la infraestructura y la seguridad del EPI, así como de otras características estructurales del servicio (por ejemplo, número de niños por adulto en cada grupo, tamaño y características de los espacios, etc.) siguiendo los estándares actuales y fichas de supervisión de los EPI. Este cuestionario fue validado previamente en Ecuador (Araujo, López Boo, Novella, Schodt y Tomé, 2015) y se basa en listas de supervisión de varios países de Latinoamérica. (Duración estimada: 45 minutos por EPI).
5. Entrevista tipo encuesta a ser aplicada a el coordinador del EPI y a el educador para recoger información individual (nivel educativo, experiencia, condiciones laborales, entre otros aspectos), así como información sobre los costos de operación, la matrícula y la asistencia del jardín y de los diferentes grupos. Para algunos elementos como la matrícula se utilizaron insumos, en particular, los listados de los centros. (Duración estimada: 45 minutos por coordinador y 45 minutos por educador).

¹⁴CLASS, por ejemplo, no posee indicadores dentro de cada ítem para su puntuación como sí tiene ITERS. Fueron estos indicadores de ITERS los que nos permitieron recuperar ítems binarios para la LV.



Instrumentos de medición de los niveles de desarrollo de los niños

Para medir el nivel de desarrollo de los niños en los EPI de este estudio se aplicaron los siguientes instrumentos:

1. Prueba Nacional de Pesquisa (PRUNAPE, por sus siglas) (Lejarraga et al., 2014). Es un instrumento de tamizaje de problemas del desarrollo infantil para niños menores de 6 años, diseñado, estandarizado y validado en el contexto argentino. Consta de 79 ítems que evalúan cuatro dominios del desarrollo: personal-social, motricidad fina, lenguaje y motricidad gruesa. El número de ítems administrados depende de la edad del niño y del objetivo de la aplicación de la prueba. Los objetivos de la administración de la prueba pueden ser: i) determinar si el niño tiene un retraso probable en su desarrollo, y ii) determinar las habilidades máximas del niño, es decir, el techo.¹⁵ Para su correcta administración el niño debe estar en buenas condiciones (buena salud, estado de vigilia, presencia del cuidador principal), el espacio libre de ruidos y de distracciones, y el encuestador adecuadamente capacitado (Lejarraga et al., 2013). El tiempo de administración oscila entre 12 y 15 minutos por niño. (Duración estimada: 15 minutos por niño).
2. La subescala de lenguaje receptivo del test de desarrollo infantil Bayley-III (Bayley, 2006). Bayley-III (de ahora en adelante, Bayley) es una herramienta de evaluación del desarrollo infantil reconocida internacionalmente por sus excelentes propiedades psicométricas y su sistema de puntuación cuantitativo (Pearson, 2018). Tiene cinco escalas; una motora, una cognitiva, una de lenguaje, una socio-emocional y una de comportamientos adaptativos. Por restricciones de tiempo en el EPI (se contaba con el niño únicamente por 10 minutos para aplicarle esta prueba) se aplicó únicamente una subescala. Se escogió la subescala de lenguaje receptivo por sugerencia de una experta en temas de primera infancia del *National Institute for Early Education Research*, al ser una de las subescalas que mejor discrimina el desarrollo del niño en las edades sujeto del estudio. Adicionalmente, se ha encontrado que el lenguaje receptivo es un importante predictor del desempeño escolar (Schady, 2012) y de resultados de largo plazo incluyendo salud mental y empleo (Schoon, Parsons, Rush y Law, 2010; Law, Rush, Schoon y Parsons, 2009). (Duración estimada: 10 minutos por niño).

3.3. Procesos de campo

3.3.1. Contacto con los EPI y selección de participantes

Previo al inicio del trabajo de campo, el 15 de mayo de 2017 se llevó a cabo una reunión con los referentes¹⁶ de los EPI de los cordones uno y dos del Gran Buenos Aires, con el objetivo de presentar el estudio y familiarizarlos con la dinámica del trabajo de campo. Posteriormente, SENAF enviaba los datos de contacto de los referentes y la firma MORI (firma principal encargada de la recolección de los datos) entablaba una comunicación vía correo electrónico y llamada telefónica. Finalmente, el referente proveía los datos de contacto de los EPI a su cargo.

¹⁵Si el objetivo de la aplicación de la prueba es el segundo, además de administrar los ítems asociados con el primer objetivo, se deberá continuar administrando ítems a la derecha hasta encontrar el techo (referirse al manual).

¹⁶El referente es aquella persona a la cual acuden las familias del barrio, ya sea para poder dejar a su hijos en la UDI o para trasladar diversas problemáticas que pueden estar atravesando (violencia doméstica, falta de trabajo, adicciones, etc.); es aquella persona que articula con el Estado las necesidades de la comunidad. Asimismo, los referentes están a cargo de organizaciones de la sociedad civil, organizaciones no-gubernamentales, fundaciones o municipios con cierto número de EPI. En nuestra muestra, cada referente tiene un promedio de 1,65 EPI a su cargo (con un máximo de 11 y un mínimo de 1). Son empleados de gobiernos municipales, provinciales u organizaciones de la sociedad civil, pero nunca del gobierno Nacional.



El contacto con el EPI se dio telefónicamente y contemplaba un mínimo de tres llamadas. Durante la primera llamada, se procuraba tener contacto con la coordinadora, se presentaba el estudio, se indagaba por la presencia de niños en el rango etario de interés del estudio, el tipo de jornada y el horario de trabajo de la educadora a cargo. También se describían las jornadas de trabajo de campo que se llevarían a cabo, se enviaba el consentimiento informado que debía ser firmado por los padres de los niños y se fijaba la fecha de visita (idealmente 5 días hábiles después del primer contacto).

La segunda llamada se planeó para realizarse dos días hábiles después del primer contacto y tenía por objetivo verificar si el EPI estaba realizando el proceso de recolección de consentimientos informados. En caso de que el EPI estuviera teniendo éxito, se confirmaba la fecha de la visita, en caso contrario se reagendaba la visita (idealmente cuatro días hábiles después). Por último, se realizaba una tercera llamada el día previo a la visita para confirmarla e informar el nombre de las investigadoras que harían las observaciones del grupo y aplicarían las entrevistas y las pruebas de desarrollo.¹⁷

Cada uno de los 95 EPI de la muestra fue visitado al menos una vez por parte de un equipo de dos investigadores, durante un total de una jornada completa o dos jornadas de medio día, para aplicar los instrumentos antes descritos con excepción del CLASS. En una submuestra de 55 EPI, se sumaba un tercer investigador para la aplicación del CLASS al mismo tiempo y en el mismo lugar que el investigador 2 aplicaba la LV y las 2 subescalas ITERS (ver Figura 1). Dado que el personal de este servicio es mayoritariamente femenino, se solicitó que al menos 1 de los investigadores del equipo de campo fuera mujer con el fin de generar más confianza entre el personal de los centros y los investigadores.

Al llegar a cada uno de los EPI, las investigadoras procedían a seleccionar al grupo de niños y su(s) educadora(s) que serían sujeto de sus observaciones a lo largo de la jornada. La selección de este grupo se basaba en el mismo criterio en todos los centros visitados. A pesar de que el objeto de estudio es la calidad del servicio en niños menores de tres años, para la selección del grupo nos enfocamos en aulas con niños entre los 24-36 meses (sala de dos) pues algunos EPI no atendían a los niños más pequeños. Este criterio de selección aseguraba una mayor comparabilidad de las muestras entre EPI. De esta forma, lo primero que se hacía al llegar al EPI era identificar cuántos grupos dentro del centro tenían niños que se encontraban en ese rango etario.¹⁸ En todos los casos, la edad de los niños se estimaba al momento del inicio del año lectivo en cada EPI. Una vez realizada la selección de un grupo por EPI, las investigadoras enfocaban el proceso de evaluación en el grupo seleccionado y en su educadora (o educadoras si hubiera más de una). Para definir los niños a los que se les administrarían las pruebas de desarrollo, en cada grupo seleccionado se repetía el proceso de selección aleatoria, pero únicamente con aquellos cuyos padres habían firmado el consentimiento informado. La PRUNAPE se administró en promedio a 10 niños por aula, y el Bayley a 7.

Con el fin de garantizar la comparabilidad de la observación de los datos de calidad entre EPI, y dada la importancia que tienen los horarios sobre las rutinas de los niños de este grupo de edad, los investigadores siguieron una rutina muy precisa en cada EPI que garantizaba que cada instrumento se aplicara en el mismo momento de la jornada para los EPI de jornada completa y para los EPI de turno mañana (que comienza alrededor de las 8 de la mañana en casi todos los centros encuestados).

¹⁷Si bien se buscó seguir el protocolo descrito, en la práctica debió realizarse un mayor número de contactos. En promedio para lograr agendar la visita debieron realizarse 10,8 llamadas telefónicas.

¹⁸Si se trataba de EPI que contaban únicamente con un grupo de niños de esta edad, entonces éste era el grupo de estudio. Si existía más de un grupo con niños en este rango de edad, se daba prioridad al que tuviera a todos los niños del grupo en ese rango. Por ejemplo, si había dos grupos, uno con niños de 0-36 meses y otro con niños de 24-36 meses, se escogía el segundo. En aquellos casos en que existía más de un grupo en el cual todos los niños tenían entre 24 y 36 meses (2 casos en nuestra muestra), entonces se escogía uno de los grupos al azar, siempre que hubiera un mínimo de 10 niños en la sala en la edad correspondiente al estudio. Para seleccionarlo al azar, las investigadoras asignaban números a cada grupo y tiraban un dado para la selección.



Para los EPI de turno tarde (37,89 % de la muestra) la aplicación de los instrumentos de la calidad de los procesos se realizó entonces durante la tarde, de nuevo alineados entre sí. En cada centro se aplicaron los instrumentos de medición de la calidad de manera simultánea. La Figura 1 describe la rutina de trabajo del equipo de investigadoras en cada centro cuando el centro es de jornada completa, así como quién estuvo a cargo de la aplicación de cada uno de los instrumentos. La nota que acompaña a la figura explica las rutinas cuando los EPI eran de turno simple.

3.3.2. Administración de instrumentos

La PRUNAPE y la totalidad de los instrumentos de medición de la calidad de los servicios, a excepción del CLASS, se administraron en los 95 EPI que conforman la muestra del presente estudio, y su aplicación estuvo a cargo de la firma MORI. Por su parte, la aplicación del CLASS estuvo a cargo de la empresa de investigaciones sociales de Ecuador, Kantar Millward Brown (Kantar de ahora en adelante), y fue aplicado en una submuestra de 55 EPI (60 % de la muestra). La submuestra corresponde a los 55 primeros EPI visitados durante el desarrollo del trabajo de campo de MORI.¹⁹ El CLASS se administró *in situ* bajo los siguientes periodos de observación: la codificación avanzó de acuerdo con ciclos de 30 minutos (20 minutos de observación y 10 minutos de registro) y un descanso de 10 minutos cada dos ciclos. Se obtuvo un mínimo de 4 ciclos, excepto para uno de los EPI, para el cual solo tuvimos 3 ciclos.²⁰

Las evaluaciones Bayley se levantaron aproximadamente un mes después de iniciado el trabajo de campo, entre el 17 de agosto y el 20 de diciembre de 2017. Esto se debió a la disponibilidad del material (hubo un retraso en el envío de los formularios desde España). La totalidad de las evaluaciones de Bayley fueron aplicadas por una única encuestadora, con acreditación de la editorial Pearson para la administración de la prueba. El Bayley se aplicó en 44 EPI (50 % de la muestra), de los cuales la mayoría coincidieron con la jornada programada por MORI para aplicar la prueba PRUNAPE.²¹ En estos casos, el Bayley se administró justo a continuación de la PRUNAPE. Para los que no coincidieron en la jornada, el periodo entre la aplicación de PRUNAPE y Bayley nunca fue superior a dos días.²²

¹⁹Para maximizar la cobertura del CLASS, las firmas debían coordinarse. En particular, MORI debía recolectar los consentimientos firmados por los padres de los niños para que Kantar pudiera entrar a los centros.

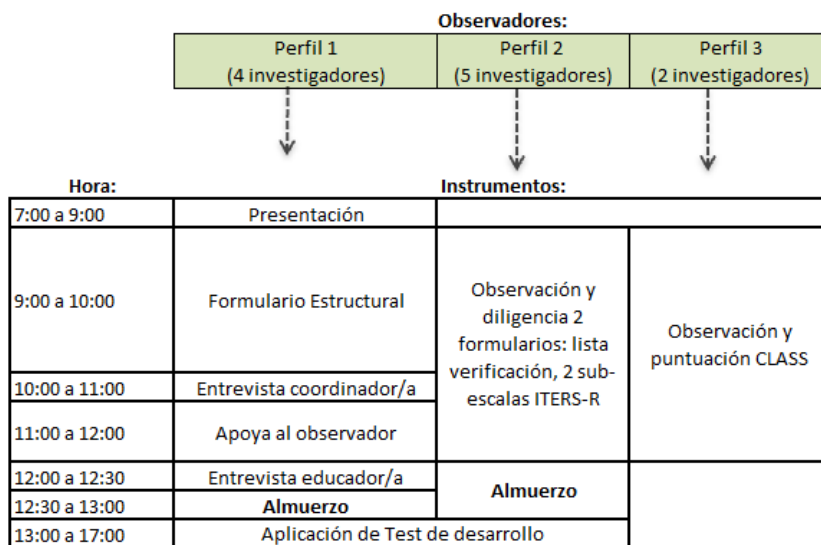
²⁰La dinámica en campo del CLASS se llevó a cabo de la siguiente manera: en 29 casos la observación se realizó en simultáneo con la aplicación de los instrumentos de calidad de sala (ITERS y Lista de Verificación) aplicados por MORI, ver Figura 1, con el fin de garantizar estricta comparabilidad en una jornada típica. En los demás casos (26), Kantar realizó la observación para puntuar el CLASS posterior a la visita de MORI. El tiempo promedio transcurrido entre la aplicación de ITERS y Lista de Verificación y la observación de CLASS fue de 9 días (un desvío estándar de 14 días).

²¹El desarrollo infantil solamente se midió en un momento del tiempo. En los análisis que utilizan estas medidas no se puede entonces controlar por el desarrollo inicial del niño, una limitación de este estudio.

²²Como Bayley y CLASS se aplicaron a submuestras de los EPI, solamente 25 EPI tienen información de la totalidad de los instrumentos.



Figura 1: Jornada de trabajo de campo



Perfil 1: educación superior, experiencia en temas de desarrollo infantil.

Perfil 2: educación superior, experiencia en temas de desarrollo infantil.

Perfil 3: educación superior, experiencia en temas de desarrollo infantil, observador CLASS certificado.

Nota: si el EPI era de medio turno (68% de los EPI), el levantamiento se realizaba en dos días, durante el primer día (mañana o tarde) se realizaba lo propuesto en esta figura para las horas de la mañana y en el segundo día (mañana o tarde) se realizaba la aplicación de las pruebas de desarrollo.



3.4. Capacitación

La capacitación de investigadores de campo partió de la elección de un equipo de 9 candidatas, cinco de ellas psicólogas, una psicopedagoga, dos licenciadas en educación y una en psicomotricidad. La capacitación en ITERS, Lista de Verificación y la PRUNAPE tuvo una duración de seis días, cuatro de los cuales fueron teóricos y dos correspondieron a días de práctica en dos EPI. Una psicopedagoga con maestría en atención temprana que cuenta con las certificaciones oficiales para la aplicación de la prueba PRUNAPE (otorgada por el Hospital Garrahan) y para el Bayley (otorgada por la editora Pearson) se encargó de la capacitación y supervisión de las prácticas.

Al culminar la capacitación, basados en el desempeño en un examen teórico y en la destreza de la aplicación de los diferentes instrumentos durante las prácticas y aplicación de los instrumentos en seis EPI que hicieron parte de la prueba piloto, se asignó a cada una de las investigadoras un rol específico que responde al diseño de la jornada típica de trabajo de campo en un EPI, tal como describe la Figura 1.

La supervisión en campo de la correcta administración del ITERS, Lista de Verificación, PRUNAPE y Bayley estuvo a cargo de la misma psicopedagoga anteriormente mencionada (se realizó la supervisión del 70 % de los formularios administrados.) Dado que PRUNAPE fue administrado por primera vez en un contexto de centros, el apéndice D brinda mayor detalle sobre el proceso y los desafíos encontrados.

Por su parte, el equipo de observadores del CLASS contaba con dos observadores experimentados de la empresa Kantar. Además de su experiencia observando y puntuado aulas en Perú y Ecuador durante más de 4 años de trabajo continuo, este equipo de codificadores recibió una capacitación adicional intensiva en el instrumento, tomó una prueba de certificación y estuvo acompañado a lo largo de todo el proceso de codificación por una investigadora certificada como entrenadora en CLASS.

3.5. Análisis

El análisis de la información en este reporte se lleva a cabo en dos partes. En la primera parte se ofrece una caracterización comprensiva de los centros (sección 4.1), luego se describen los resultados de cada uno de los instrumentos que dan cuenta de la calidad de los servicios y del ambiente de los EPI (sección 4.2) y del desarrollo cognitivo, motor y del lenguaje de los niños objeto de estudio (sección 4.3). Para los instrumentos que evalúan el desarrollo infantil se calcula tanto el puntaje bruto como el puntaje estandarizado internamente por edad. Este último se computa utilizando regresiones no paramétricas para eliminar los efectos de edad. Para el Bayley se prefiere estandarizar internamente los resultados y no las referencias internacionales pues no está claro que dichas referencias se pueden aplicar de manera adecuada a la población argentina.²³

En la segunda parte (sección 4.4), a partir de la LV se identifica empíricamente un subconjunto de 32 ítems críticos de la calidad y que son factibles de recopilar de forma rutinaria como parte de los esfuerzos de monitoreo de los EPI. Una vez identificados estos ítems, se realiza un análisis detallado de sus propiedades psicométricas. En particular, se revisa su consistencia interna, es decir, el grado en que los ítems de la escala miden el constructo subyacente y su validez concurrente, es decir, el grado en el que el instrumento se correlaciona con otros instrumentos que miden calidad. Para la consistencia interna se calculan los Alfa de Cronbach a nivel de constructo e ítem, y las correlaciones de Pearson entre el ítem y el puntaje de la escala. De otra parte, la validez concurrente se revisa a través de correlaciones entre el subconjunto de constructos y el CLASS y el ITERS. Adicionalmente,

²³El análisis de las propiedades psicométricas de los instrumentos de calidad en sala se encuentra en el Apéndice E, y el de las pruebas de desarrollo infantil en el Apéndice F.



se computan las correlaciones entre el instrumento corto y una batería de variables con las que se esperaba que estuviera correlacionado – en particular, variables que dan cuenta del desarrollo del niño y de la calidad estructural del EPI-. En todos los casos se usan umbrales establecidos en la literatura para revisar si las propiedades psicométricas se satisfacen (Furr y Bacharach, 2013; Araujo et al., 2015).

4. Resultados

4.1. Caracterización de los EPI

La muestra está compuesta por 95 EPI de 14 partidos de los cordones uno y dos del Gran Buenos Aires más el partido Pilar. El 47,37 % de los EPI se encuentran en el partido la Matanza y le siguen los partidos Tigre y Avellaneda con el 7,37 % y 6,32 %, respectivamente (ver Tabla 4.1). La distribución de EPI por partido sigue la distribución de la población en el Gran Buenos Aires, siendo La Matanza el partido con el mayor número de habitantes, ya que representa cerca del 30 % de la población en los partidos del presente estudio. (Ministerio de Economía PBA, 2016).

Como muestra la Figura 2, solo el 32,63 % ofrece servicios en jornada completa y el resto de EPI se dividen entre jornada simple en la mañana (29,47 %) y en la tarde (37,89 %). Además, como se muestra en la Tabla 4.1, el 48,42 % opera bajo un convenio con el municipio y el 34,74 % con convenio con SENAF (no son excluyentes este tipo de convenios). La edad promedio de los niños PRUNAPE (9,78 por EPI) es 33,76 meses y la de los niños Bayley (6,86 por EPI) es 35,05. De los 781 casos en los que la madre reportó la edad gestacional, el 16,39 % nacieron prematuros. Finalmente, la Figura 3 muestra que, si bien la mayoría de los niños en la sala de dos años tienen entre 25 y 36 meses, el 39 % está fuera de este rango etario. Por otro lado, la mayoría de niños en el EPI son mayores de 36 meses.

Los EPI operan con un promedio de 16,64 empleados, 8,73 educadoras y 128,06 niños inscritos (ver Tabla 4.1). De esta manera los coeficientes de atención son de 7,73 niños por empleado y 15,25 por educadora (es importante comparar este coeficiente de atención con el de los jardines del Ministerio de Educación: para este grupo etario (2-4), en el 2010 este coeficiente era de 11,8 niños por educadora, ver detalles en Araujo, Lopez Boo y Puyana (2013)). Por su parte, la sala de dos tiene en promedio 20,14 niños, y como se mencionó anteriormente, la mayoría se encuentra en el grupo etario que pretende atender. El promedio de años de funcionamiento de los EPI es 22,38, operan durante 11 o más meses al año y en el 65 % de los casos las familias deben realizar un pago, cuyo valor promedio mensual por niño es de \$106,99 pesos argentinos (equivalente al 1,21 % del salario mínimo).²⁴ La tendencia a realizar pagos es común en varios servicios públicos de América Latina y el Caribe (Araujo, et al., 2013). El 84 % de los EPI tienen estrategias de inclusión para niños con discapacidad y el 74 % estrategias de matriculación y participación para niños de minorías. El promedio de niños pertenecientes a minorías es de 32,64 (representando un cuarto de los inscritos en el EPI).²⁵ En lo que se refiere al registro de la información y el contenido del programa, el 74 % mantiene registros de información sobre peso y talla y solo el 66 % tiene guías curriculares (Tabla 2).

Se analizan las características de los EPI y su personal por tipo de jornada (no reportado). No se observan diferencias en las características del EPI o de su personal, solo vemos una mayor proporción de EPI de jornada completa (53 %) en partidos completamente urbanizados, en relación a los de jornada simple (34 %). También se observa que los salarios de las educadoras son alrededor

²⁴Los centros reciben en promedio un total de \$13.701,54 pesos argentinos de las familias/128.06 (que es el número promedio de niños inscritos por centro)=AR\$106,99.

²⁵Para la definición de minoría en el contexto de este estudio, ver nota en la Tabla 2.



de AR\$3.400 pesos más altos en los EPI de jornada completa, lo que refleja la mayor cantidad de horas trabajadas.

Tabla 1: La muestra: EPI y Niños

	N	Media/ Porcentaje	DE
Niños			
Edad (meses) niños PRUNAPE ^o	930	33,76	6,52
Edad (meses) niños Bayley ^o	302	35,05	5,48
Prematuro (%) [*]	781	16,39	
Prematuro menor de 2 años (%) ^{†*}	930	1,08	
EPI			
Número de educadoras del EPI	95	8,73	4,00
Número de empleados del EPI	95	16,64	6,59
Número de niños inscritos	95	128,06	81,20
Número de niños presentes	95	64,42	48,05
Niños (as) / número de empleados	95	7,73	3,64
Niños (as) / número de educadores	95	15,25	7,35
Población 0 a 4 en hogares con NBI (%) [‡]	93	22,80	
Jornada completa (%)	95	32,63	
Jornada simple: mañana (%)	95	29,47	
Jornada simple: tarde (%)	95	37,89	
Convenio con el municipio (%)	95	48,42	
Convenio con fundación u ONG (%)	95	9,47	
Convenio con SENAF (%)	95	34,74	
Partido			
Avellaneda (%)	95	6,32	
Berazategui (%)	95	1,05	
General San Martín (%)	95	5,26	
José C. Paz (%)	95	3,16	
La Matanza (%)	95	47,37	
Lanús (%)	95	2,11	
Lomas de Zamora (%)	95	6,32	
Malvinas Argentinas (%)	95	3,16	
Moreno (%)	95	3,16	
Pilar (%)	95	2,11	
Quilmes (%)	95	6,32	
San Fernando (%)	95	1,05	
San Isidro (%)	95	3,16	
San Miguel (%)	95	2,11	
Tigre (%)	95	7,37	

^o Edad en el momento de la aplicación de la prueba.

^{*} Variable reportada por los padres en el consentimiento informado.

[†] Para el cálculo de la edad en la PRUNAPE únicamente es relevante si el niño nació prematuro o no en el caso de que tenga menos de 2 años.

[‡] Fuente: SIEMPRO.

Nota: N corresponde al número de observaciones.



Figura 2: Proporción EPI por tipo jornada

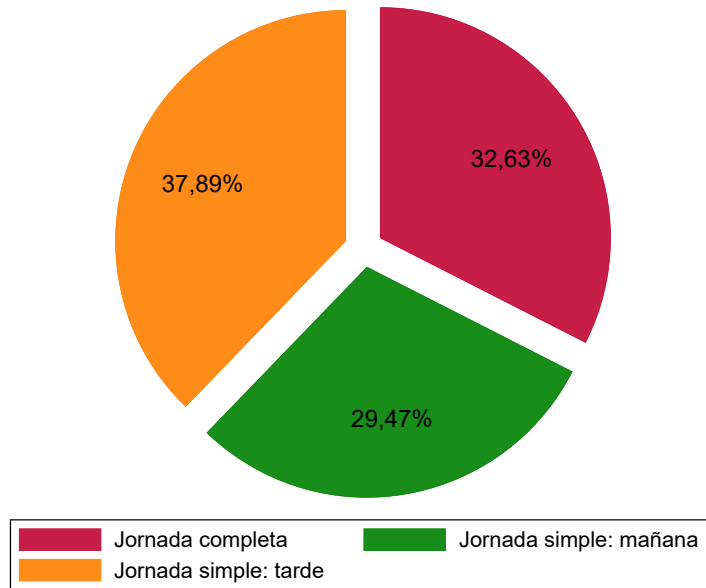


Figura 3: Composición etaria

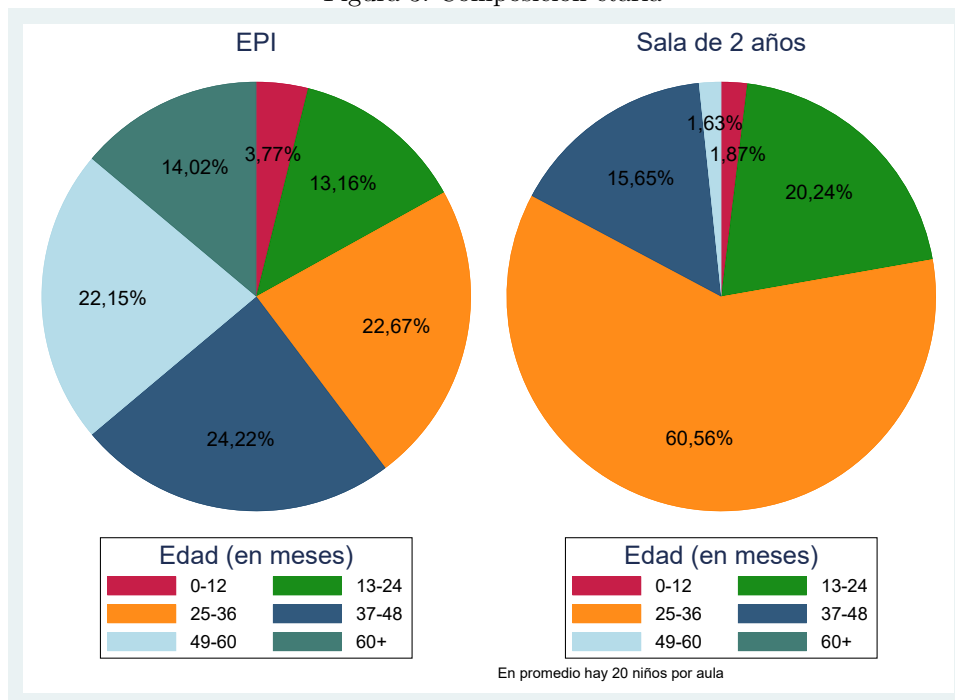




Tabla 2: Características EPI: funcionamiento, inclusión y manejo de información

	N	Media	DE	P-10	P-90
Funcionamiento					
Número de años de funcionamiento	95	22,38	9,93	8,00	32,00
Meses al año que ofrece atención	95	11,11	0,74	10,00	12,00
Días de vacaciones	80	26,20	14,24	12,00	40,00
Familias deben realizar algún pago	95	0,65	0,48	0,00	1,00
Pago promedio mensual por EPI (Pesos)	91	13701,54	25880,59	0,00	40000,00
Inclusión					
Estrategias para inclusión de niños con discapacidad*	95	0,84	0,37	0,00	1,00
Número de niños con discapacidad	95	2,05	2,16	0,00	5,00
Estrategias de participación para niños de minorías†	95	0,74	0,44	0,00	1,00
Número de niños pertenecientes a minorías	95	32,64	37,58	0,00	100,00
Apoyo sobre practicas de crianza‡	95	0,85	0,36	0,00	1,00
Apoyo sobre salud y nutrición§	95	0,78	0,42	0,00	1,00
Apoyo sobre cuidado prenatal°	95	0,21	0,41	0,00	1,00
Información y Currícula					
EPI tiene información sobre peso y talla	95	0,74	0,44	0,00	1,00
El EPI tiene guías curriculares	95	0,66	0,48	0,00	1,00

* Cualquier estrategia que el centro esté implementando que considera facilita o favorece la inclusión de niños con discapacidad. Ejemplos de estrategias mencionadas son: facilitar asesoramiento a los padres, rampas para acceder al comedor y asistencia psicológica.

† Cualquier estrategia que el centro esté implementando que considera favorece la participación de minorías. En este estudio, después de consultarlo con el comité de ética y varios especialistas, se concluyó que “minoría” en este contexto se asocia mayoritariamente con provenir de países limítrofes. Ejemplos de estrategias mencionadas son: respeto a costumbres culturales, incluir en charlas información de otros países e incentivar a los padres a participar activamente de las actividades del centro.

‡ Si el centro brinda información o apoyo a los padres, ya sea a través de comunicaciones o reuniones, sobre cuidado de los niños y buenas prácticas de crianza.

§ Si el centro brinda información o apoyo a los padres, ya sea a través de comunicaciones o reuniones, sobre salud, nutrición y desarrollo infantil.

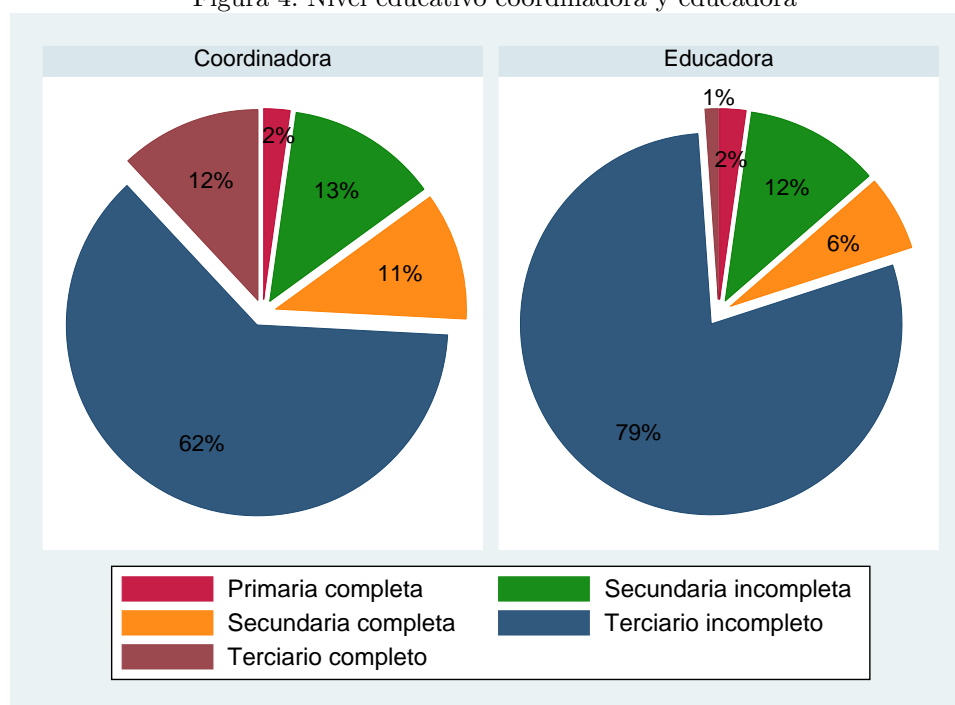
° Si el centro brinda información o asistencia a madres embarazadas y cuidado prenatal.



Características de las coordinadoras y educadoras

En la Tabla 3 se observa que el 74 % de las coordinadoras ha alcanzado la educación terciaria o superior (solo el 12 % ha culminado), el promedio de años de educación es 14,15 (lo que equivale a secundaria completa más un año de estudios técnicos -nivel terciario-), y el 48 % ha recibido un título en primera infancia. En lo que se refiere a la condición laboral, el salario mensual líquido promedio es de \$13.092,82 pesos (o 1,48 veces el salario mínimo), el 46 % tiene un empleo por nombramiento, el 70 % tiene vacaciones remuneradas y el 67 % realiza aportes a la seguridad social. Por otro lado, las educadoras a cargo de la sala de dos tienen un promedio de años de educación de 13,89, el 80 % ha alcanzado el nivel terciario o superior (solo el 1 % lo ha completado) y el 38 % tiene un título en primera infancia. Las condiciones laborales de la educadora son menos favorables que las de la coordinadora, ya que tienen un salario menor (\$8.568,39 pesos) y una menor proporción tiene empleo por nombramiento (20 %), aportes a la seguridad social (39 %) y vacaciones remuneradas (48 %). Esto posiblemente se relacione con las diferencias en el nivel educativo terciario completo (Figura 4). El salario promedio de las educadoras de los EPI de este estudio también es menor que el de las educadoras de jardines que dependen del Ministerio de Educación en la provincia de Buenos Aires (AR\$14.491²⁶), aunque las maestras de jardines similares en la Ciudad de Buenos Aires han cursado 2 años más de estudios que las educadoras de esta muestra, llegando a completar 16 años de educación (Araujo et al., 2013).

Figura 4: Nivel educativo coordinadora y educadora



²⁶Este valor corresponde a la cifra reportada por la Coordinación General de Estudio de Costos/DNPPE/SIyCE/ME para julio de 2018 dividida por 1,248 para tener en cuenta la inflación de 2017.



Tabla 3: Estadísticas descriptivas coordinadora y educadora

	N	Media	DE	P-10	P-90
Coordinadora					
Educación terciaria o superior [†]	93	0,74	0,44	0,00	1,00
Título en primera infancia	94	0,48	0,50	0,00	1,00
Años de educación	94	14,15	2,75	10,00	17,00
Salario mensual líquido (pesos)	90	13092,82	9115,79	5000,00	25681,50
Aportes a seguridad social	94	0,67	0,47	0,00	1,00
Vacaciones remuneradas	94	0,70	0,46	0,00	1,00
Empleo: con nombramiento	94	0,46	0,50	0,00	1,00
Empleo: por contrato	94	0,15	0,36	0,00	1,00
Empleo: voluntariado	94	0,09	0,28	0,00	0,00
Empleo: bonificaciones	94	0,29	0,45	0,00	1,00
Educadora					
Educación terciaria o superior [†]	95	0,80	0,40	0,00	1,00
Título en primera infancia	95	0,38	0,49	0,00	1,00
Años de educación	95	13,89	2,19	10,00	16,00
Salario mensual líquido (pesos)	93	8568,39	5216,39	3800,00	13500,00
Aportes a seguridad social	95	0,39	0,49	0,00	1,00
Vacaciones remuneradas	95	0,48	0,50	0,00	1,00
Ha trabajado con niños (0-60) más de 5 años	95	0,47	0,50	0,00	1,00
Empleo: con nombramiento	95	0,20	0,40	0,00	1,00
Empleo: por contrato	95	0,23	0,42	0,00	1,00
Empleo: voluntariado	95	0,05	0,22	0,00	0,00
Empleo: bonificaciones	95	0,45	0,50	0,00	1,00

[†] Incluye terciario incompleto.

Nota: Primera Infancia: Pedagogía, Desarrollo Infantil, Educación, Puericultura o Educación Pre-escolar.



Variables de calidad estructural: características físicas del EPI

El acceso a un teléfono es el único servicio básico en el que los EPI no tienen una cobertura del 100 %. En cuanto a la división física del EPI, si bien el 100 % tiene cocina, el 72 % tiene sala múltiple y el 86 % tiene inodoros adaptados para niños (61 % en buen estado), solo el 26 % de los jardines que ofrecen jornada completa y, por lo tanto, que deberían proveer almuerzo, tienen comedor. Adicionalmente, dado que la normativa de estos espacios es atender a niños desde los 45 días, sorprende que solo el 42 % tenga sala de bebés. El bajo porcentaje puede deberse a los requerimientos adicionales de estas salas, en términos de personal y espacio, que suponen un mayor costo por niño para los centros (UNICEF, 2019). Cabe resaltar la baja accesibilidad para personas con discapacidad (24 %) a pesar del alto auto reporte (84 %) de las coordinadoras afirmando que el EPI cuenta con estrategias de inclusión. Con respecto a la disponibilidad de algunos materiales, se observa que 49 % tiene juegos infantiles en buen estado, el 74 % tiene mesas y sillas suficientes acordes a la edad de los niños, que a su vez se consideran seguras en un 89 % de los casos. Por el lado de los elementos que ofrecen una infraestructura segura, las estadísticas no son muy alentadoras, pues solo el 3 % tiene un detector de gas, el 14 % tiene detector de humo, el 29 % ha hecho simulacros de evacuación y el 48 % tiene señales de evacuación. Adicionalmente, solo el 69 % tiene puertas o cerramiento y el 78 % tiene juegos infantiles de motricidad gruesa seguros (Tabla 4 y Tabla 5).

Los problemas más frecuentes que tienen los espacios físicos de los EPI son tomacorrientes sin protección (68 %), interruptores superpuestos (35 %) y la falta de buena ventilación natural (56 %). Adicionalmente, solo el 61 % de los artefactos usados para cocinar está en buen estado (Tabla 5). A pesar de que el 80 % de los EPI reporta que hay mantenimiento de las áreas externas, únicamente en el 63 % no hay acumulación de escombros y en el 54 % no se empoza el agua. En lo que se refiere a riesgos inminentes dentro del EPI, la Tabla 5 muestra que solo un 11 % tiene la garrafa o tubo fuera del centro, en el 69 % no hay presencia de material líquido inflamable y un 75 % no tiene objetos colgados mal asegurados.

Tabla 4: Estadísticas descriptivas calidad estructural (1)

	N	Porcentaje
Servicios básicos		
EPI tiene teléfono	95	92 %
División física del EPI		
EPI con acceso vehicular por todo el año	95	78 %
Accesibilidad para personas con discapacidad	95	24 %
EPI tiene sala múltiple	95	72 %
EPI tiene cocina	95	100 %
EPI tiene comedor*	31	26 %
EPI tiene Sala de 2 años	95	98 %
EPI tiene Sala de bebés	95	42 %
EPI tiene cerramiento	95	54 %
EPI tiene escaleras	95	43 %
Juegos infantiles en buen estado	89	49 %
Los juegos infantiles son seguros	89	78 %

*Cifra para EPI jornada completa.



Tabla 5: Estadísticas descriptivas calidad estructural (2)

	N	Porcentaje
Inodoros en buen estado	95	61 %
Inodoros adaptados para niños	95	86 %
Mesas y sillas suficientes	95	74 %
Mesas y sillas seguras	95	89 %
Sala múltiple con uso para juegos y aprendizaje	68	90 %
Existe facilidad de movilidad	95	73 %
Seguridad		
Detector de gas	93	3 %
Detector de humo	95	14 %
El EPI ha hecho simulacros de evacuación	95	29 %
Señales de evacuación	94	48 %
Escaleras seguras para niños	50	52 %
Salidas de emergencia	95	61 %
Cerramiento en buen estado	66	55 %
Tiene puertas o cerramiento	95	69 %
Los juegos infantiles son seguros	89	78 %
Botiquín de primeros auxilios	95	96 %
Vulnerabilidades		
Interruptores empotrados	95	93 %
Interruptores superpuestos	95	35 %
Toma corriente dentro de la pared	95	83 %
Toma corriente con protección	95	32 %
Buena iluminación natural	95	44 %
Buena ventilación natural	95	44 %
Tiene buena ventilación con ventilador	95	43 %
Buena calefacción	95	46 %
Buen aire acondicionado y calefacción	95	26 %
Químicos en un lugar seguro	23	91 %
Cocina y manejo de alimentos		
Heladera en buen estado	93	81 %
Cocina (artefacto) en buen estado	95	61 %
Alimentos NO están en contacto con el piso	95	76 %
Los niños NO tienen acceso a la cocina	95	77 %
Durante la visita: desayuno	95	64 %
Durante la visita: almuerzo	95	81 %
Durante la visita: merienda	95	55 %
Mantenimiento externo		
No hay acumulación de escombros	95	63 %
El agua no se empoza en áreas externas	95	54 %
Hay mantenimiento de áreas externas	95	80 %
Riesgos		
Garrafa o tubo fuera del jardín	95	11 %
No hay material lúdico inflamable	95	69 %
No hay objetos colgados mal asegurados	95	75 %
Los niños NO tienen libre acceso a pozo/tanque	69	90 %
Muebles aéreos asegurados	92	84 %
Ventanas NO rotas	95	88 %
No hay objetos cortantes	95	92 %
No hay animales	95	93 %
No hay cintas/cordones con los que puedan asfixiarse	95	95 %
No hay equipos electrónicos en áreas húmedas	95	97 %



4.2. Caracterización de la calidad de los EPI

CLASS

La Tabla 6 presenta la distribución del instrumento CLASS por dominio y dimensión. Se puede observar para el dominio de *Apoyo emocional y conductual* que el puntaje promedio es de 5,12 de un total de 7 puntos, y prácticamente la totalidad de la distribución se encuentra en el nivel medio/medio alto de puntajes posibles para este instrumento (Figura 5). Es interesante notar que el percentil 90 de los EPI en *Clima positivo*, *Clima negativo* y *Orientación de la conducta* alcanzan el nivel de excelencia. Por su parte, el puntaje promedio para el dominio de *Apoyo comprometido con el aprendizaje* es del nivel bajo/medio bajo, 2,48 de un total de 7 puntos, y casi la totalidad de su distribución se encuentra centrada en estos niveles (ver Figura 6). Es solo en *Modelaje lingüístico* que el percentil 90 de los EPI alcanzan el nivel medio. Es importante agregar que no se encuentran diferencias significativas en los puntajes CLASS entre EPI de jornada simple o EPI de jornada completa (no reportado). Así, los resultados CLASS sugieren que las educadoras de los EPI a veces brindan interacciones y actividades a los niños que estimulan el desarrollo de sus habilidades cognitivas y promueven su uso de lenguaje.

Tabla 6: Estadísticas descriptivas CLASS

	N	Media	DE	P-10	P-90
Apoyo emocional y conductual					
Total	55	5,12	0,51	4,35	5,75
Clima positivo	55	5,40	0,75	4,25	6,25
Clima negativo	55	6,73	0,40	6,00	7,00
Sensibilidad del educador	55	4,72	0,75	3,75	5,50
Consideración por la perspectiva del niño	55	3,07	0,56	2,25	3,75
Orientación de la conducta	55	5,65	0,54	5,00	6,25
Apoyo comprometido con el aprendizaje					
Total	55	2,48	0,67	1,44	3,25
Facilitación del aprendizaje	55	2,45	0,61	1,75	3,50
Calidad de la retroalimentación	55	2,06	0,63	1,25	2,75
Modelaje lingüístico	55	2,94	1,21	1,25	4,50

Nota: El puntaje CLASS va de uno a siete.

En términos comparativos, los hallazgos son bastante interesantes. Por ejemplo, en otro estudio muy parecido a este en el que se aplicó el CLASS en una muestra de centros públicos de cuidado infantil (Centros Infantiles de Buen Vivir) en Ecuador a través de videos, el puntaje promedio para el *Apoyo emocional y conductual* fue de 3,7, mientras para el *Apoyo comprometido con el aprendizaje* fue bajo, de 1,8; así los niños en los EPI experimentan un cuidado más sensible, donde se contiene más a los niños emocionalmente y se les ayuda a regular sus deseos, sentimientos y comportamientos. Son resultados no disimilares a los de un estudio de guarderías públicas de los Estados Unidos (Mashburn, 2008; Clifford y Bryant, 2003).

Desagregando el CLASS en sus respectivas dimensiones, se observa un puntaje de 5,12 en *Clima positivo* del dominio *Apoyo emocional y conductual*, lo cual refleja que los niños experimentan momentos de calidez en las interacciones con su educadora; en algunas ocasiones sonríen y se ríen con los niños o utilizan frases como “Te quiero” y “¡Me encanta ser tu educador!” Además, por lo



Figura 5: CLASS: Apoyo emocional y conductual

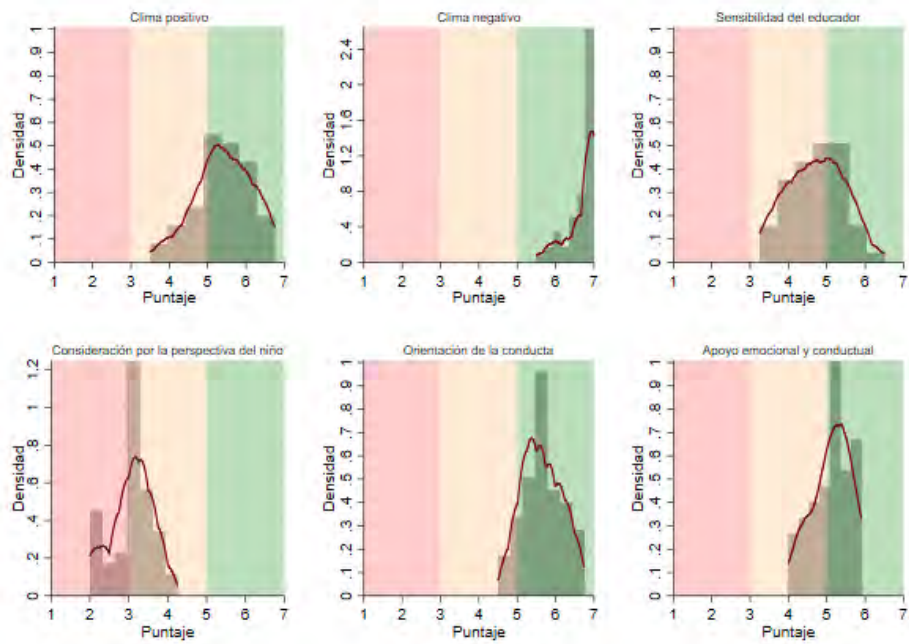
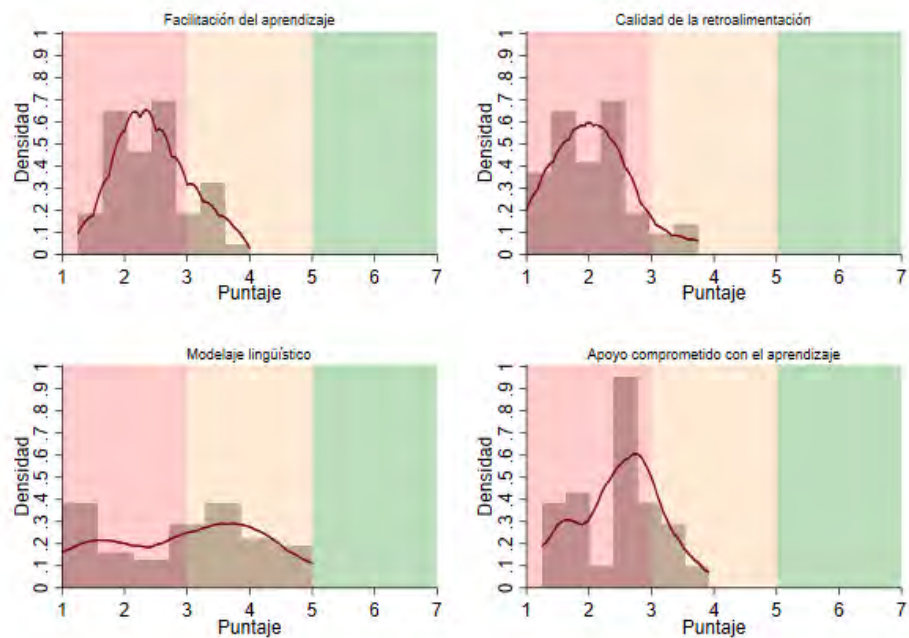


Figura 6: CLASS: Apoyo comprometido con el aprendizaje





general se observa un ambiente de respeto, con evidencia de contacto visual entre la educadora y los niños y en su orientación física hacia ellos. La calificación media alta de 4,72 en la dimensión *Sensibilidad del educador* es evidencia de que las educadoras están pendientes de la mayoría de los niños en el grupo durante gran parte de la jornada; por ejemplo, en algunas ocasiones les ofrecen ayuda, confort y juguetes, y los supervisan de cerca. En la dimensión de *Consideración por la perspectiva del niño*, que recibe el puntaje más bajo del dominio, se observa que los niños se encuentran a veces restringidos en su libertad de moverse, conversar y agarrar juguetes; además, no se observa suficiente evidencia del apoyo al autocuidado e independencia de los niños; por ejemplo, en el momento de guardar los juguetes, la educadora pierde la oportunidad de apoyar la autonomía y responsabilidad de los niños al hacerlo ella. En *Apoyo comprometido con el aprendizaje* se observa poca evidencia de interacciones eficaces; por ejemplo, la calificación de *Facilitación del aprendizaje* de 2,45 significa que los educadores durante gran parte de la observación están supervisando el grupo, en vez de guiar el aprendizaje y el desarrollo de los niños durante las actividades o rutinas con su participación activa, preguntas y comentarios. En *Calidad de la retroalimentación*, que recibió un 2,06, se observa en muy pocas ocasiones que los educadores provean andamiaje verbal o físico para que un niño logre hacer algo que no puede hacer por sí solo, o alienten la metacognición para ayudar a los niños a pensar sobre sus procesos de aprendizaje. Lo anterior quiere decir que, por ejemplo, la educadora no da seguimiento a las respuestas y acciones de los niños o no fomenta o reafirma los esfuerzos de los niños.

En conclusión, algunos de los resultados CLASS son alentadores, más que todo los del dominio *Apoyo emocional y conductual*, porque sugieren que muchos niños reciben una atención cálida y sensible en un porcentaje importante de los EPI. Sin embargo, dadas las altas brechas en desarrollo cognitivo que existen en todos los países de la región entre los niños de los estratos socioeconómicos más bajos y aquellos de estratos más altos documentadas en la literatura, y que se amplían precisamente en la primera infancia (Berlinski y Schady, 2015), los resultados – especialmente los de *Apoyo comprometido con el aprendizaje* - sugieren áreas de mejora de cara al futuro.

ITERS-R

El ITERS está diseñado para ser aplicado con un criterio de gradualidad según el cual, cuando el observador detecta que un aula no cumple con un conjunto de indicadores consecutivos, le otorga al ítem el puntaje correspondiente al nivel del indicador más alto alcanzado y ya no le administra los siguientes sino que pasa al siguiente ítem. En otras palabras, el protocolo tradicional del ITERS no requiere evaluar los ítems que se encuentran por encima de este punto de corte (Harnes, Cryer y Clifford, 2006). No obstante, dado que el ITERS no había sido aplicado antes en Argentina, en este estudio se decidió aplicar todos los ítems, más allá del punto de corte que dictan los lineamientos del instrumento, con el objeto de probar en qué medida los supuestos de gradualidad que están implícitos en el diseño de la escala son apropiados para la realidad de los centros argentinos. En la Tabla 7 se encuentran los puntajes ITERS de los centros, según el protocolo tradicional.

Los promedios de las subescalas ITERS están en el nivel mínimo/mínimo bueno con *Escuchar y hablar* en 3,03 e *Interacción* en 3,94. Es importante agregar que no se encuentran diferencias en los puntajes ITERS entre EPI de jornada simple o EPI de jornada completa. En *Escuchar y hablar*, una dimensión clave para el grupo etario menor a los 3 años, el ítem de *Ayuda a los niños a entender el lenguaje* se encuentra en un 3,82 y el ítem *Ayuda a los niños a utilizar el lenguaje* en un 3,70, que es el nivel medio-bajo (“mínimo”) de calidad según la escala ITERS. En términos prácticos, este hallazgo significa que solo en un tercio de los EPI observados había evidencia de educadores que utilizaron “palabras sencillas y descriptivas para los objetos y acciones cuando se comunica con los niños”. Además, se observaron muy pocas conversaciones bidireccionales con niños. Se encontró que



Tabla 7: Estadísticas descriptivas ITERS

	N	Media	DE	P-10	P-90
Escuchar y hablar					
Total	86	3,03	0,81	2,00	4,00
Ayuda a los niños a entender el lenguaje	94	3,82	1,22	2,00	5,00
Ayuda a los niños a utilizar el lenguaje	94	3,70	1,24	3,00	5,00
Uso de libros	88	1,63	0,94	1,00	3,00
Interacción					
Total	90	3,94	1,08	2,38	5,25
Supervisión del juego y aprendizaje	94	4,11	1,34	2,00	6,00
Interacción entre los niños	95	4,20	1,13	3,00	5,00
Interacción entre el personal y niños	95	3,83	1,55	2,00	6,00
Disciplina	91	3,53	1,42	1,00	5,00
Total ITERS					
Total	82	3,52	0,91	2,14	4,71

Nota: El puntaje ITERS va de uno a siete.

los educadores en cierta medida hablan con los niños, y por lo general, de manera agradable; a veces, ellos responden a las vocalizaciones o intentos no verbales de los niños a comunicarse, pero la mayoría no hace intentos explícitos para estimular su desarrollo lingüístico. El último ítem de la subescala, *Uso de libros*, está en el rango más bajo de la distribución con 1,63, evidencia de una ausencia generalizada de libros y actividades de lectura con los niños. Específicamente, no hay una variedad de libros apropiados y accesibles a niños y, en los EPI que sí cuentan con libros, los educadores no los utilizan con los niños. En resumen, los resultados sugieren que los EPI están realizando algunas prácticas que favorecen el desarrollo de los niños, pero aún enfrentan un desafío en términos de mejorar los procesos que específicamente contribuyen al desarrollo de las habilidades de lengua y pre-lectura de los niños a quienes atienden. La Figura 7 muestra la distribución de los puntajes y, al igual que el promedio, indica que los puntajes en los ítems de *Ayuda a los niños a entender el lenguaje* y *Ayuda a los niños a utilizar el lenguaje* son substancialmente mayores que los puntajes en *Uso de libros*.

El promedio de la segunda subescala que aplicamos, *Interacción*, es de 3,94. El desglose de los ítems se encuentra en la Tabla 7 y la distribución de los puntajes en la Figura 8. Este resultado implica que, en promedio, los educadores se mostraban atentos a los niños y a sus responsabilidades de cuidado; por ejemplo, dejaban que los niños interactuaran entre sí durante una gran parte del día; mostraban una cierta medida de afecto físico y verbal durante el día; a veces, estructuraban el día y las actividades para evitar conflictos entre niños; y, a veces, prestaban atención a los niños que se portaban bien (en vez de solo llamar la atención a los que se portaban de una manera inapropiada). Se debe destacar que, si bien los puntajes promedio de esta subescala y los ítems que la componen indican un nivel de calidad mínimo, un porcentaje importante de las distribuciones de los primeros tres ítems -*Supervisión del juego y aprendizaje* (43,62%), *Interacción entre los niños* (45,26%) e *Interacción entre el personal y niños* (41,05%)- tienen un puntaje de 5 o mayor (nivel de calidad bueno o superior).²⁷

²⁷ Además, comprobamos que no se cumple el criterio de gradualidad del ITERS, en particular, para los ítems *Interacción entre el personal y niños* y *Ayuda a los niños a entender el lenguaje*, el puntaje alternativo proporciona información adicional. Se observa que para estos dos ítems el puntaje según el protocolo alternativo es más alto que el



Figura 7: ITERS: Escuchar y hablar

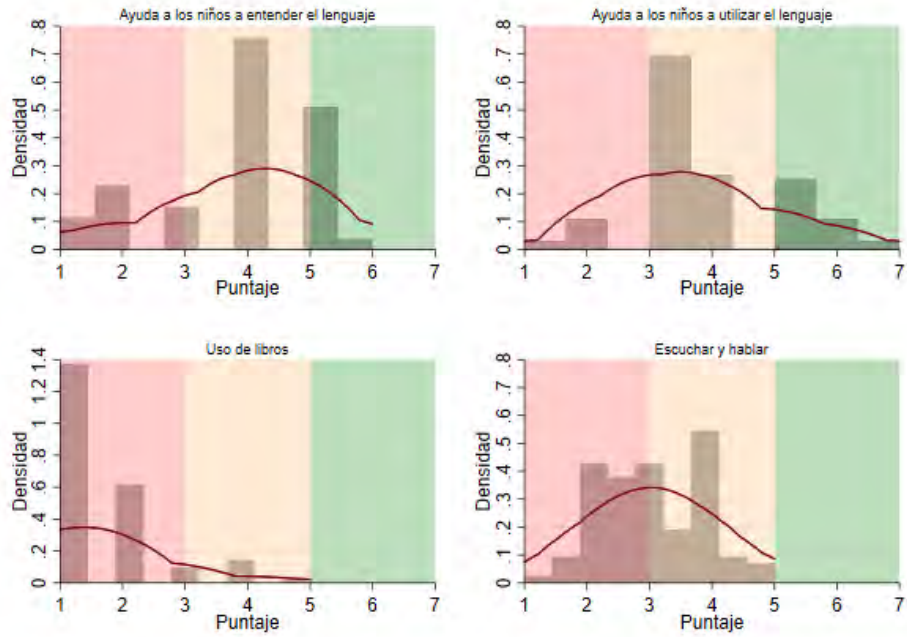
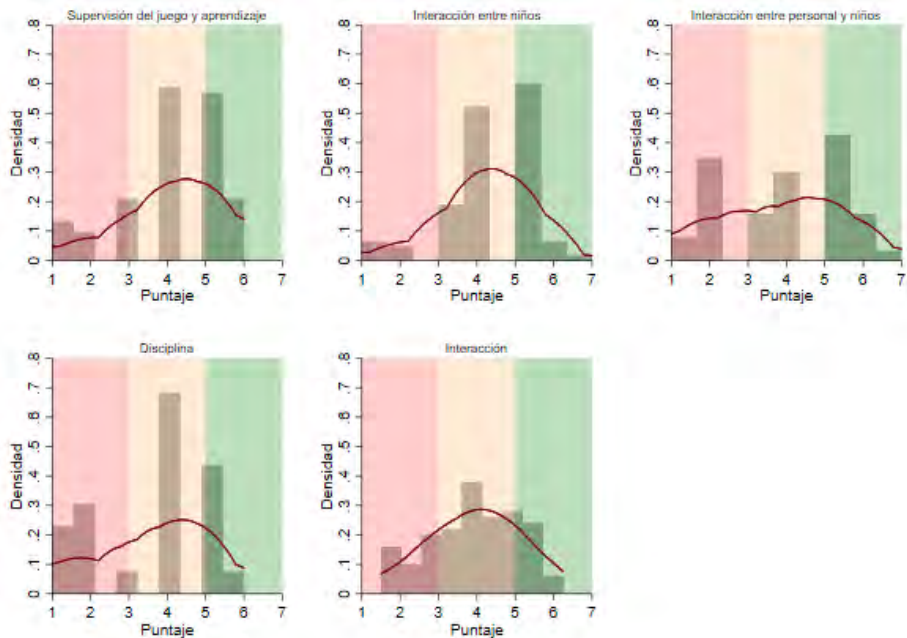


Figura 8: ITERS: Interacción





La Tabla 8 resume los resultados de los instrumentos CLASS e ITERS para diferentes programas de la región. Se observa que los puntajes obtenidos en todas las subescalas en ambos instrumentos por los EPI del presente estudio son claramente superiores a los de muestras comparables en Bolivia y Ecuador. Con relación a los centros en Perú, los puntajes también tienden a ser más altos en el CLASS, lo que no sucede en las subescalas del ITERS. Respecto al CLASS, resalta la magnitud de la diferencia (alrededor de 1,5 puntos) en el dominio *Apoyo emocional y conductual*, y en particular en *Clima positivo* y *Orientación de la conducta*; así como en *Modelaje lingüístico* que presenta puntajes de casi el doble de los encontrados en Perú y Ecuador. En *Clima negativo* y *Facilitación del aprendizaje y del desarrollo*, por el contrario, prácticamente no se observan diferencias entre los países. Con respecto a la primera dimensión, son todos ambientes cálidos en los que no se observa presencia de manejos de la disciplina negativa, lo cual es algo positivo en la región. Con relación a la segunda dimensión, los puntajes son bajos, lo que sugeriría que es un área de mejora en América Latina.

Tabla 8: Otros estudios: ITERS y CLASS

	Bolivia, CBPVB	Ecuador, CIBV	Perú, CM	Argentina, EPIs
ITERS				
Escuchar y hablar	1,3	1,5	3,3	3,0
Interacción	1,4	3,3	5,0	3,9
Observaciones	100	404	602	95
CLASS				
Total	-	2,6	3,1	4,1
<i>Apoyo emocional y conductual</i>	-	3,6	3,9	5,1
Clima positivo	-	3,3	3,4	5,4
Clima negativo	-	6,6	6,9	6,7
Sensibilidad del educador	-	3,4	3,3	4,7
Consideración por la perspectiva del niño	-	2,0	3,1	3,1
Orientación de la conducta	-	2,9	3,0	5,7
<i>Apoyo comprometido con el aprendizaje</i>	-	1,6	1,8	2,5
Facilitación del aprendizaje y del desarrollo	-	2,1	2,5	2,5
Calidad de la retroalimentación	-	1,3	1,3	2,1
Modelaje lingüístico	-	1,6	1,5	2,9
Observaciones	100	404	602	55

Fuente: Los datos de Bolivia y Perú provienen de Berlinski y Schady (2015). Los datos de Ecuador de Araujo, López Boo, et al. (2015). Datos para Argentina, cálculos propios.

Lista de Verificación

Los resultados de la LV en la Tabla 9 ofrecen una imagen bastante favorable de las prácticas en los EPI. Por ejemplo, el puntaje total muestra que en más de tres cuartos de los EPI observados, los educadores inician intercambios verbales y físicos positivos con los niños y responden a los pedidos de atención verbal y no-verbal de los niños de manera positiva y apropiada; además, los educadores participan de manera activa y positiva en las responsabilidades del cuidado y las actividades con los niños. Algunas áreas por mejorar son las prácticas de salud para reducir la propagación de gérmenes (como lavarse las manos adecuadamente cuando es necesario para proteger la salud) y el acceso a

tradicional. Esto sugiere que, en estos casos se observan interacciones de mayor calidad pero no se están cumpliendo todos los requisitos previos para alcanzar un puntaje más alto según protocolo tradicional ITERS. En definitiva, son áreas donde, con relación al puntaje tradicional obtenido, se necesitarían menos esfuerzos por parte de la educadora para alcanzar un nivel de calidad mayor.



suficientes materiales y libros para el juego libre. Los resultados de LV son, en su mayoría, más positivos que los del ITERS o el CLASS. Una posible explicación de la diferencia puede ser que mientras la LV se centra en la presencia o ausencia (0 o 1) de ciertas interacciones y elementos clave (funcionando como un tamizaje), los otros instrumentos tienen un mayor enfoque en los niveles de calidad – es decir, la frecuencia, intensidad, eficacia y complejidad – de los mismos aspectos de calidad, lo que se ve reflejado en que tengan puntajes continuos (1 a 7).

Tabla 9: Estadísticas descriptivas Lista de Verificación (% de buenas prácticas)

	N	Media	DE	P-10	P-90
Receptividad (8 ítems)	95	0,83	0,16	0,63	1,00
Organización (1 ítem)	95	0,91	0,29	1,00	1,00
Materiales de aprendizaje (5 ítems)	95	0,70	0,28	0,40	1,00
Involucramiento (5 ítems)	95	0,61	0,32	0,20	1,00
Variedad (2 ítems)	95	0,45	0,41	0,00	1,00
Rutinas de cuidado personal (3 ítems)	95	0,95	0,17	0,67	1,00
Siesta (3 ítems)	74	0,86	0,29	0,33	1,00
Prácticas de salud (3 ítems)	95	0,44	0,33	0,00	1,00
Escuchar y hablar (7 ítems)	95	0,74	0,22	0,43	1,00
Actividades (3 ítems)	95	0,57	0,32	0,33	1,00
Interacción (18 ítems)	95	0,78	0,21	0,39	1,00
Estructura del programa (8 ítems)	95	0,73	0,23	0,38	1,00
Promedio					
Total - LV*	95	0,73	0,15	0,54	0,90

* Se excluyen los ítems de Siesta.

Nota: La interpretación de cada línea es la siguiente. Si en Receptividad la media es 0,83, esto significa que en promedio un EPI cuenta con buenas prácticas en el 83% de los ítems de Receptividad y habría que trabajar en el restante 27% de los ítems. Ver estadísticas por cada uno de los 66 ítems en la Tabla C.2, Apéndice C.

En resumen, los instrumentos de calidad en sala y sus respectivas subdimensiones aplicados para este estudio revelan niveles de calidad de tendencia media o medio-baja en los EPI. Los instrumentos que más comúnmente se usan en la literatura para evaluar la calidad de este tipo de servicios, el CLASS y el ITERS, presentan puntajes que principalmente se concentran en la parte media de la distribución de calidad; un punto importante a notar es que los dos instrumentos parecen registrar niveles más altos de calidad en prácticas de cuidado cálido y sensible, y niveles más bajos en prácticas de estimulación de los procesos de pensamiento, lenguaje y pre-lectura. Esto sigue los patrones encontrados para estos instrumentos a nivel internacional (Slot, 2018; Lemay, Lehrer y Naud, 2017).

4.3. Pruebas de Desarrollo Infantil

Bayley III

Como el Bayley no se aplicó a todos los niños que tienen la PRUNAPE, realizamos pruebas de diferencias de medias en características de los niños y del EPI, entre la muestra a la que se le aplicó Bayley y la muestra a la que no (ver Tabla C.3). Únicamente se encontró una diferencia



estadísticamente significativa (de acuerdo con el reporte de si el niño nació prematuro o no), lo que sugiere que no existen diferencias entre las dos muestras.

En la Tabla 10 se puede observar que el puntaje promedio de nuestra muestra es de 33,73 (DE=4,36). En un estudio de los centros AEIOTU en Colombia, el promedio de las 3 cohortes con niños de 1 a 4 años (edad promedio 28 meses) fue de 23,78 en la misma subescala de lenguaje receptivo (NIEER, 2012). Por otro lado, un estudio reciente de Bogotá (Rubio, Araujo, et al., 2016) presenta una media levemente superior 34,35 y DE=4,04. Ambos estudios de Colombia son representativos de niños de ingresos bajos y medios, como este estudio. En definitiva, el desarrollo de lenguaje de los niños del GBA es similar al de niños colombianos semejantes. Si tomamos la población de referencia de los Estados Unidos (lo cual es debatible, pero sirve para poner en contexto los puntajes), se encuentra que los niños en los EPI incluidos en este estudio se encuentran en la misma curva que los norteamericanos. Es decir, el 17% de niños de esta muestra se encuentra con un desvío estándar por debajo de la norma internacional, mientras que la distribución (normal) de niños de referencia muestra un 15,8% niños por debajo de un desvío estándar (puntaje no incluido en la tabla).

Tabla 10: Estadísticas descriptivas Bayley III- lenguaje receptivo

	N	Media	DE	P-10	P-90
Hombre	302	0,45	0,50	0,00	1,00
Edad (meses)	302	35,06	5,48	29,00	42,00
Prematuro menor a 2 años	302	0,00	0,06	0,00	0,00
Puntaje bruto	302	33,73	4,36	28,00	39,00
Puntaje estandarizado por edad	302	-0,02	0,97	-1,18	1,16

Nota 1: Muestra: 44 EPI. No hay errores en la aplicación de ninguna de las pruebas.

Nota 2: La información de prematuro proviene de la prueba PRUNAPE.

Nota 3: Si bien las madres fueron invitadas para la administración de los tests, en el 98% de los casos fue la educadora quien acompañó al niño durante la aplicación (en el 2% la madre).

La Figura 9 muestra las diferencias en el puntaje Bayley de acuerdo con dos medidas de condiciones sociodemográficas del partido en el que se encuentra el EPI al que asiste el niño: i) el porcentaje de niños de 0 a 4 en hogares con necesidades básicas insatisfechas (NBI), y ii) el nivel de urbanización. Para determinar si las distribuciones del puntaje en Bayley entre las submuestras (niños en EPI localizados en partidos con porcentaje de niños de 0 a 4 en hogares con NBI (i) mayor a 24%, y (ii) menor a 24%, y niños en EPI localizados en partidos (i) completamente urbanizados, y (ii) parcialmente urbanizados) son estadísticamente diferentes, se utiliza la prueba de Kolmogorov de igualdad de distribuciones. De acuerdo con esta prueba, en ninguno de los dos casos, existe evidencia estadística para afirmar que las distribuciones son diferentes.

PRUNAPE

En este estudio nos planteamos dos objetivos con la administración de la prueba, identificar si el niño presenta un retraso probable en su desarrollo, usando la prueba PRUNAPE tal cual fue diseñada; y construir un puntaje continuo que nos permitiera generar una distribución y comparar al niño con sus pares. En el Apéndice D se describe en detalle la prueba y su aplicación.

En la Tabla 11, los datos de PRUNAPE muestran un porcentaje de rezago de 21 puntos por-



centuales (pp) más que en el Estudio de Nutrición, Desarrollo Psicomotor y Análisis Toxicológico (ENDUPAT) en la cuenca Matanza-Riachuelo en 2014, donde 51,0 % de los niños no pasaba la PRUNAPE; y, como era de esperarse por el nivel socioeconómico, de 32 pp más que en el estudio de San Isidro (porcentaje de no pase del 40,2 %) (Lejarraga et al., 2014).²⁸ Al revisar por dimensión se encuentra que Lenguaje es el área con mayor rezago, con un porcentaje de 46 % de niños que no pasan la prueba. Las áreas con menor rezago son Motricidad fina y Motricidad gruesa con porcentajes de no pasa de 32 % y 37 % respectivamente.

Tabla 11: Estadísticas descriptivas PRUNAPE

	N	Media	DE	P-10	P-90
Edad (meses)	930	33,76	6,52	26,00	42,00
Prematuro	781	0,01	0,11	0,00	0,00
Personal-social: no pasa	859	0,41	0,49	0,00	1,00
Motricidad fina: no pasa	859	0,32	0,47	0,00	1,00
Lenguaje: no pasa	859	0,46	0,50	0,00	1,00
Motricidad gruesa: no pasa	859	0,37	0,48	0,00	1,00
Puntaje personal-social	683	14,65	2,58	12,00	17,00
Puntaje motricidad fina	683	12,48	1,70	11,00	15,00
Puntaje lenguaje	683	10,93	3,56	6,00	15,00
Puntaje motricidad gruesa	683	18,40	2,06	16,00	21,00
Puntaje estandarizado personal-social	683	0,05	0,85	-0,91	0,93
Puntaje estandarizado motricidad fina	683	0,04	0,99	-1,17	1,26
Puntaje estandarizado lenguaje	683	0,02	1,00	-1,34	1,08
Puntaje estandarizado motricidad gruesa	683	0,01	0,98	-1,05	1,17
Total PRUNAPE: no pasa*	859	0,72	0,45	0,00	1,00
Número de áreas con no pase	859	1,56	1,18	0,00	3,00

* No se incluye la subescala de motricidad gruesa. Ver Apéndice D para justificación.

Nota 1: Número de pruebas inválidas no incluidas 71.

Nota 2: Si bien las madres fueron invitadas para la administración de los tests, en el 98 % de los casos fue la educadora quien acompañó al niño durante la aplicación (en el 2 % la madre).

Nota 3: No pasa en el área X es igual al: porcentaje de niños que fracasan por lo menos en una pauta tipo A o en dos pautas tipo B del área X .

Nota 4: Los niños que no pasan la prueba son sospechosos o están en riesgo de padecer un retraso en el desarrollo.

En la Figura 10 se presentan las diferencias en el puntaje para cada una de las áreas de la PRUNAPE de acuerdo con dos medidas de condiciones sociodemográficas del partido en el que se encuentra el EPI al que asiste el niño: i) el porcentaje de niños de 0 a 4 años en hogares con NBI, y ii) el nivel de urbanización. De nuevo se utiliza la prueba de Kolmogorov de igualdad de distribuciones para determinar si las distribuciones del puntaje en la PRUNAPE entre las dos submuestras y para cada una de las medidas son estadísticamente diferentes. Únicamente se encuentra que las distribuciones son diferentes en el área de motricidad fina por nivel de NBI (a un nivel de confianza del 90 %). Contrariamente a lo esperado, los niños en partidos con un menor porcentaje de NBI

²⁸Las condiciones de evaluación (distractores en el EPI, falta de la madre/educadora principal acompañando al niño, exceso de rapidez de encuestadores en la aplicación del PRUNAPE) son posibles factores que explican el mayor número de fallas (ver Apéndice D).



obtuvieron puntajes más bajos en esta área. Se debe notar que no se encontró diferencia alguna en motricidad fina al utilizar nivel de urbanización.

Adicionalmente, para ambas medidas de desarrollo se revisó si existían diferencias en el desarrollo del niño por tipo de jornada, completa – simple (resultados no presentados). Se encontraron diferencias estadísticamente significativas. En particular, el puntaje en la prueba Bayley y el puntaje en todas las áreas de la PRUNAPE son en promedio mayores en los EPI que operan en jornada simple.

En resumen, los instrumentos de desarrollo infantil muestran que: (i) el nivel de lenguaje receptivo de los niños que asisten a los EPI es semejante al de regiones similares, tales como Bogotá; (ii) el porcentaje de niños que asisten a estos EPI y que no pasan la PRUNAPE es muy alto comparado con el baremo nacional (el Apéndice D discute algunas razones metodológicas que pueden estar detrás de estas altas tasas); y iii) la evidencia sugiere que, en general, no existen diferencias en el desarrollo de los niños de acuerdo con condiciones sociodemográficas del partido en el que se encuentra ubicado el EPI. Esto último puede deberse a que los EPI siguen un criterio de focalización por lo que es probable que las características de los niños atendidos y sus hogares sean similares independientemente de la ubicación del centro. Es importante recalcar que sólo se encuentran representados niños en EPI y no en centros privados, cuya matrícula aumenta en las localidades de mayores ingresos.



Figura 9: Distribución de puntajes de la subescala de lenguaje receptivo del Bayley por condiciones sociodemográficas del partido

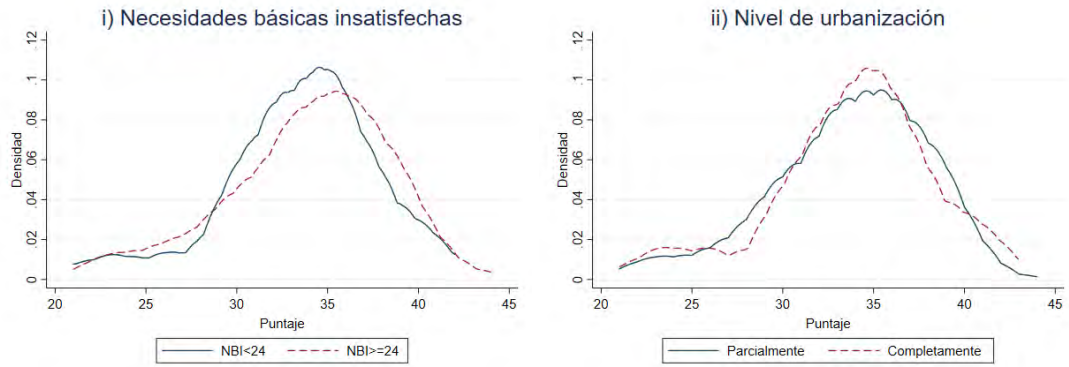
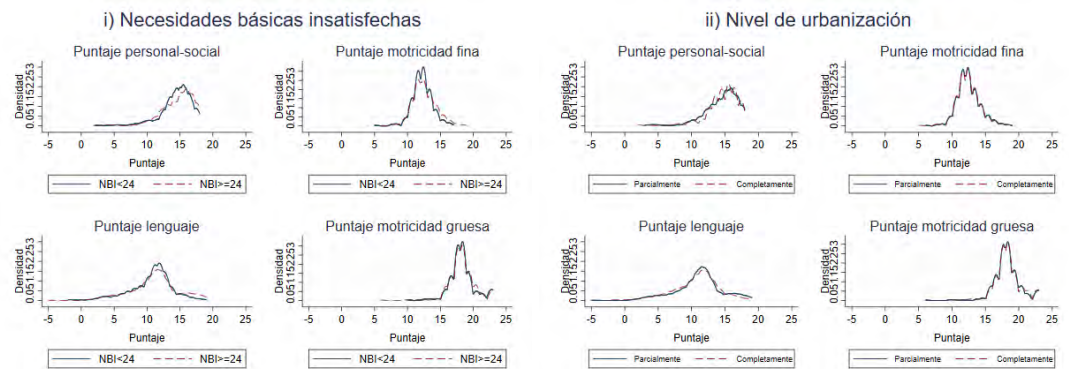


Figura 10: Distribución de puntajes de las áreas de la PRUNAPE por condiciones sociodemográficas del partido





4.4. Un instrumento costo-efectivo para monitorear la calidad de procesos en el aula

Luego de verificar algunas propiedades psicométricas de la Lista de Verificación, tales como la confiabilidad entre observadores, la consistencia interna, la validez concurrente y sus asociaciones con varias variables relevantes (ver Apéndice E), en esta sección analizamos en detalle cada uno de los ítems que componen la Lista con el objetivo de seleccionar un sub-conjunto de ellos que puedan constituir un instrumento simplificado y costo-efectivo para ser usado de manera frecuente para la supervisión de los EPI a gran escala. Idealmente, este instrumento debería: a) focalizarse en los aspectos más informativos de la calidad de procesos; y b) ser de utilidad para los mecanismos de supervisión de los EPI, en particular, ser útil para que los supervisores puedan estructurar su retroalimentación a las educadoras y los procesos internos de SENAF/MDS en cuanto a la mejora continua de los servicios de primera infancia. Mientras CLASS y ITERS son instrumentos que capturan varios aspectos de la calidad de procesos, también son complejos, consumen 3 horas de observación por sala y es costoso su entrenamiento y administración frecuente a gran escala.

La sección 3.2 describió brevemente como se obtuvieron los 12 constructos y 66 ítems de la LV. Para asegurar una buena calidad de la observación y los datos, y por restricciones de tipo operativo (recursos, tiempo de capacitaciones, perfil de supervisores), se decidió acortar la LV a un número de ítems menor que permitiera una capacitación de 1 día y no más de 2 horas de observación. Para seleccionar entonces los ítems que harían parte de la lista corta se aplicaron los siguientes criterios:

1. El ítem no podía tener más del 10% de observaciones faltantes. Este criterio se impuso para descartar ítems que no son relevantes en todas las situaciones de cuidado infantil.
2. El ítem debía tener suficiente variabilidad, en particular el 10% de las observaciones deberían ser diferentes de 0 ó de 1. Este criterio se impuso pues ítems con poca variabilidad no aportan información y no permiten captar diferencias entre los EPI.
3. Para factores con valor propio superior a 1, la carga factorial del ítem debía ser igual o mayor a 0,5. El criterio de valor propio mayor a 1, o regla de Kaiser-Guttman (1954), asegura que la varianza explicada por el factor es mayor que la varianza explicada por una única de las variables observadas. Adicionalmente, un ítem con carga factorial baja sugiere que no está significativamente relacionado con el factor a medir y, por lo tanto, no debería ser incluido. Si bien la literatura especializada utiliza a menudo un umbral de 0,4, en este trabajo se decidió usar un umbral mayor para reducir la cantidad de ítems resultantes por restricciones de tipo operativo.

Una vez recolectada la información en nuestra muestra, se identificó que 6 de estos ítems no cumplían con el primer criterio, y que 3 ítems presentaban casi la mitad de las observaciones sin información²⁹, es decir no cumplían el segundo criterio. Así es que estos 9 ítems se descartaron de la lista. Con los 57 ítems restantes se computó el valor propio de cada factor (los ítems se asignan al constructo preestablecido teóricamente) y la carga factorial de cada ítem. Se excluyó de este ejercicio el constructo *Organización* pues al estar conformado por un único ítem estos estadísticos no se pueden computar (ver resultados en las Tablas G.1 y G.2). Una vez se aplica el tercer criterio, se reduce el instrumento a 32 ítems sobre los cuales se conducen el resto de los análisis (a esta lista se le llamará “LV corta”). Los ítems incluidos en esta lista se presentan en la Tabla 12.³⁰

²⁹Corresponden, por ejemplo, al caso particular de ítems relacionados con la siesta, situación que por la dinámica de los EPI no fue posible observar.

³⁰Esta lista resulta en 9 subescalas (dado que los ítems de *Siesta*, *Variedad* y *Actividades* quedaron completamente eliminados luego del ejercicio descripto. Nótese que los ítems que presentan la mayor cantidad de correlaciones positivas y significativas con CLASS e ITERS son también los que hacen parte del instrumento final (referirse a la Tabla E.6). La lista de ítems descartados y el criterio utilizado para excluirlos se presenta en la Tabla G.3.



Tabla 12: LV corta: formato e ítems incluidos

Panel I : Información sobre la visita al aula, el niño y la educadora	
Visita	
Fecha de la visita	
Duración	
Nombre EPI	
Nombre del responsable	
Pertence a una red	
Dirección	
Partido	
Teléfono	
Educadoras	
Número de educadoras presentes	
Educación de la educadora	
Niños	
Número de niños presentes	
Rango de edades de niños presentes	
Panel II : Listado para monitoreo de calidad de procesos a escala (LV, 32 ítems)	
	Puntuación
Receptividad	
1. Educador/a habla o vocaliza espontáneamente al niño al menos dos veces	[0, 1]
7. Educador/a habla en un tono de voz que transmite sentimientos positivos hacia el niño	[0, 1]
Organización	
9. El espacio donde juega el niño es seguro	[0, 1]
Materiales de aprendizaje	
11. Carrito de juguete, coche de supermercado, coche de bebés, andador, patineta o triciclo disponible	[0, 1]
14. Juguetes de coordinación compleja entre manos y ojos	[0, 1]
Involucramiento	
15. Educador/a conscientemente promueve el progreso en el desarrollo del niño	[0, 1]
16. Educador/a se enfoca en juegos apropiados para la edad priorizando la atención personalizada	[0, 1]
17. Educador/a estructura los períodos de juego del niño	[0, 1]
18. Educador/a provee juegos que desafían al niño a desarrollar habilidades nuevas	[0, 1]
Rutinas de cuidado personal	
22. Se saluda a la mayoría de los niños afectuosamente	[0, 1]
23. El personal saluda a cada niño y a su madre o padre y dirige una salida agradable y organizada	[0, 1]
24. Los problemas con la separación de la madre o del padre, o la salida del jardín, se tratan con sensibilidad	[0, 1]
Prácticas de salud	
28. Por lo general, el personal actúa para reducir la propagación de gérmenes	[0, 1]
29. El personal y los niños se lavan las manos adecuadamente el 75 % de las veces cuando es necesario para proteger la salud	[0, 1]
Escuchar y hablar	
32. El personal habla con los niños frecuentemente a lo largo del día durante las rutinas y los juegos	[0, 1]
36. El personal agrega palabras a las acciones que llevan a cabo al responder a los niños durante el día	[0, 1]
37. El personal es hábil en cuanto interpretar los intentos de comunicación de los niños y hace seguimiento apropiadamente	[0, 1]
Interacción	
43. El personal se muestra atento a todo el grupo incluso mientras trabaja con un niño o un grupo pequeño	[0, 1]
44. El personal reacciona rápidamente para solucionar los problemas de una manera reconfortante y fortalecedora	[0, 1]
45. El personal juega con los niños y muestra interés y conocimiento por lo que hacen	[0, 1]
46. El personal ayuda y anima a los niños cuando es necesario	[0, 1]
48. El personal facilita las interacciones positivas entre niños	[0, 1]
49. El personal modela la interacción social positiva	[0, 1]
52. Interacción positiva frecuente entre el personal y los niños durante el día	[0, 1]
53. El personal y los niños se muestran normalmente relajados, hablan agradablemente y sonríen con frecuencia	[0, 1]
55. El personal por lo general mantiene suficiente control para evitar problemas	[0, 1]
56. El personal usa métodos de disciplina positiva de manera efectiva	[0, 1]
57. Se presta atención frecuentemente cuando los niños se portan bien	[0, 1]
Estructura del Programa	
59. El personal organiza actividades de juego como parte del horario diario	[0, 1]
63. El personal participa activamente para facilitar el juego de los niños durante el día	[0, 1]
64. El personal por lo general es positivo y tolerante con los niños durante el juego en grupo	[0, 1]
65. El personal es flexible y adapta la actividad a medida que los niños se unen o dejan el grupo	[0, 1]



Validación de la Lista de Verificación corta

La consistencia interna: En la Tabla 13 se presentan los resultados de la consistencia interna y la validez de constructo para los ítems definitivos. En particular, en la primera columna se presenta el Alfa de Cronbach de cada constructo y de los ítems. El alfa del ítem corresponde al alfa del constructo si no se incluyera ese ítem, por lo que lo deseable es que sea menor que el alfa del constructo. En la segunda columna se presenta el coeficiente de correlación de Pearson del ítem con el puntaje del constructo. Finalmente, en las últimas dos columnas se muestra el valor propio del constructo y la carga factorial de cada ítem, respectivamente. Como se observa, en la mitad de los casos el alfa del constructo es mayor a 0,7, valor que se considera un nivel aceptable de coherencia interna. *Materiales de Aprendizaje* es la única escala con un alfa bajo, inferior a 0,5, por lo que se debe tener cuidado al utilizarla. Se debe notar además que este constructo se enfoca en la disponibilidad de materiales y no en las interacciones educador-niño. Para todos los ítems menos uno, el Alfa de Cronbach disminuiría si se excluyera la pregunta del constructo, lo que indica altos niveles de coherencia interna entre los ítems de la escala. Por último, todos los ítems tienen una carga factorial alta, de por lo menos 0,4, lo que significa que contribuyen a la medición de la misma.

La validez concurrente -nuestra muestra y datos de Ecuador-: Para verificar la validez concurrente se revisaron las correlaciones entre la LV corta y otros instrumentos que miden la calidad, en particular ITERS y CLASS. Los resultados de este ejercicio se presentan en la Tabla 14. Se observa que *Involucramiento*, *Escuchar y hablar*, *Interacción* y *Estructura del programa*, así como el puntaje total, están positiva y significativamente correlacionados con las subescalas y puntajes totales del ITERS (alto) y del CLASS (moderado-bajo), mientras que *Receptividad*, *Organización*, *Rutinas de cuidado personal* y *Prácticas de salud* solamente están correlacionadas significativamente con el ITERS, y estas correlaciones son bajas. En contraste, el constructo *Materiales de aprendizaje* no está correlacionado ni con el ITERS ni con el CLASS, lo cual es esperable dado que está más enfocado en variables estructurales que en procesos. En conjunto, estos resultados sugieren una alta validez de la LV corta.

Para fortalecer la conclusión anterior, se hizo un análisis de umbrales de calidad. En particular se revisó si las correlaciones entre los constructos de la LV corta y el ITERS y CLASS se mantenían para aulas con puntajes de calidad medios a altos. Los resultados de este ejercicio sugieren que la LV corta logra discriminar aulas con diferentes niveles de calidad, aun en el rango alto de la distribución. En particular, encontramos que las correlaciones estadísticamente significativas entre los constructos de la LV corta y la subescala de *Interacción* del ITERS se mantienen. No obstante, la significancia estadística de las correlaciones con la dimensión de *Apoyo emocional y conductual* del CLASS desaparecen para aulas con niveles de calidad medio alto o superior, principalmente por falta de precisión, por lo que es importante repetir el análisis en otros estudios similares para tener evidencia más sólida. En el Apéndice G se presenta una descripción detallada del ejercicio y los resultados.

Como un ejercicio de validación adicional se replicó la Tabla 14 utilizando los datos secundarios del estudio representativo de centros públicos en Ecuador. En particular, se volvieron a estimar las correlaciones entre los constructos que componen la Lista de Verificación corta y los puntajes obtenidos en el ITERS y en el CLASS (ver Tabla G.4). Las magnitudes de las correlaciones en Ecuador son similares a las obtenidas para el Gran Buenos Aires y los rankings se mantienen, es decir, los dominios del CLASS y las subescalas del ITERS que se asocian más o menos fuerte con los constructos de la LV corta en el Gran Buenos Aires son los mismos que en el estudio de Ecuador. Adicionalmente, aun cuando los puntajes obtenidos en el estudio de Ecuador son más bajos y presentan menor varianza, la significancia de las correlaciones se mantiene e incluso tiende a ser mayor. Cabe aclarar que en Ecuador no se administró la LV corta como tal, sino que se recuperaron



Tabla 13: Consistencia interna y validez de constructos de la LV corta

	Alfa	Correlación	Valor propio	Carga
Receptividad	0,59		0,60	
P1.		0,85***		0,55
P7.		0,83***		0,55
Materiales de aprendizaje	0,47		0,40	
P11.		0,82***		0,45
P14.		0,80***		0,45
Involucramiento	0,75		1,60	
P15.	0,70	0,74***		0,61
P16.	0,65	0,80***		0,69
P17.	0,68	0,76***		0,65
P18.	0,72	0,72***		0,57
Rutinas de cuidado personal	0,67		1,13	
P22.	0,49	0,79***		0,69
P23.	0,65	0,79***		0,55
P24.	0,60	0,78***		0,60
Prácticas de salud	0,78		1,05	
P28		0,92***		0,72
P29		0,89***		0,72
Escuchar y hablar	0,69		1,22	
P32	0,76	0,68***		0,43
P36	0,52	0,82***		0,70
P37	0,46	0,85***		0,73
Interacción	0,87		4,25	
P43	0,86	0,60***		0,59
P44	0,85	0,73***		0,72
P45	0,86	0,65***		0,60
P46	0,86	0,59***		0,52
P48	0,85	0,71***		0,67
P49	0,86	0,60***		0,57
P52	0,85	0,73***		0,70
P53	0,85	0,72***		0,70
P55	0,86	0,56***		0,53
P56	0,86	0,64***		0,58
P57	0,85	0,68***		0,62
Estructura del programa	0,72		1,61	
P59	0,62	0,77***		0,74
P63	0,69	0,75***		0,54
P64	0,70	0,67***		0,55
P65	0,63	0,79***		0,69

Nota. Si el constructo no tiene al menos 3 ítems, no se puede calcular el Alfa de Cronbach del ítem.



Tabla 14: Correlación bivariada entre constructos de la LV corta y CLASS e ITERS

	1.Dominio: Apoyo emocional	2.Dominio: Apoyo comprometido	3.Total CLASS	4.Subescala: Escuchar y hablar	5.Subescala: Interacción	6.Total ITERS
1. Receptividad (2 ítems)	0,14	0,01	0,09	0,16	0,25**	0,24**
2. Organización (1 ítem)	0,23*	0,14	0,21	0,21**	0,22**	0,24**
3. Materiales de aprendizaje (2 ítems)	0,03	0,05	0,04	-0,01	0,10	0,07
4. Involucramiento (4 ítems)	0,35***	0,30**	0,35***	0,58***	0,51***	0,59***
5. Rutinas de cuidado personal (3 ítems)	-0,06	-0,16	-0,11	0,13	0,27***	0,24**
6. Prácticas de salud (2 ítems)	0,19	0,23*	0,22	0,24**	0,24**	0,26**
7. Escuchar y hablar (3 ítems)	0,32**	0,27**	0,32**	0,73***	0,63***	0,73***
8. Interacción (11 ítems)	0,35***	0,47***	0,43***	0,62***	0,88***	0,86***
9. Estructura del programa (4 ítems)	0,28**	0,32**	0,32**	0,53***	0,65***	0,67***
10. Total LV corta (32 ítems)	0,41***	0,44***	0,45***	0,71***	0,86***	0,88***

Nota: La muestra de las columnas 1-3 son los 55 EPI donde se levantó CLASS, para las demás son 95 EPI. * p<0.10, ** p<0.05, *** p<0.01.

los ítems de la aplicación del ITERS y del CC-IT-HOME. En resumen, estos resultados solidifican la validez concurrente de la Lista de Verificación corta.

Correlaciones con otras variables relevantes: Se revisaron las correlaciones entre los constructos de la Lista de Verificación corta y diferentes variables estructurales del EPI, incluyendo características de la coordinadora y la educadora a cargo de la sala observada, características físicas del EPI y su funcionamiento. La literatura ha encontrado que las variables estructurales están significativamente correlacionadas con la calidad de proceso al facilitar las interacciones de alta calidad (Howes, Phillips y Whitebook, 1992; Clarke-Stewart, Gruber y Fitzgerald, 1994; NICDH, 2000; Kagan, 2010). Sin embargo, por sí solas no garantizan un buen cuidado (Hamre y Pianta, 2007). Al revisar las correlaciones entre esta batería de variables estructurales y el CLASS, ITERS y LV (Tabla E.7) se encontraron pocas correlaciones significativas, por lo que tampoco esperábamos que la LV corta estuviera correlacionada con todas estas variables estructurales. Los resultados del ejercicio para la LV corta se presentan en la Tabla 15.

El *Involucramiento* y la *Estructura del programa* son constructos que están positiva y significativamente correlacionados con características que dan cuenta de las condiciones laborales de la coordinadora y la educadora, incluyendo el salario mensual, realizar aportes a seguridad social y vacaciones remuneradas. Contrariamente a lo que se esperaría, la *Receptividad* está negativamente correlacionada con el nivel educativo de la educadora y las variables que dan cuenta de las condiciones laborales. No obstante, la significancia de estas correlaciones desaparece o se invierte para



el puntaje total. Por su parte, el número de niños en el aula se correlaciona negativamente con *Organización*, *Materiales de aprendizaje* e *Involucramiento*. Este último hallazgo está en línea con lo que indica la literatura especializada.

Con relación al funcionamiento del EPI dos hallazgos llaman la atención. Por un lado, el número de meses al año en que ofrece atención el EPI está positivamente correlacionado con 5 de los 9 constructos y el total. Por otro lado, el pago promedio mensual que deben realizar las familias está negativamente correlacionado con el *Involucramiento*, la *Interacción*, *Escuchar y hablar* y el total, lo cual puede explicarse si los EPI que “piden” pago mensual a las familias son los que cuentan con menos recursos disponibles y, por ende, les resulta más difícil promover buenas prácticas. Variables relacionadas con las estrategias de inclusión, los servicios a padres y la disponibilidad de información sobre peso y talla no se correlacionan con ninguno de los factores, tal vez porque en los EPI hay un porcentaje muy bajo de niños con discapacidad o minorías y los servicios a padres son básicos (variables no reportadas en la tabla). También se debe tener en cuenta que el objetivo *per se* de verificar peso y talla o incorporar estrategias de inclusión no es afectar la calidad en el aula. Estas permiten referir al niño a servicios especializados de ser necesario y están asociadas a normativas relacionadas con los derechos del niño, lo que ayudaría a explicar la ausencia de correlaciones estadísticamente significativas. En términos generales se ve que la Lista de Verificación corta presenta correlaciones dentro de los parámetros esperados.

Correlación entre la LV corta y el desarrollo del niño: Por último, se revisaron las correlaciones entre la LV corta y el desarrollo del niño. En la literatura se han reportado de manera consistente relaciones entre indicadores de la calidad de proceso y el desarrollo de los niños (La Paro, Pianta y Stuhlman, 2004; Hamre y Pianta, 2007). Luego, esperábamos que CLASS, ITERS y LV estuvieran significativamente asociados con los instrumentos de medición del desarrollo infantil. Los resultados se presentan en la Tabla 16.

El total y los constructos de la LV que estaban asociados con alguna de las medidas de desarrollo de los niños usando los 66 ítems (Tabla E.9), siguen estándolo al restringir a los 32 ítems de la LV corta (con excepción de la correlación marginalmente significativa entre el puntaje total de la LV y la subescala de motricidad fina de la PRUNAPE). Resaltan los constructos de: *Organización* (4 asociaciones positivas incluyendo Bayley), *Prácticas de salud* (1 asociación positiva con Motricidad fina de PRUNAPE), *Escuchar y hablar* (3 asociaciones positivas incluyendo Bayley), *Interacción* (2 asociaciones positivas incluyendo Bayley), *Estructura del programa* (2 asociaciones positivas con Bayley y Lenguaje de PRUNAPE), que tienen asociaciones muy consistentes con varios de los dominios del desarrollo (ver Tabla 16). Asimismo, se encuentran solo dos asociaciones negativas: una entre *Rutinas de cuidado personal* y el total de PRUNAPE y la otra entre *Prácticas de salud* y Bayley. Ambos dominios de la LV consisten de pocos ítems (3) y los constructos son de naturaleza más estructural, con lo que estos resultados deberían ser analizados en mayor profundidad en investigaciones futuras. La magnitud de estas relaciones tiende a aumentar en la LV corta.

Tabla 15: Correlación bivariada entre constructos de la LV corta y variables estructurales del EPI

	1. Receptividad	2. Organización	3. Materiales de aprendizaje	4. Involucramiento	5. Rutinas de cuidado personal	6. Prácticas de salud	7. Escuchar y hablar	8. Interacción	9. Estructura del programa	10. Total LV corta
Coordinadora										
1. Educación terciaria o superior	0,10	0,20**	0,17	0,03	0,05	-0,01	-0,02	-0,06	0,07	0,02
2. Título en DTI	-0,04	0,09	0,07	0,10	-0,01	0,02	0,11	0,01	0,06	0,07
3. Salario mensual líquido (pesos)	-0,07	0,13	0,15	0,36***	0,09	0,17	0,18*	0,04	0,26**	0,23**
4. Aportes a seguridad social	-0,10	0,23**	-0,02	0,12	0,02	0,13	0,11	0,16	0,27***	0,19*
5. Vacaciones remuneradas	-0,08	0,10	0,08	0,10	-0,05	0,14	0,09	0,17	0,23**	0,18*
6. Empleo: con nombramiento	-0,05	0,08	0,03	0,10	0,02	0,15	0,09	0,06	0,10	0,11
Educadora										
7. Educación terciaria o superior	-0,19*	0,11	-0,12	0,07	-0,09	0,00	0,02	-0,14	-0,04	-0,09
8. Título en DTI	-0,12	-0,04	-0,01	0,25**	0,10	0,13	0,15	0,02	0,09	0,12
9. Salario mensual líquido (pesos)	-0,23**	0,02	0,06	0,22**	0,04	0,18*	0,16	0,00	0,18*	0,13
10. Aportes a seguridad social	-0,20*	0,04	-0,01	0,18*	0,15	0,12	0,04	0,03	0,18*	0,11
11. Vacaciones remuneradas	-0,21**	-0,05	0,09	0,20*	0,07	-0,02	0,11	0,15	0,24**	0,17*
12. Empleo: con nombramiento	-0,14	0,07	0,01	0,04	0,09	0,01	0,00	0,00	0,01	0,01
Estructura										
13. Niños/ educadores (EPI)	0,12	-0,05	-0,14	-0,10	-0,14	0,07	0,08	0,05	-0,05	0,00
14. Número de niños en el aula (Sala de 2)	-0,15	-0,25**	-0,23**	-0,24**	-0,02	-0,15	0,05	-0,05	0,00	-0,14
15. Índice estructural (sin imputaciones)	-0,26**	0,08	-0,01	0,02	-0,03	0,14	0,03	0,00	0,05	0,02
16. Índice estructural†	-0,28**	0,01	-0,00	0,10	-0,18**	0,10	0,13	0,08	0,09	0,08
Funcionamiento										
17. Número de años de funcionamiento	0,20*	-0,18*	-0,02	0,13	-0,02	-0,01	0,23**	0,13	0,05	0,14
18. Meses al año que ofrece atención	0,25**	0,47***	0,04	0,20*	0,36***	-0,12	0,14	0,24**	0,17	0,27***
19. Días de vacaciones	0,01	0,15	0,03	0,18	0,15	0,09	0,19*	0,16	0,29**	0,24**
20. Familias deben realizar algún pago	0,05	0,07	0,02	-0,06	0,01	-0,02	-0,14	0,00	0,01	-0,03
21. Pago promedio mensual (pesos)	0,07	0,13	0,12	-0,22**	0,04	-0,05	-0,23**	-0,18*	-0,10	-0,18*

† Como control se incluye una variable dicotómica por cada variable utilizada en la construcción del índice que indica si para esa variable se realizó o no una imputación. Nota: Número de observaciones 95 (puede variar ligeramente por valores omitidos). Correlaciones con Estrategias de Inclusión (ver variables en la Tabla 2) no muestran significancia estadística. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



Tabla 16: Asociación entre la Lista de Verificación corta y el desarrollo del niño

	Bayley						PRUNAPE					
	Lenguaje Receptivo		Personal Social		Motricidad Fina		Lenguaje		Motricidad Gruesa		Total	
Receptividad (2 ítems)	0,87	(0,52)	0,10	(0,12)	-0,16	(0,12)	-0,21	(0,17)	-0,04	(0,17)	-0,10	(0,10)
Organización (1 ítem)	0,85***	(0,24)	0,16	(0,10)	0,10	(0,09)	0,40***	(0,15)	0,61***	(0,16)	0,41***	(0,12)
Materiales de Aprendizaje (2 ítems)	0,29	(0,29)	-0,09	(0,10)	0,01	(0,11)	0,04	(0,15)	0,01	(0,17)	-0,01	(0,12)
Involucramiento (4 ítems)	0,25	(0,46)	0,16	(0,10)	0,02	(0,10)	-0,03	(0,15)	-0,08	(0,19)	0,03	(0,13)
Rutinas de Cuidado Personal (3 ítems)	-0,78	(0,52)	-0,11	(0,22)	-0,21	(0,23)	-0,33	(0,31)	-0,44	(0,28)	-0,34**	(0,16)
Prácticas de Salud (2 ítems)	-0,37**	(0,18)	0,08	(0,10)	0,25***	(0,10)	-0,14	(0,10)	-0,12	(0,13)	-0,01	(0,08)
Escuchar y Hablar (3 ítems)	0,40*	(0,24)	0,19*	(0,10)	0,25*	(0,13)	0,05	(0,14)	0,16	(0,14)	0,18	(0,12)
Interacción (11 ítems)	0,55*	(0,29)	0,29**	(0,14)	0,18	(0,19)	-0,09	(0,20)	-0,10	(0,23)	0,07	(0,16)
Estructura del Programa (4 ítems)	2,05***	(0,75)	0,22	(0,18)	0,09	(0,17)	0,51***	(0,19)	-0,13	(0,36)	0,28	(0,22)
Total IV corta (32 ítems)	0,83*	(0,49)	0,41**	(0,18)	0,26	(0,26)	-0,02	(0,28)	-0,07	(0,30)	0,17	(0,21)
N	292		657		657		657		657		657	

Nota: Errores estandar en paréntesis (agrupados a nivel EPI). Todas las regresiones incluyen efectos fijos de referente y partido. * p<0.10, ** p<0.05, *** p<0.01.



Junto con la LV corta, existen otras iniciativas que buscan desarrollar herramientas de observación del aula para países de ingresos bajos y medios. Resaltamos el “Instrumento para la evaluación de la calidad de los servicios para niños de cero a tres años” (IECS-03)(Ortiz, et al., 2018) y el “Inventario de caracterización de ambientes para educación inicial” (INCA-EI) (Vásquez, et al., 2018), instrumentos adaptados al contexto uruguayo que evalúan la calidad del servicio de la educación inicial. El primero fue diseñado para aulas con niños entre 0 a 3 años y el segundo para aulas con niños entre 3 a 5 años. Ambos instrumentos miden aspectos de la calidad de proceso, pero a diferencia de la LV corta, esta dimensión no es el foco de la medición. Además, son más extensos y exigen un tiempo de observación mayor. Otra iniciativa semejante a la LV corta por i) su objetivo de uso, se espera se utilice en los sistemas de diagnóstico y monitoreo, al igual que en los planes de desarrollo profesional; ii) su énfasis en aspectos de la calidad de procesos; y iii) ser escalable, instrumento de fácil aplicación, escrito de manera concisa y simple, incluye materiales complementarios con ejemplos específicos, es abierto y no requiere de observadores con niveles muy altos de educación ni experiencia para su aplicación, es el *Teach classroom observational tool* (Molina, et al., 2018). La principal diferencia entre el *Teach* y la LV corta es que el primero fue desarrollado para evaluar aulas de primaria, mientras que el segundo para evaluar aulas que atienden niños menores de 3 años.

5. Conclusiones

El principal hallazgo respecto a la calidad global en la sala es que, si bien los EPI estudiados tienen márgenes de mejora en aspectos estructurales y de procesos, los centros muestran algunas fortalezas, en particular en el dominio de *Apoyo emocional y conductual* (CLASS) donde prácticamente la totalidad de la distribución se encuentra en el nivel medio/medio alto de puntajes posibles. En comparación con programas de otros países de la región, el modelaje del lenguaje aparece mejor puntuado, aunque es mejorable respecto de lo esperado en estos ambientes. Puntajes altos en la dimensión *Clima positivo* de CLASS reflejan que los niños experimentan muchos momentos de calidez en las interacciones con sus educadores. El dominio de *Apoyo comprometido con el aprendizaje* de CLASS, por el contrario, se encuentra en el nivel bajo/medio bajo sugiriendo que las educadoras de los EPI brindan pocas interacciones y actividades a los niños que apoyen el desarrollo de sus habilidades cognitivas y promuevan su uso de lenguaje. Los resultados anteriores sugieren que el fortalecimiento de planes de capacitación y supervisión de recursos humanos debería priorizar temas que apunten a mejores interacciones, particularmente en lo que se refiere al desarrollo cognitivo y al lenguaje infantil.

Las dos subescalas del ITERS se encuentran en el nivel medio/medio bajo (aunque se debe notar que una proporción importante de EPI obtuvieron puntajes medio altos o superiores en tres de los ítems de la subescala de *Interacción*): solo en un tercio de los EPI observados hubo evidencia de educadoras que utilizaron palabras sencillas y descriptivas para los objetos y acciones cuando se comunicaban con los niños y se observaron muy pocas conversaciones bidireccionales con niños. Si bien las educadoras hablan de forma agradable, pocas veces responden a las vocalizaciones o intentos no-verbales de los niños de comunicarse y, en general, no se observan intentos explícitos para estimular su desarrollo lingüístico. Estas son competencias más complejas que requieren una capacitación, un desarrollo profesional y hasta una modelación más específica. A su vez, el puntaje obtenido en el ítem de *Uso de libros* (1,6 puntos sobre 7) evidencia una ausencia de libros apropiados y actividades de lectura con los niños pequeños. En definitiva, es prioritario que los gobiernos locales (municipales o provinciales) y nacionales, construyan competencias para maximizar el bienestar y promover el aprendizaje en todas las áreas del desarrollo.

Las educadoras observadas en este estudio cuentan en promedio con 14 años de educación y están



a cargo de un promedio de 15 niños por sala. Estos niños tienen en promedio 2 años y medio y la mitad sufre rezagos en su desarrollo, como ya ha sido documentado en otros estudios en población vulnerable del GBA. En cuanto a infraestructura y materiales, se manifiesta que hay necesidades básicas de infraestructura que, aun cuando no definen el desarrollo del niño, sí afectan su bienestar y salud.

Por otra parte, todos los instrumentos presentaron excelentes propiedades psicométricas en su primera aplicación en Argentina y la Lista de Verificación se correlacionó bien con los instrumentos internacionales más complejos administrados en paralelo y con el salario de la coordinadora del centro en particular.

La principal contribución de este estudio fue identificar empíricamente un subconjunto de 32 ítems que se centran en aspectos críticos de la calidad de procesos y que se pueden recopilar de forma rutinaria como parte de los esfuerzos de monitoreo de los servicios de cuidado infantil. Es importante destacar que este instrumento, no solo presenta excelentes propiedades psicométricas, sino que también logra discriminar aulas de diferentes niveles de calidad, aún en el rango alto de la distribución.

Los ítems de esta lista son muy específicos y usan un lenguaje simple y directo para que se puedan completar eficientemente a medida que se desarrolla la visita al EPI. Adicionalmente, la especificidad de los ítems es importante para garantizar que no haya ambigüedad en su interpretación y calificación. No obstante, para lograr mayor confiabilidad es fundamental la capacitación de los supervisores que lo administren. En resumen, las ventajas operativas principales de esta lista de ítems son: su corta duración en campo (dos horas); la posibilidad de recolectar datos con personal no profesional, siempre que la capacitación sea la adecuada; el costo razonable de la capacitación y de la administración frecuente, dada la simplicidad de la herramienta, y su solidez técnica ya mencionada. La limitación principal tiene que ver con el puntaje binario que se centra en la presencia o ausencia (0 a 1) de ciertas interacciones y elementos clave, mientras que instrumentos más complejos, caros y largos tienen un mayor enfoque en los niveles de calidad, es decir, en la frecuencia, intensidad, eficacia y complejidad de los mismos aspectos de calidad, lo que se ve reflejado en que tengan puntajes continuos (1 a 7).

Dentro de la lista, los ítems relacionados con los constructos de *Involucramiento*, *Escuchar y hablar*, *Interacción* y *Estructura del programa* parecen estar más altamente correlacionados con los aspectos de calidad identificados en CLASS e ITERS. Como una prueba adicional de validez, los puntajes de esta lista (en particular, aquellos en las escalas de *Organización*, *Interacción* y *Escuchar y hablar*) se correlacionan positivamente con la mayoría de las dimensiones de desarrollo de los niños estudiadas (Lenguaje receptivo del Bayley, y Personal-social, Motricidad fina y Lenguaje de PRUNAPE).

Además de todos estos elementos, la lista también incluye unos pocos elementos de calidad estructural en la visita al EPI, tales como: duración de la visita del supervisor, número de niños en la sala, rango de edad de niños en la sala, número de educadoras y características de las educadoras.

Mientras este reporte se publica, la Subsecretaría de Primera Infancia del Gobierno Argentino está planificando la capacitación de los supervisores territoriales de los EPI de todo el país en el uso de esta lista corta, que ha sido bautizada como MIRAr (por **M**onitoreo de **I**nteracciones **R**eceptivas en salas de **A**rgentina). Para ese ejercicio, un equipo local de especialistas en primera infancia adaptó mínimamente el lenguaje de algunos de los 32 ítems con el fin de reflejar el lenguaje y la cultura de distintas regiones del país (ver SENAF/BID, 2019). Este trabajo de campo a nivel nacional servirá como un primer pilotaje a escala que no requiera un perfil profesional para la aplicación. Serán necesarios más pilotajes de la aplicación de esta herramienta para refinarla y para determinar si lo más idóneo en un sistema de mejora continua es el supervisor regular o un agente externo. Aplicar pilotos con diferentes perfiles de aplicador, diferentes tipos de capacitación, diferentes frecuencias



(semestral versus anual) de aplicación, para no “cansar” al ojo del observador ni a la educadora, entre otros, permitirán mejoras y enriquecerán la literatura actual.

En términos de recomendaciones de políticas públicas, es importante destacar que aquellos programas que, además del monitoreo frecuente, quieran enfocarse en evaluar exhaustivamente su calidad de procesos, el uso del CLASS puede resultar atractivo dadas sus asociaciones altas con los constructos de interacciones educador-niño. Es deseable y existe la posibilidad de integrar el CLASS (o el ITERS) dentro de un sistema de monitoreo a escala, con algunos ajustes para reducir los costos de administración. Por ejemplo en Estados Unidos (*Head Start*, n.d.), el CLASS se aplica cada dos o tres años y se administra con menos ciclos de lo indicado en los protocolos originales (es decir, menos tiempo de observación). Para esto, contar con sistemas de carga de datos a nivel nacional/regional resulta indispensable. El ITERS también ha sido integrado en algunos prescolares en Nueva Jersey, Estados Unidos, para evaluar las prácticas en estos centros. Los “Maestros Máster”, que cumplen una función de guía y apoyo para los maestros, administran este instrumento como parte del sistema de monitoreo.

Sin embargo, en contextos con restricciones de recursos, la administración del ITERS y del CLASS puede ser de difícil implementación ya que implica contar con recursos humanos especializados para la aplicación del instrumento, procesos de capacitación extensos, periodos de observación largos, etc. En este sentido, este estudio provee a los programas de la región que operan en contextos con fuertes restricciones de recursos una base concreta para mejorar en el corto plazo el monitoreo de la calidad de proceso y lograr un sistema de mejora continua. En situaciones ideales, estos ítems deberían ser complementados con el uso de los instrumentos globales, aplicando el ITERS o el CLASS con una periodicidad menor y/o aplicándolos a una submuestra. La información conjunta de los instrumentos (listas de verificación para monitoreo y CLASS o ITERS) permitirá dar una dirección a largo plazo a las políticas de profesionalización del personal de primera infancia y a las políticas de apoyo a centros del sector. Será de extrema importancia, asimismo, el acompañamiento cercano y articulado de los municipios, de las provincias y de la nación en el desarrollo y regulación del funcionamiento de estos Espacios de Primera Infancia que, al integrarse al PNPI, reciben beneficios pero, también, aceptan obligaciones supervisadas por el MDS.



Apéndices A-F

A. Cálculos de poder

Tabla A.1: Efectos mínimos detectables

# grupos (EPI)	Número promedio de niños por EPI		
	5	10	15
CIC=0.1			
90	0,156	0,129	0,118
100	0,148	0,122	0,112
CIC=0.2			
90	0,177	0,156	0,149
100	0,168	0,148	0,141
CIC=0.3			
90	0,196	0,180	0,174
100	0,186	0,170	0,165



B. Características de las medidas de calidad utilizadas

Tabla B.1: Características principales de las medidas de calidad de cuidado infantil

	CLASS	ITERS-R	LV
Edad (meses)	15-36	0-30	0-36
Tipo de variable	Proceso	estructural/proceso	estructural/proceso
Numero de dimensiones/subescalas	8	7 (1 estructural)	12 (3 estructural)****
Numero de ítems	8	39	66 (LVC=32)
Método de evaluación	Observación (directa o video)	Observación directa y encuesta	Observación directa y encuesta
Capacitación oficial?	Si	No	No
Puntaje	1-7	1-7 *	0,1 (Si/no)
Tiempo mínimo de administración (horas)	2	3.5	2
Costo total (US\$) **	902,90	22,90	0
Adaptado para la región LAC?	Si	Si	Si
Traducción en español	Si	Si	Si
Validez ***	Si	Si	En proceso

* El ITERS está diseñado para ser aplicado con un criterio de gradualidad según el cual para cada Ítem, cuando el observador detecta que un aula no cumple con un conjunto de indicadores consecutivos, le otorga al Ítem el puntaje correspondiente al nivel del indicador más alto alcanzado y ya no le administra los siguientes sino que pasa al siguiente Ítem.

** Este costo corresponde al costo de los materiales por observador / codificador. CLASS es la única excepción donde el costo incluye el entrenador oficial obligatorio. Los precios eran válidos el 28 de diciembre de 2015.

*** Esta característica indica la existencia de estudios sobre la validez y confiabilidad del instrumento (López Boo et al., 2019).

**** *Siesta* se fue en el primer paso y *Variedad y Actividades* en el análisis de carga factorial resultando en 9 subescalas la LV corta.



C. Tablas Adicionales

Tabla C.1: Análisis factorial confirmatorio para constructos de interacción educador-niño

Instrumento	Constructo latente									
	Sensitividad/Receptividad	Estimulación de lenguaje y cognitiva	Apoyo a las interacciones entre pares	Consideración positiva/calidez	Afecto positivo	Reciprocidad	Mutualidad	Atención conjunta	Estilorientación de la regulación de la conducta	Actitud intrusiva
CLASS										
1. Clima Positivo	0.97		0.88	0.97	0.44	0.94		0.98		
2. Clima Negativo										0.41
3. Sensibilidad del educador	0.97			0.97				0.95		
4. Consideración por la perspectiva de		0.38		0.58		0.60	0.49			0.60
5. Orientación de la conducta									0.46	
6. Facilitación del aprendizaje		0.53					0.94	0.74	0.40	0.72
7. Calidad de la retroalimentación		0.48				0.71	0.81			
8. Modelaje lingüístico		0.53	0.80							
ITERS-R										
1. Espacio y muebles				0.37	0.58					
2. Rutinas de cuidado personal	0.38									
3. Escuchar y hablar	0.45	0.76								
4. Actividades		0.50				0.27		0.25		
5. Interacción	0.39			0.37	0.85					0.81
6. Estructura del programa		0.69								
7. Padres y personal							0.05			0.41
CC-IT-HOME										
1. Responsividad	0.36	0.65		0.34	0.75					0.76
2. Aceptación	0.05			0.04	0.25					
3. Organización		0.53								
4. Materiales de aprendizaje		0.62	0.32							
5. Involucramiento	0.28		0.43			0.27				0.62
6. Variedad		0.50								
Pruebas de bondad de ajuste										
Chi-cuadrada	808.30	860.86	145.03	438.01	24.50	66.33	39.01	25.10	345.00	
Grados de libertad	27.00	54.00	5.00	20.00	9.00	9.00	5.00	5.00	14.00	
RMSEA	0.27	0.19	0.26	0.23	0.07	0.13	0.13	0.10	0.24	
CFI	0.60	0.59	0.74	0.76	0.98	0.90	0.93	0.98	0.69	
CD	0.97	0.87	0.85	0.97	0.85	0.90	0.91	0.98	0.85	

Nota: cargas factoriales mayores a 0.3 en negrita, mayores a 0.7 en cursiva.

Fuente: Dormal y López Bóo (2018), con base en datos de Ecuador.



Tabla C.2: Estadísticas Descriptivas Lista de Verificación (% de buenas prácticas)

	N	Media	DE	P-10	P-90
1. Educador/a habla o vocaliza espontáneamente al niño al menos dos veces	95	0,91	0,29	1,00	1,00
2. Educador/a responde verbalmente a vocalizaciones o verbalizaciones del niño	95	0,92	0,28	1,00	1,00
3. Educador/a le dice o indica al niño el nombre de un objeto o persona durante la visita	95	0,75	0,44	0,00	1,00
4. Cuando el educador/a habla, lo que se dice se distingue con facilidad, es claro, se escucha bien	95	1,00	0,00	1,00	1,00
5. Educador/a conversa libremente y fácilmente	95	0,96	0,20	1,00	1,00
6. Educador/a espontáneamente elogia al niño al menos dos veces	95	0,46	0,50	0,00	1,00
7. Educador/a habla en un tono de voz que transmite sentimientos positivos hacia el niño	95	0,92	0,28	1,00	1,00
8. Educador/a acaricia o besa al niño al menos una vez	95	0,76	0,43	0,00	1,00
9. El espacio donde juega el niño es seguro	95	0,91	0,29	1,00	1,00
10. Juegos o equipos disponibles que permiten al niño ejercitar sus músculos	95	0,71	0,46	0,00	1,00
11. Carrito de juguete, coche de supermercado, coche de bebés, andador, patineta o triciclo disponible	95	0,52	0,50	0,00	1,00
12. Muñeco de peluche o muñeco suave, juegos para imitar roles disponibles	95	0,69	0,46	0,00	1,00
13. Equipo que facilita el aprendizaje -móviles, mesa y sillas pequeñas, silla de comer alta, corral- disponible	95	0,96	0,20	1,00	1,00
14. Juguetes de coordinación compleja entre manos y ojos	95	0,63	0,48	0,00	1,00
15. Educador/a conscientemente promueve el progreso en el desarrollo del niño	95	0,51	0,50	0,00	1,00
16. Educador/a se enfoca en juegos apropiados para la edad priorizando la atención personalizada	95	0,55	0,50	0,00	1,00
17. Educador/a estructura los períodos de juego del niño	95	0,67	0,47	0,00	1,00
18. Educador/a provee juegos que desafían al niño a desarrollar habilidades nuevas	95	0,43	0,50	0,00	1,00
19. Educador/a mantiene al niño dentro de su rango visual, y lo mira seguido	95	0,87	0,33	0,00	1,00
20. El educador/a lee cuentos con el niño al menos 3 veces por semana	95	0,49	0,50	0,00	1,00
21. El niño tiene acceso a 3 o más libros en el jardín	95	0,41	0,49	0,00	1,00
22. Se saluda a la mayoría de los niños afectuosamente	95	0,97	0,18	1,00	1,00
23. El personal saluda a cada niño y a su madre o padre y dirige una salida agradable y organizada	95	0,94	0,24	1,00	1,00
24. Los problemas con la separación de la madre o del padre, o la salida del jardín, se tratan con sensibilidad	95	0,95	0,22	1,00	1,00
25. Hay suficiente supervisión para los niños durante la siesta	46	0,80	0,40	0,00	1,00
26. Las cunas (o colchonetas o camas infantiles) son para dormir y no para jugar durante largos períodos de tiempo	47	0,81	0,40	0,00	1,00
27. La supervisión es agradable, atenta y cariñosa	73	0,88	0,33	0,00	1,00
28. Por lo general, el personal actúa para reducir la propagación de gérmenes	95	0,34	0,48	0,00	1,00
29. El personal y los niños se lavan las manos adecuadamente el 75% de las veces cuando es necesario para proteger la salud	95	0,22	0,42	0,00	1,00
30. Los niños son atendidos adecuadamente para satisfacer sus necesidades de salud adentro y afuera	95	0,76	0,43	0,00	1,00
31. El contenido de la conversación es generalmente alentador y positivo más que desalentador y negativo	95	0,93	0,26	1,00	1,00
32. El personal habla con los niños frecuentemente a lo largo del día durante las rutinas y los juegos	95	0,76	0,43	0,00	1,00
33. La comunicación verbal es personalizada	95	0,88	0,32	0,00	1,00
34. Hay una medida moderada de respuesta positiva verbal o no verbal a los intentos de los niños de comunicarse durante el día*	95	0,96	0,20	1,00	1,00
35. Hay algunos intentos de interpretar correctamente lo que los niños están intentando comunicar durante el día	95	0,92	0,28	1,00	1,00
36. El personal agrega palabras a las acciones que llevan a cabo al responder a los niños durante el día	95	0,34	0,48	0,00	1,00
37. El personal es hábil en cuanto interpretar los intentos de comunicación de los niños y hace seguimiento apropiadamente	95	0,39	0,49	0,00	1,00
38. Hay como mínimo un juego de bloques (6 o más bloques del mismo tipo) accesibles diariamente	95	0,65	0,48	0,00	1,00
39. Los materiales son accesibles diariamente, durante gran parte del día	95	0,33	0,47	0,00	1,00
40. Los materiales se encuentran organizados por tipo	95	0,73	0,45	0,00	1,00
41. El personal puede ver, oír y alcanzar a los niños fácilmente, con no más de unos pocos descuidos momentáneos	95	0,93	0,26	1,00	1,00
42. La atención se centra en las responsabilidades del cuidado, no en otras tareas o intereses	95	0,95	0,22	1,00	1,00
43. El personal se muestra atento a todo el grupo inclusive mientras trabaja con un niño o un grupo pequeño	95	0,83	0,38	0,00	1,00
44. El personal reacciona rápidamente para solucionar los problemas de una manera reconfortante y fortalecedora	95	0,75	0,44	0,00	1,00
45. El personal juega con los niños y muestra interés y conocimiento por lo que hacen	95	0,71	0,46	0,00	1,00
46. El personal ayuda y anima a los niños cuando es necesario	95	0,60	0,49	0,00	1,00
47. La interacción entre niños es positiva durante la mayor parte del día	95	0,95	0,22	1,00	1,00
48. El personal facilita las interacciones positivas entre niños	95	0,55	0,50	0,00	1,00
49. El personal modela la interacción social positiva	95	0,78	0,42	0,00	1,00
50. Hay sonrisas, conversaciones y muestras de afecto ocasionales para con los niños durante el día	95	0,98	0,14	1,00	1,00
51. El personal por lo general responde de una manera comprensiva para ayudar a los niños dolidos, enojados o molestos	95	0,78	0,42	0,00	1,00
52. Interacción positiva frecuente entre el personal y los niños durante el día	95	0,74	0,44	0,00	1,00
53. El personal y los niños se muestran normalmente relajados, hablan agradablemente y sonrían con frecuencia	95	0,80	0,40	0,00	1,00
54. El personal nunca usa un castigo físico ni disciplina severos	95	0,98	0,14	1,00	1,00
55. El personal por lo general mantiene suficiente control para evitar problemas	95	0,84	0,37	0,00	1,00
56. El personal usa métodos de disciplina positiva de manera efectiva	95	0,51	0,50	0,00	1,00
57. Se presta atención frecuentemente cuando los niños se portan bien	95	0,58	0,50	0,00	1,00
58. El personal reacciona de manera consistente al comportamiento de los niños	95	0,84	0,37	0,00	1,00
59. El personal organiza actividades de juego como parte del horario diario	95	0,92	0,28	1,00	1,00
60. El horario de rutinas básicas es flexible e individualizado para satisfacer las necesidades de los niños	95	0,68	0,47	0,00	1,00
61. El horario proporciona equilibrio entre las actividades en el interior y las actividades al aire libre	95	0,76	0,43	0,00	1,00
62. No hay períodos largos de espera durante las transiciones entre acontecimientos diarios	95	0,57	0,50	0,00	1,00
63. El personal participa activamente para facilitar el juego de los niños durante el día	95	0,76	0,43	0,00	1,00
64. El personal por lo general es positivo y tolerante con los niños durante el juego en grupo	95	0,88	0,32	0,00	1,00
65. El personal es flexible y adapta la actividad a medida que los niños se unen o dejan el grupo	95	0,79	0,41	0,00	1,00
66. Se hacen modificaciones al ambiente, programa y horario para que niños puedan participar en muchas actividades con otros	95	0,46	0,50	0,00	1,00

Nota: La interpretación de cada línea es la siguiente. Si en la pregunta 1 la media es 0,91, esto significa que el 91% de centros observados cuentan con buenas practicas en este ítem, y habría que trabajar en este tema en el restante 9% de centros.

* se ignora o se dan respuestas negativas a los niños en poca o ninguna medida



Tabla C.3: Diferencias en características de los niños y de los EPI entre la muestra a la que se le aplicó Bayley y la muestra a la que no

	N	Se aplicó Bayley (N=302)		No se aplicó Bayley (N=628)		P-valor
		Media	DE	Media	DE	
Características Niño						
Edad (meses)*	930	25,39	5,53	24,69	6,51	0,32
Prematuro**	781	0,00	0,06	0,02	0,13	0,06
Total PRUNAPE: No Pasa	859	0,76	0,43	0,80	0,40	0,22
Características EPI						
Partido: La Matanza	930	0,44	0,50	0,47	0,50	0,73
Jornada Completa	930	0,31	0,47	0,33	0,47	0,86
Niños(as)/Número de Empleados	930	8,14	3,55	7,83	3,57	0,65
Niños(as)/Número de Educadores	930	16,07	6,67	15,26	7,44	0,56
Número de niños en el aula	930	20,75	7,96	21,02	7,33	0,83
Años de educación de la coordinadora	918	14,72	2,48	14,03	2,71	0,16
Años de educación de la educadora	930	13,78	2,06	14,11	2,18	0,39

*Edad al 1 de enero de 2017. **La información de prematuro proviene de la prueba PRUNAPE. Nota. Desvío estándar (DE) clustered a nivel de EPI. P valores de diferencia de medias menores o iguales a 0,10 en negrita.



D. La aplicación de la PRUNAPE en el contexto de los EPI

*

Buscando cumplir con las instrucciones de la administración de la PRUNAPE, en la carta de consentimiento informado se invitó a los padres a acompañar al niño durante la aplicación de la prueba. Sin embargo, solo en 2% de los casos la madre acompañó al niño. Como respuesta, se buscó que la educadora a cargo del aula estuviese durante la totalidad de la administración de la PRUNAPE.

A pesar de los esfuerzos para satisfacer las condiciones que supone la administración de la prueba, la dinámica de los EPI y la magnitud del estudio impusieron tres grandes retos. Primero, la rutina informal de los EPI: si bien se ofrece una jornada específica, los padres llevan y recogen a los niños en diferentes horarios, lo cual se traduce en un menor tiempo disponible para la administración de las pruebas y en un mayor engranaje de la logística en campo. De manera similar, los EPI manejan horarios de alimentación y de siesta que no pueden ser interrumpidos y que tienen un impacto directo sobre la disposición del niño a colaborar. Además, el niño era retirado de su sala y de su grupo de pares - su dinámica habitual- para llevarlo a una situación individual y no habitual (otra sala, la dirección del EPI, etc). Por último no era homogénea la cantidad de educadoras a cargo de la sala. Segundo, se encuentra la limitación del espacio físico: por la estructura y uso de los espacios fue difícil encontrar un lugar completamente aislado, libre de distracciones. Tercero, la disponibilidad de la educadora: cuando solo había una educadora y se evaluaban dos niños en simultáneo, esta debía alternarse y priorizaba el acompañamiento del niño más inhibido.

Los últimos dos factores implican una dificultad para neutralizar completamente las variables de contexto que proveen distractores al encuadre de la evaluación y es un factor que exige en los niños un control atencional, muchas veces mayor al esperado a su edad. El contexto institucional de los EPI no siempre puede ser equiparable, por más esfuerzo que se realice, a la aplicación de la prueba en un contexto clínico, con mayor “asepsia” de estímulos y con la presencia del /los cuidadores primarios junto al niño.

En el mejor de los casos el niño era acompañado por su educadora o educadora auxiliar, o por la coordinadora del EPI. Niños que estaban en reciente adaptación al jardín o naturalmente inhibidos e introvertidos en la sala, se mostraban desconfiados ante la evaluación. Algunos de ellos, a partir del manejo de las investigadoras de la situación de evaluación, lograban regular su comportamiento y responder a las consignas. Otros, rehusaban responder a las propuestas, lloraban o manifestaban su rechazo a la evaluación en forma manifiesta. Muchas de estas evaluaciones quedaron incompletas por este motivo.

En particular, el puntaje continuo (calculado como la suma de pautas que el niño completa con éxito, siguiendo a Rubio-Codina et al., 2016) presentó más errores que “no pasa PRUNAPE”, lo que es explicado por el número de pautas exigidas en cada caso. Mientras que el puntaje discreto solamente exige que el niño no apruebe 1 pauta A ó 2 pautas B (Item 6.9, páginas 117 y 118 manual PRUNAPE,), la administración del puntaje continuo es mucho más exigente, ya que requiere que TODOS los ítems estén puntuados, entre las 3 “pase”(P) de la base³¹ hasta las 3 “falla” (F) del techo.

Al analizarlo con mayor detalle se encuentra que, de las 930 pruebas administradas, en 71 (8%) no se administraron todas las pautas A y B por lo que no es posible decir nada al respecto sobre el desarrollo de estos niños. El dominio con mayor número de observaciones sin registrar la totalidad de pautas A fue Personal-social, mientras que Lenguaje fue el dominio con mayor número de errores

* Análisis de Claudia Viñuela Ortiz

³¹Para hallar la base se deben administrar los ítems de la línea de edad y continuar hacia la izquierda hasta encontrar tres pases consecutivos.



en el registro de las pautas B. El cálculo del techo fue el error más frecuente seguido del cálculo de la base, de manera que en 247 casos no es posible calcular el puntaje continuo. Para Motricidad gruesa llama la atención el alto porcentaje (20 %) de pruebas en las que no se halló el techo. Esto puede explicarse porque en algunas ocasiones los niños se sentían inhibidos y rechazaban realizar ítems y porque es el último dominio por evaluar y los niños ya se encontraban cansados.

A continuación se hace un análisis más profundo de los mayores retos encontrados.

Rutina del EPI:

La dinámica para la evaluación PRUNAPE fue configurándose y optimizándose según la organización horaria de cada EPI. La extensión horaria en que los niños asistían a cada EPI determinaba la exigencia del ritmo en la evaluación requerido para relevar la cantidad de niños mínima por EPI. La exigencia del tiempo, sumado a que muchos niños requieren más tiempo que otros para “entrar en confianza”, fue un factor potencial de eventuales errores de precisión en algunas de las evaluaciones.

En el caso de los EPI de jornada completa, PRUNAPE era administrada luego del almuerzo. En alguno de los EPI la dinámica horaria permitía el retiro de los niños en diversos momentos, según conveniencia de cada familia. Por ello, esta situación exigía priorizar la evaluación de los niños que se retiraban antes para no perder la posibilidad de administrar la técnica a los niños.

En los casos en que el EPI fuera de jornada simple, en la segunda visita se evaluaba con PRUNAPE a los niños, con consentimiento informado firmado por sus responsables parentales. Por ende, se debía recibir con anterioridad a la segunda visita, el listado de los niños con consentimiento informado. No siempre la misma pareja de evaluadoras que había visitado el EPI y observado la sala en la primera visita, era la misma que realizaba la evaluación PRUNAPE en la segunda jornada. Esta dinámica llevó a anticipar la confección previa de los protocolos con los datos básicos de los niños (edad, nombre, etc.) para maximizar los tiempos disponibles para la evaluación. Ello pudo realizarse gracias a que las investigadoras estuvieron disponibles a realizar la tarea de confección de los formularios en su domicilio, previo a la segunda visita al EPI.

Disponibilidad de un adulto referente para el niño:

Si bien en la mayoría de los casos se contó con una gran colaboración de parte del personal de los EPI, no siempre los niños contaron con un adulto referente que pudiera estar acompañándolo durante el 100 % del tiempo de la evaluación de PRUNAPE. En los casos en que había dos educadoras trabajan en la sala, una de ellas se quedaba con el grupo y la otra educadora acompañaba de manera alternada a 2 niños, evaluados en paralelo por dos investigadoras.

Así es que algunos (minoría) de los niños no contaban por momentos con la compañía del adulto referente, o se distraía buscando su presencia, o con lo que su otro compañero estaba haciendo. Por ende, la investigadora debía posponer algunas preguntas al educador, para puntuar algunos ítems que requieren, de no ser observados, del reporte del adulto que acompaña al niño. Ello predisponía algunas omisiones en la puntuación de estos ítems, quedando los protocolos incompletos.

Limitación del espacio físico:

El espacio disponible para ubicar dos mesitas para poder evaluar a dos niños en forma simultánea, fue otro factor a considerar. La mayoría de las técnicas psicométricas que evalúan el desarrollo infantil, suponen preparar un lugar libre de distracciones o con la menor cantidad de distractores posibles.



La ubicación y disposición de los materiales para presentar a los niños fue motivo en muchos de los EPI de un desafío, que no siempre concluyó en las condiciones óptimas para la evaluación.

La falta de espacio físico en muchos de los EPI sólo permitió trabajar en espacios muy reducidos, por lo que las dos investigadoras evaluaban a dos niños simultáneamente, en dos espacios contiguos, no pudiendo evitar la distracción de los niños por esta cercanía. En otros casos, el único lugar factible de ser utilizado, era un pasillo, lugar de paso de niños, y adultos.

Fueron utilizados espacios abiertos, cuando el clima lo permitía (patios, parques, a veces, con juegos a la vista, que los niños asociaban con el momento de esparcimiento).

Cabe destacar una situación particular respecto a las respuestas de un alto porcentaje de niños en las pruebas propuestas del área Motricidad gruesa. Dichas pruebas suponen habilidades de los niños respecto de sus movimientos gruesos y el uso de su cuerpo en el espacio amplio así como la utilización de diversas destrezas motoras, muchas de ellas a realizar por imitación directa del adulto. (Ej. salta con ambos pies, patear una pelota, lanza una pelota, párate en un pie por 5 segundos). Si bien la mayoría de estas propuestas se ofertaban luego de otras previas desplegadas sobre la mesa, cuando el niño ya había “entrado en confianza” con la investigadora, muchos niños mostraban su oposición o inhibición a siquiera pararse de la mesa para responder a dichas consignas. Otros se paraban, pero se inhibían progresivamente al pedido de las distintas pruebas. Debe resaltarse que las mismas se presentaban en un marco lúdico. (Ej. patear la pelota “para hacer un gol”, saltar con ambos pies “para tocar el techo”). Para facilitar las evaluaciones en esta área de Motricidad gruesa incluso se hacía participar muchas veces a la educadora que lo acompañaba (cuidando que diera en forma precisa la consigna), para facilitar que el niño pudiese mostrar la habilidad esperada. No siempre daba resultado.

Por ello, es interesante reflexionar acerca del eventual sentimiento de seguridad que implica para los niños pequeños el estar ante una mesa y la posible inseguridad que sienten al ser invitados a mostrar sus habilidades en el espacio amplio.

En un intento, entonces, de analizar el porcentaje alto de niños reacios a mostrar sus habilidades motrices, parados, fuera de la mesa en que habían “jugado”, se propone tomar para la reflexión algunos aportes teóricos del psicomotricista y escritor argentino Calmels (2011) respecto a la “construcción de un espacio habitado por el niño en edades tempranas”. Dice el autor: “El conocimiento del espacio requiere del cuerpo y de la acción. La sala de trabajo no se presenta como ingenua ni homogénea; el niño va a ir habitando este espacio a través de sus acciones, del juego corporal con los objetos” ... “La génesis del espacio habitado por el niño, después del cuerpo del adulto y de la cuna, tiene un punto de organización, que es el rincón”. Puede pensarse la situación en la mesa, rincón en función de refugio, en que el niño está sentado, inmóvil desde sus movimientos gruesos, desde la cual se le pide que transicione hacia la situación de pie; de la inmovilidad al movimiento. Al decir de Calmels, “De singular importancia es el momento en el que pueden salir de él ... (el rincón) ... y lo hacen a través de la pared. Al carecer de un sostén interiorizado que les permita moverse con comodidad por la sala, la pared los ayuda a entrar en tarea”. Al respecto, se observó recurrentemente en muchos de los niños evaluados, su necesidad de apoyo en la pared o en la silla o la mesa en la que habían estado sentados. Consecuentemente con la indicación de administración de alguno de los ítems de motricidad gruesa, debía pedírseles que no se sostuvieran en ningún objeto para dar respuesta a las consignas (ej. patear la pelota sin estar apoyado en la pared o sostenido en algún objeto; pararse/saltar en un pie, sin estar sostenido). “El centro de los espacios destinados al movimiento se connota como espacio vacío; el que accede al centro se expone, hace presente su existencia en un lugar que no posee paredes”. (Calmels, 2011) ¿Es probable que muchos de los niños, aún teniendo la habilidad motriz buscada, rehusaran a mostrarla? Siendo precisos en lo estipulado para la administración de la PRUNAPE, muchos de estos niños puntuaron Rehusa, en las pruebas



de motricidad gruesa, dado que, al mostrarse no disponibles, o ni siquiera intentar responder/imitar las consignas de las investigadoras, no puede corroborarse que fallaran en dichos ítems.

E. Análisis inicial de las propiedades psicométricas de los instrumentos de calidad de proceso en el GBA

Confiabilidad entre observadores

La confiabilidad de un instrumento es una de las propiedades psicométricas más importantes pues indica si el instrumento evalúa a los individuos con precisión. Para dar cuenta de ello existen diferentes indicadores, entre ellos la confiabilidad entre observadores, que mide el grado de acuerdo o consenso entre los observadores, es decir, indica qué tan homogéneo está siendo el juicio emitido por distintas personas que están evaluando a un mismo individuo o aula. Si este grado de acuerdo es bajo puede significar que la escala está mal construida o que es necesario dar mayor capacitación a los observadores, dado que el error de medición producido en su uso es alto.

Durante las prácticas y pilotos de este estudio se realizaron 3 días de ejercicios de confiabilidad entre duplas, observador-supervisor, este último considerado como un observador máster dado su alto grado de experiencia administrando el instrumento.³² Los resultados de estos ejercicios se presentan en la Tabla E.1. Para todos los instrumentos se obtuvo una confiabilidad mayor al 90 %, menos en un caso en que se obtuvo una confiabilidad del 80 %.³³ Este corresponde al ejercicio de la LV con una duración de la observación de treinta minutos. Para la Lista de Verificación se realizaron dos tipos de observaciones, una de treinta minutos de duración y otra con una duración de dos horas. Para el primer caso, el resultado puede considerarse de calidad media, que puede ser debido a que corresponden a ejercicios en los que la observación solamente se realizó durante 30 minutos y en momentos de la jornada de los EPI en donde no hay una interacción típica entre niños y educador, sino que está permeada por otros individuos y otro tipo de actividades donde la interacción puede ser mínima (horario de llegada, hora de almuerzo). De todas formas, en todos los casos se logró una confiabilidad de por lo menos el 80 %, que es lo que se busca para asegurar que los individuos se están evaluando de la misma manera.

³²Mientras el observador administra el test, el supervisor puntúa el test independientemente, es decir, sin conocer de los resultados del observador.

³³En ITERS se obtuvo un porcentaje mayor para la subescala *Escuchar y hablar*.



Tabla E.1: Confiabilidad entre observadores durante el piloto (% de acuerdo)

Área	
PRUNAPE	
Personal Social	98,9 %
Motricidad Fina	97,0 %
Lenguaje	98,0 %
Motricidad Gruesa	98,8 %
Toda la prueba	98,2 %
LISTA DE VERIFICACION	
Toda la escala (30 minutos)	80,3 %
Toda la escala (2 horas)	90,7 %
CLASS	90,0 %
ITERS	91,5 %

Nota 1. Para el test de Bayley, por haber sido una sola aplicadora, no se cuenta con ejercicio de confiabilidad.

Nota 2. Se realizó al menos un ejercicio de confiabilidad por observador.

Consistencia interna

Por otro lado, las Tablas E.2-E.4 resumen las medidas de consistencia interna de los tres instrumentos de calidad en sala. En la primera columna se presenta el Alfa de Cronbach de cada subescala (la línea en negrita de cada panel) y de los ítems (las líneas a continuación). El alfa del ítem corresponde al alfa de la subescala si no se incluyera ese ítem, por lo que lo deseable es que sea menor que el alfa de la subescala. En la segunda columna se presenta el coeficiente de correlación de Pearson del ítem con el puntaje de la subescala. Los resultados muestran que en el 50 % de los casos los alfas de Chronbach asociados a las subescalas (o el total) son iguales o mayores a 0,7, lo que se considera niveles aceptables de coherencia interna, y solo en el 22 % de los casos son menores de 0,6. Por otro lado, en el 65 % de los casos, el alfa asociado al ítem es menor al alfa de la subescala, lo que sugiere un alto nivel de coherencia interna. Esta calidad de los datos constituye una ventaja metodológica importante del presente estudio, la cual nos permite confiar en los resultados de este reporte y en las recomendaciones que se desprenden del mismo.



Tabla E.2: Consistencia Interna de los Instrumentos de Calidad

	Alfa	Correlación
CLASS		
Apoyo Emocional	0,88	
Clima positivo	0,81	0,95***
Clima negativo	0,88	-0,71***
Sensibilidad del educador	0,82	0,92***
Consideración por la perspectiva del niño	0,88	0,73***
Orientación de la conducta	0,86	0,81***
Apoyo Comprometido	0,67	
Facilitación del aprendizaje	0,74	0,59***
Calidad de la retroalimentación	0,29	0,90***
Modelaje lingüístico	0,67	0,89***
Total CLASS	0,88	
ITERS		
Escuchar y hablar	0,49	
Ayuda a los niños a entender el lenguaje	0,13	0,84***
Ayuda a los niños a utilizar el lenguaje		0,80***
Uso de libros	0,74	0,37***
Interacción	0,80	
Supervisión del juego y aprendizaje	0,75	0,79***
Interacción entre los niños	0,75	0,78***
Interacción entre el personal y niños	0,77	0,80***
Disciplina	0,74	0,82***
Total ITERS	0,82	
LV		
Receptividad	0,50	
P1.	0,38	0,63***
P2.	0,47	0,43***
P3.	0,52	0,46***
P5.	0,50	0,28***
P6.	0,43	0,64***
P7.	0,40	0,58***
P8.	0,50	0,50***
Materiales de Aprendizaje	0,63	
P10.	0,57	0,66***
P11.	0,54	0,71***
P12.	0,55	0,68***
P13.	0,63	0,42***
P14.	0,56	0,68***
Involucramiento	0,72	
P15.	0,66	0,72***
P16.	0,62	0,79***
P17.	0,65	0,73***
P18.	0,66	0,72***
P19.	0,75	0,43***



Tabla E.3: Consistencia Interna de los Instrumentos de Calidad

	Alfa	Correlación
Variedad	0,50	
P20.		0,82***
P21.		0,81***
Rutinas de Cuidado Personal	0,67	
P22.	0,49	0,79***
P23.	0,65	0,79***
P24.	0,60	0,78***
Siesta	0,72	
P25.	0,44	0,86***
P26.	0,98	0,62***
P27.	0,29	0,91***
Prácticas de Salud	0,61	
P28	0,31	0,83***
P29	0,32	0,81***
P30	0,78	0,60***
Escuchar y Hablar	0,70	
P31	0,68	0,51***
P32	0,62	0,74***
P33	0,69	0,47***
P34	0,69	0,44***
P35	0,68	0,53***
P36	0,64	0,71***
P37	0,64	0,73***
Actividades	0,43	
P38	0,22	0,73***
P39	0,26	0,71***
P40	0,48	0,60***
Interacción	0,86	
P41	0,86	0,36***
P42	0,87	0,16
P43	0,86	0,59***
P44	0,85	0,75***
P45	0,85	0,64***
P46	0,86	0,58***
P47	0,86	0,35***
P48	0,85	0,66***
P49	0,86	0,59***
P50	0,86	0,33***
P51	0,86	0,49***
P52	0,85	0,72***
P53	0,85	0,75***
P54	0,87	0,23**
P55	0,86	0,55***
P56	0,86	0,62***
P57	0,85	0,65***
P58	0,86	0,53***



Tabla E.4: Consistencia Interna de los Instrumentos de Calidad

	Alfa	Correlación
Estructura del Programa	0,65	
P59	0,58	0,68***
P60	0,63	0,51***
P61	0,67	0,35***
P62	0,64	0,52***
P63	0,59	0,62***
P64	0,60	0,58***
P65	0,54	0,76***
P66	0,67	0,43***
Total LV	0,90	

Nota. La muestra corresponde a los 95 EPI, para CLASS los 55 donde se midió. No se incluye la pregunta 4 de la Lista de Verificación pues no presenta variabilidad

Validez concurrente entre medidas de calidad de procesos

Para verificar la validez concurrente se revisaron tanto las correlaciones entre las subescalas del instrumento, como las correlaciones con los otros instrumentos de medición de la calidad de procesos. En la Tabla E.5 se observa que los constructos de ITERS y del CLASS que miden aspectos de proceso están moderadamente correlacionados, evidencia de la validez de las medidas entre sí.³⁴ También se observa que todas las dimensiones de CLASS están positiva y altamente asociadas entre sí, lo que es un signo de validez concurrente de la medida en la Argentina. De la misma manera, se observa que casi todas las subescalas de ITERS se correlacionan entre sí de manera significativa (correlaciones moderadas a altas), excepto la subescala *Uso de libros*, que no parece asociada de manera significativa con las otras subescalas del instrumento. *Uso de libros* solo está asociada con una de las otras subescalas del ITERS y con dos subescalas del CLASS, y estas correlaciones son bajas. Finalmente, la última línea de la Tabla E.5 muestra que la Lista de Verificación está positivamente asociada a todas las dimensiones de CLASS e ITERS con excepción de *Clima negativo* y *Facilitación del aprendizaje y el desarrollo* de CLASS, y *Uso de libros* de ITERS.³⁵ Esto indica que la Lista de Verificación capturaría las medidas de calidad de instrumentos más complejos. Cabe resaltar que los componentes de CLASS con los que no está asociado (*Clima negativo* y *Facilitación del aprendizaje*) son también aquellos que presentan menor correlación con ITERS.³⁶

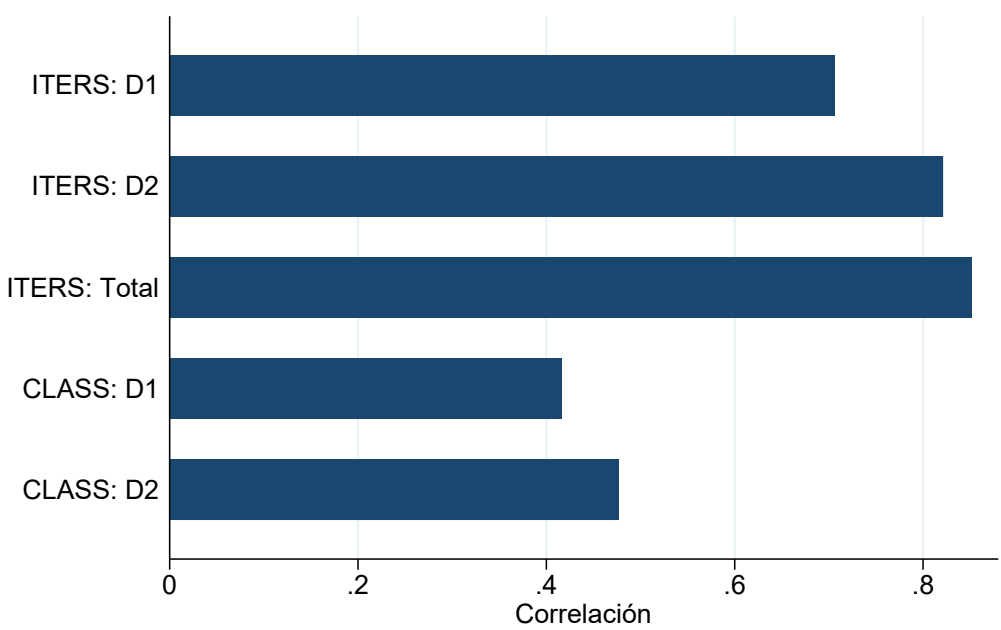
³⁴También usamos el puntaje alternativo ITERS (sin gradualidad, no reportado). En general, los resultados son similares a los de la Tabla E.5. Sin embargo, en el puntaje alternativo hay una asociación positiva entre *Uso de libros* y la subescala de *Escuchar y hablar*. Tal vez para esta subescala, la proporción de respuestas correctas está asociada de manera significativa pero el orden en el cual los encuestadores preguntaban el tema en el puntaje tradicional no es óptimo para capturar el contexto.

³⁵La Tabla E.6 contiene las correlaciones de la Lista de Verificación desagregada por constructos e ítems, y las subescalas del ITERS y los dominios del CLASS. Como se espera, los ítems de *Escuchar y hablar*, *Interacción* y *Estructura del programa*, se correlacionan positiva y significativamente con las subescalas del ITERS instrumento del cual provienen originalmente. Justamente también son estos ítems los que mejor correlación tienen con los dominios de CLASS.

³⁶También se observa (tabla no reportada por razones de espacio) que algunos constructos de la LV están más altamente correlacionados entre sí que otros, tales como: *Receptividad*, *Involucramiento*, *Escuchar y hablar*, *Interacción*, y *Estructura del programa*, todas variables proxies de la calidad de procesos. Otros constructos como *Organización*, *Variada*, *Rutinas de cuidado personal*, *Prácticas de salud* y *Actividades* no están alta ni significativamente corre-



Figura E.1: Correlación bivariada entre LV y resto



Nota: Todas las correlaciones son significativas al 1%.
ITERS: D1 es Escuchar y hablar; D2 es Interacción. CLASS: D1 es Apoyo emocional y conductual; D2 es Apoyo comprometido con el aprendizaje.

A manera de resumen, en la Figura E.1 se presentan las correlaciones entre puntajes totales y dimensiones principales con la LV. Se observa que las correlaciones entre medidas son de moderadas a altas. Las correlaciones entre LV y las medidas tradicionales de calidad son alentadoras y avalarían el uso de este instrumento como herramienta de supervisión rigurosamente validada. Por ejemplo, las correlaciones con ITERS son todas mayores a 0,70 (0,85 con el promedio), mientras que las correlaciones de LV con CLASS son mayores a 0,40.

lacionados entre sí; es decir, parece que no están capturando la misma medida de calidad. Estos resultados sugieren que tal vez no son las mejores preguntas para capturar la calidad de los EPI y que son dominios candidatos para condensar la Lista de Verificación como se realizó en la sección 4.4.

Tabla E.5: Correlación bivariada entre CLASS, ITERS y LV

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
CLASS																						
1. Clima Positivo	1,00																					
2. Clima Negativo	0,55***	1,00																				
3. Sensibilidad del Educador	0,87***	0,61***	1,00																			
4. Consideración por la Perspectiva del Niño	0,72***	0,38***	0,55***	1,00																		
5. Orientación de la Conducta	0,73***	0,60***	0,71***	0,32***	1,00																	
6. Dominio: Apoyo Emocional	0,95***	0,71***	0,92***	0,73***	0,81***	1,00																
7. Facilitación del Aprendizaje	0,37***	0,27***	0,42***	0,33***	0,33***	0,42***	1,00															
8. Calidad de la Retroalimentación	0,74***	0,32***	0,66***	0,60***	0,44***	0,69***	0,50***	1,00														
9. Modelaje Lingüístico	0,72***	0,24***	0,53***	0,75***	0,35***	0,65***	0,21	0,71***	1,00													
10. Dominio: Apoyo Comprometido	0,78***	0,33***	0,65***	0,74***	0,45***	0,74***	0,59***	0,90***	0,89***	1,00												
11. Total CLASS	0,94***	0,58***	0,86***	0,78***	0,70***	0,95***	0,53***	0,84***	0,81***	0,91***	1,00											
ITERS																						
12. Ayuda a los Niños a Entender el Lenguaje	0,34***	0,17	0,20	0,11	0,33***	0,28**	0,17	0,33**	0,26*	0,32**	0,32**	1,00										
13. Ayuda a los Niños a Utilizar el Lenguaje	0,39***	0,12	0,34**	0,22	0,29**	0,34***	0,12	0,38***	0,29**	0,33**	0,36***	0,59***	1,00									
14. Uso de Libros	0,15	0,13	0,03	0,29**	0,01	0,14	0,03	0,14	0,27**	0,22	0,19	0,03	-0,07	1,00								
15. Subescala: Escuchar y hablar	0,43***	0,21	0,29**	0,29**	0,32**	0,38***	0,16	0,42**	0,40***	0,42***	0,43***	0,84***	0,80***	0,37***	1,00							
16. Supervisión del Juego y Aprendizaje	0,44***	0,11	0,24*	0,35***	0,26*	0,35***	0,17	0,48***	0,40***	0,45***	0,42***	0,55***	0,55***	-0,03	0,56***	1,00						
17. Interacción entre los Niños	0,27***	0,19	0,25*	0,18	0,11	0,25*	0,15	0,38***	0,19	0,28**	0,28**	0,45***	0,56***	-0,03	0,51***	0,52***	1,00					
18. Interacción entre el Personal y Niños	0,22	-0,11	0,15	0,11	-0,05	0,11	-0,18	0,14	0,20	0,11	0,12	0,38***	0,54***	-0,08	0,45***	0,45***	0,53***	1,00				
19. Disciplina	0,43***	0,29**	0,35***	0,29**	0,34**	0,41***	0,15	0,40***	0,41***	0,44***	0,45***	0,57***	0,50***	0,00	0,56***	0,56***	0,52***	0,50***	1,00			
20. Subescala: Interacción	0,44***	0,14	0,31**	0,30**	0,21	0,36***	0,07	0,46***	0,39***	0,41***	0,41***	0,61***	0,67***	-0,04	0,65***	0,79***	0,78***	0,80***	0,82***	1,00		
21. Total ITERS	0,48***	0,18	0,33**	0,32**	0,27**	0,40***	0,12	0,49***	0,43***	0,45***	0,45***	0,75***	0,78***	0,11	0,84***	0,77***	0,75***	0,74***	0,79***	0,96***	1,00	
LV																						
22. Total - LV	0,50***	0,17	0,32**	0,35***	0,32**	0,42***	0,12	0,55***	0,45***	0,48***	0,47***	0,59***	0,64***	0,16	0,71***	0,66***	0,63***	0,58***	0,75***	0,82***	0,85***	1,00

Nota: Los coeficientes representan correlaciones de Person. Los números en las columnas equivalentes a la variable con el mismo número en los líneas. La muestra de las columnas 1-11 son los 55 EPI donde se levantó CLASS, para las demás columnas los 95 EPI.

* $p < 0,10$, ** $p < 0,05$, *** $p < 0,01$



Tabla E.6: Correlación Bivariada entre LV y Otros Instrumentos, por Ítem

	1. Escuchar y hablar	2. Interacción	3. Total ITERS	4. Apoyo emocional	5. Apoyo comprometido	6. Total CLASS
Receptividad	0,35***	0,44***	0,44***	0,23*	0,24*	0,25*
p1.	0,20*	0,16	0,19*		0,12	0,10
p2.	0,21**	0,32***	0,31***	0,17	-0,07	0,07
p3.	0,34***	0,29***	0,33***	0,03	0,22	0,12
p4.*						
p5.	0,08	0,19*	0,16	0,12	-0,07	0,04
p6.	0,22**	0,26**	0,27***	0,21	0,27**	0,25*
p7.	0,08	0,26**	0,21**	0,14	-0,06	0,05
p8.	0,06	0,11	0,10	-0,01	0,16	0,07
Organización	0,21**	0,22**	0,24**	0,23*	0,14	0,21
p9.	0,21**	0,22**	0,24**	0,23*	0,14	0,21
Materiales de Aprendizaje	0,02	0,08	0,06	0,10	0,15	0,13
p10.	-0,02	-0,07	-0,06	0,14	0,13	0,15
p11.	-0,09	0,12	0,05	-0,10	-0,07	-0,09
p12.	0,03	0,03	0,03	-0,04	0,07	0,01
p13.	0,19*	0,20**	0,22**	0,31**	0,32**	0,34**
p14.	0,08	0,04	0,06	0,15	0,15	0,16
Involucramiento	0,62***	0,55***	0,62***	0,38***	0,31**	0,37***
p15.	0,45***	0,41***	0,47***	0,33**	0,29**	0,33**
p16.	0,43***	0,37***	0,43***	0,21	0,22*	0,23*
p17.	0,42***	0,48***	0,50***	0,24*	0,22	0,24*
p18.	0,45***	0,29***	0,38***	0,29**	0,17	0,25*
p19.	0,37***	0,33***	0,37***	0,21	0,12	0,18
Variedad	0,21**	-0,04	0,05	0,11	0,22	0,17
p20.	0,15	-0,02	0,05	0,05	0,15	0,10
p21.	0,19*	-0,05	0,04	0,14	0,23*	0,19
Rutinas de Cuidado Personal	0,13	0,27***	0,24**	-0,06	-0,16	-0,11
p22.	0,11	0,18*	0,17*	0,03	-0,09	-0,02
p23.	0,16	0,24**	0,23**	0,00	-0,16	-0,08
p24.	0,03	0,19*	0,15	-0,10	-0,03	-0,08
Siesta	0,30***	0,48***	0,45***	0,24	0,20	0,24
p25.	0,26*	0,31**	0,32**	0,11	0,30	0,21
p26.	0,04	0,08	0,07	0,22	0,02	0,14
p27.	0,33***	0,51***	0,48***	0,21	0,25	0,24
Prácticas de Salud	0,23**	0,23**	0,25**	0,19	0,18	0,20
p28.	0,25**	0,25**	0,27***	0,26*	0,31**	0,30**
p29.	0,18*	0,18*	0,20*	0,07	0,10	0,09
p30.	0,08	0,09	0,09	0,10	-0,02	0,05
Escuchar y Hablar	0,72***	0,67***	0,75***	0,32**	0,32**	0,34**
p31.	0,37***	0,39***	0,42***	0,26*	0,27**	0,28**
p32.	0,42***	0,51***	0,52***	0,19	0,20	0,21
p33.	0,18*	0,25**	0,25**	-0,01	0,01	0,00
p34.	0,23**	0,18*	0,22**	0,11	0,26*	0,19
p35.	0,35***	0,43***	0,44***	0,19	0,24*	0,23*
p36.	0,64***	0,43***	0,55***	0,30**	0,30**	0,33**
p37.	0,65***	0,55***	0,64***	0,26*	0,13	0,22
Actividades	0,21**	0,15	0,19*	0,23*	0,31**	0,28**
p38.	0,07	0,11	0,10	0,14	0,17	0,16
p39.	0,14	0,09	0,12	0,09	0,17	0,13
p40.	0,23**	0,11	0,17	0,25*	0,29**	0,29**
Interacción	0,62***	0,91***	0,88***	0,34**	0,43***	0,41***
p41.	0,44***	0,36***	0,43***	0,36***	0,44***	0,42***
p42.	0,15	0,21**	0,21**	0,29**	0,25*	0,29**
p43.	0,45***	0,47***	0,50***	0,34**	0,46***	0,42***
p44.	0,47***	0,66***	0,64***	0,24*	0,26*	0,27**
p45.	0,34***	0,55***	0,52***	0,30**	0,36***	0,35***
p46.	0,31***	0,57***	0,52***	0,18	0,23*	0,22
p47.	0,07	0,39***	0,30***	-0,09	-0,15	-0,13
p48.	0,51***	0,58***	0,61***	0,12	0,33**	0,23*
p49.	0,41***	0,53***	0,53***	0,18	0,19	0,19
p50.	0,19*	0,23**	0,23**	-0,04	0,10	0,03
p51.	0,22**	0,53***	0,46***	0,04	-0,01	0,02
p52.	0,42***	0,63***	0,61***	0,21	0,30**	0,27**
p53.	0,39***	0,62***	0,59***	0,16	0,21	0,19
p54.	0,10	0,28***	0,23**	0,04	-0,08	-0,01
p55.	0,40***	0,51***	0,51***	0,30**	0,34**	0,34**
p56.	0,44***	0,61***	0,60***	0,26*	0,31**	0,30**
p57.	0,37***	0,60***	0,56***	0,25*	0,37***	0,32**
p58.	0,28***	0,42***	0,40***	0,14	0,05	0,10
Estructura del Programa.	0,52***	0,61***	0,63***	0,27**	0,37***	0,34**
p59.	0,37***	0,43***	0,44***	0,16	0,27**	0,22
p60.	0,11	0,24**	0,21**	-0,22	-0,04	-0,15
p61.	0,23**	0,18*	0,21**	0,03	0,01	0,03
p62.	0,43***	0,43***	0,47***	0,32**	0,28**	0,33**
p63.	0,42***	0,49***	0,51***	0,28**	0,22	0,27**
p64.	0,38***	0,39***	0,42***	0,10	0,20	0,16
p65.	0,42***	0,61***	0,59***	0,22	0,22	0,24*
p66.	0,01	0,01	0,01	0,19	0,33**	0,27**

*Este ítem no presenta variabilidad. Esta tabla contiene la correlación de Pearson entre los constructos de ITERS, de CLASS y la LV (66 ítems). A diferencia de la Tabla E.5 esta tabla presenta un análisis por constructo e ítem. * p<0.10, ** p<0.05, *** p<0.01.



La asociación entre la calidad de los procesos, las características de los EPI y el desarrollo infantil

En esta sección se revisan las asociaciones entre los instrumentos de calidad de proceso y las características estructurales de los EPI y el desarrollo infantil. La literatura ha encontrado que las variables estructurales están significativamente correlacionadas con la calidad de proceso al facilitar las interacciones de alta calidad (Howes, Phillips y Whitebook, 1992; Clarke-Stewart, Gruber y Fitzgerald, 1994; NICDH, 2000; Kagan, 2010). Sin embargo, por sí solas no garantizan un buen cuidado (Hamre y Pianta, 2007). Luego, esperábamos encontrar algunas correlaciones significativas, aunque moderadas en magnitud, entre los instrumentos que dan cuenta de la calidad de los procesos y las características estructurales de los EPI. Con relación al desarrollo infantil, se han encontrado de manera consistente relaciones entre los indicadores de la calidad del proceso y el desarrollo de los niños (La Paro, Pianta y Stuhlman, 2004; Hamre y Pianta, 2007). Luego, esperábamos que CLASS, ITERS y LV estuvieran significativamente asociados con los instrumentos de medición del desarrollo infantil. Se debe notar que una limitación central de este estudio es que mide el desarrollo infantil en un solo momento del tiempo, por lo que no se pueden computar las asociaciones con el cambio en el desarrollo del niño, que es lo que se acostumbra en esta literatura (Mashburn et al., 2008).

En la Tabla E.7 se presentan las correlaciones entre las medidas de calidad y las características estructurales (características de la coordinadora, educadora y elementos de la infraestructura y funcionamiento del EPI³⁷). Se observan pocas correlaciones significativas. CLASS solo presenta una correlación baja ($r=0,28$) con la variable dicotómica que representa si la educadora posee un título de primera infancia. Los Ítems del ITERS que captan ciertos elementos estructurales como, por ejemplo, si los espacios físicos son apropiados para realizar actividades lúdicas con los niños y si cuentan con ciertos materiales didácticos como libros, bloques y juguetes (y si los mismos son accesibles) son los que presentan las correlaciones estadísticas más fuertes con las características estructurales. En particular, la subescala *Escuchar y hablar* está positivamente correlacionada con los salarios de la coordinadora y de la educadora, mientras que la subescala de *Interacción* presenta una correlación marginalmente significativa con la variable de vacaciones remuneradas de la coordinadora. Por último, la LV también se correlaciona con las variables de vacaciones remuneradas y salario mensual de la coordinadora del centro. Además, se encuentra una correlación positiva y significativa con el número de días de vacaciones. En suma, las medidas de calidad global presentan pocas correlaciones con aspectos estructurales de los centros y las asociaciones que existen son pequeñas en magnitud. Esto era lo esperado dado que miden diferentes aspectos de la calidad del servicio.

En la Tabla E.8 se observan correlaciones negativas entre el dominio *Apoyo comprometido con el aprendizaje* del CLASS y el Bayley y, en particular, en la dimensión de *Modelaje lingüístico*, un resultado no anticipado. En *Modelaje lingüístico* es posible que la correlación negativa se dé porque los educadores están hablando más en aulas donde en promedio los niños tienen un menor desarrollo del lenguaje (inversión compensatoria), generado una situación de relación inversa. Por lo general los grupos de niños con mayores rezagos en su desarrollo también son los que requieren mayor atención y ayuda (e interacción verbal) de sus educadores.

También se observa una correlación significativa y positiva entre la subescala *Interacción* del ITERS y el desarrollo del niño, aunque el Ítem individual de *Uso de libros* presenta correlaciones

³⁷A partir de la información sobre infraestructura y seguridad del EPI se construyó un índice de estructura. Para la construcción de este índice se seleccionó previamente un conjunto de variables claves. Para no perder observaciones ni excluir información relevante sobre la infraestructura de los EPI, se imputó cualquier información faltante con cero. Con las variables completas para los 95 EPI, se construyó el índice por medio del método de componentes principales. Como alternativa, se crea un segundo índice utilizando el mismo método, pero restringiendo el conjunto de preguntas de infraestructura a aquellas con mayor variabilidad y sin ninguna observación faltante. A este índice se le llama Índice de estructura sin imputaciones.



negativas y estadísticamente significativas. En *Uso de libros* es posible que en los centros con más libros los educadores no están leyendo los cuentos de una forma que fomenta la participación activa, el análisis y razonamiento y las destrezas de lectura y lenguaje tempranos en los niños. Un cuento leído de manera no apropiada para la edad/nivel de desarrollo de los niños, o de una manera en la cual los niños no participan activamente, no promueve su desarrollo (Andrew et al., 2018). Este resultado se debe interpretar con cautela pues este ítem no funcionó bien en su aplicación, presentó una correlación baja con los otros ítems de la subescala y afectó negativamente la consistencia interna.

Por su parte, el puntaje total de la Lista de Verificación no resultó estar asociado con Bayley. Sin embargo, sí se asocia significativamente con las áreas de Motricidad fina y Personal-social de PRUNAPE. Al desagregar la Lista de Verificación por constructos, se encontraron asociaciones positivas entre varios de los constructos con el desarrollo de los niños. Resaltan los constructos de: *Organización* (4 asociaciones positivas, altas, incluyendo Bayley), *Prácticas de salud* (1 asociación positiva, moderada, con Motricidad fina de PRUNAPE), *Escuchar y hablar* (2 asociaciones positivas, altas, incluyendo Bayley), *Interacción* (2 asociaciones positivas, muy altas, incluyendo Bayley), *Estructura del programa* (1 asociación positiva, moderada, con Lenguaje de PRUNAPE), que tienen asociaciones muy consistentes con varios de los dominios del desarrollo (ver Tabla E.9). Asimismo, se encuentran solo dos asociaciones negativas: una entre *Rutinas de cuidado personal* y el total de PRUNAPE y la otra entre *Prácticas de salud* y Bayley. Ambos dominios de la LV constan de pocos ítems (3) y los constructos son de naturaleza más estructural, con lo que estos resultados deberían ser analizados en mayor profundidad en investigaciones futuras.

Para contextualizar los resultados, la Tabla E.10 analiza la asociación entre el desarrollo de los niños y las características de la coordinadora, la educadora, el coeficiente de atención niño/educadora y las condiciones estructurales de la sala. En esta, se observa que contar con coordinadoras con mayor nivel educativo está asociado con mayores puntajes en Bayley, mientras que poseer un título especializado en primera infancia se asocia con mayores puntajes en lenguaje (tanto Bayley, como en la área de lenguaje de la PRUNAPE). Sin embargo, no parece existir una asociación entre el nivel educativo de la educadora o su título de especialización y los puntajes en Bayley o PRUNAPE. Se debe recordar que una de las limitaciones centrales del estudio es que mide el desarrollo infantil en un solo momento de tiempo, por lo que no se puede controlar por el desarrollo inicial del niño para distinguir si maestras de mayores calificaciones o con especialización en DIT logran estimular un mayor progreso en el desarrollo de los niños matriculados en su sala. Adicionalmente, aunque una coordinadora con vacaciones remuneradas y aportes a la seguridad social, es decir, con mayor estabilidad laboral, está asociada con mayores puntajes en las pruebas de lenguaje (PRUNAPE), existe una asociación negativa entre los aportes a la seguridad social de la educadora y varias dimensiones del desarrollo de los niños. Esto invita a ser investigado con mayor profundidad ya que, por el contrario, la relación es positiva y alta para los aportes de la coordinadora en su relación con lenguaje de PRUNAPE. Además, la experiencia en el trabajo con niños está positivamente asociada con varias dimensiones de PRUNAPE. En cuanto a las condiciones del EPI, y como era de esperar, un mayor número de niños por empleado o educador está asociado negativamente al Bayley. Por último, el índice estructural está asociado con mayores puntajes en las áreas de motricidad fina y gruesa de la PRUNAPE, aunque las asociaciones solo son estadísticamente significativas para la variable en la que no se imputan valores cuando hay faltantes. El índice incluye variables relacionadas con la disponibilidad de material como, por ejemplo, si el EPI tiene sala múltiple donde los niños pueden jugar y que facilita la estimulación de las habilidades motoras del niño, lo que podría explicar la relación hallada.



Tabla E.7: Correlación Bivariada entre Medidas de Calidad y Variables Estructurales del EPI

	1.Dominio: Apoyo emocional	2.Dominio: Apoyo comprometido	3.Total CLASS	4.Subescala: Escuchar y hablar	5.Subescala: Interacción	6.Total ITERS	7.Total LV
Coordinadora							
1.Educación Terciaria o Superior	0,07	-0,02	0,03	0,02	-0,05	-0,02	0,09
2.Título en DTI	0,08	-0,04	0,03	0,14	0,03	0,07	0,09
3.Salario Mensual Líquido (pesos)	0,01	-0,11	-0,05	0,25**	0,02	0,11	0,22**
4.Aportes a Seguridad Social	0,11	0,15	0,13	0,14	0,16	0,17	0,19*
5.Vacaciones Remuneradas	0,09	0,21	0,15	0,17	0,18*	0,19*	0,21**
6. Empleo: con nombramiento	-0,01	0,10	0,04	0,17	0,08	0,12	0,11
Educadora							
7.Educación Terciaria o Superior	0,06	-0,09	-0,01	0,01	-0,12	-0,08	-0,11
8.Título en DTI	0,28**	0,09	0,21	0,06	0,06	0,07	0,10
9.Salario Mensual Líquido (pesos)	0,11	0,03	0,08	0,18*	-0,01	0,06	0,13
10.Aportes a Seguridad Social	-0,12	-0,11	-0,12	-0,03	-0,03	-0,03	0,08
11.Vacaciones Remuneradas	0,08	0,22	0,15	0,15	0,06	0,10	0,16
12.Empleo: con nombramiento	0,02	-0,02	0,00	-0,04	0,02	0,00	0,01
Estructura							
13.Niños/ Educadores (EPI)	-0,10	-0,08	-0,10	0,03	0,01	0,02	-0,01
14.Número de niños en el aula (Sala de 2)	-0,14	-0,10	-0,13	-0,02	-0,17	-0,13	-0,15
15.Índice Estructural (Sin imputaciones)	-0,02	0,06	0,01	0,07	0,03	0,05	0,04
16.Índice Estructural [†]	0,09	0,14	0,12	0,15	0,11	0,14	0,08
Funcionamiento							
17.Número de Años de Funcionamiento	-0,05	-0,16	-0,10	-0,10	-0,17*	-0,16	-0,10
18.Meses al Año que Ofrece Atención	-0,12	-0,21	-0,17	-0,10	-0,07	-0,09	-0,10
19.Días de Vacaciones	0,11	0,04	0,09	0,21*	0,13	0,17	0,23**
20.Familias deben realizar algún pago	-0,16	0,02	-0,09	-0,04	0,05	0,02	-0,01
21.Pago promedio mensual (Pesos)	-0,20	-0,15	-0,19	-0,18*	-0,11	-0,15	-0,16

[†] Como control se incluye una variable dicotómica por cada variable utilizada en la construcción del índice que indica si para esa variable se realizó o no una imputación. Nota: En las correlaciones de CLASS el número de observaciones es 55, en el resto 95 (puede variar ligeramente por missings). * p<0.10, ** p<0.05, *** p<0.01.



Tabla E.8: Asociación entre CLASS, ITERS, LV y desarrollo del niño

	Bayley						PRUNAPE					
	Lenguaje Receptivo		Personal Social		Motricidad Fina		Lenguaje		Motricidad Gruesa		Total	
CLASS												
Clima Positivo	0,12	(0,15)	0,09*	(0,05)	0,05	(0,05)	0,15*	(0,08)	0,12**	(0,06)	0,14**	(0,05)
Clima Negativo	0,28***	(0,06)	0,09**	(0,04)	0,04	(0,06)	-0,01	(0,08)	0,00	(0,06)	0,04	(0,04)
Sensibilidad del Educador	0,25*	(0,14)	0,11***	(0,04)	0,03	(0,06)	0,08	(0,07)	0,05	(0,08)	0,09**	(0,04)
Consideración por la Perspectiva del Niño	-0,30	(0,19)	-0,02	(0,07)	-0,04	(0,06)	-0,00	(0,11)	-0,02	(0,06)	-0,02	(0,06)
Orientación de la Conducta	0,31***	(0,05)	0,10**	(0,04)	-0,01	(0,05)	0,00	(0,07)	-0,01	(0,07)	0,03	(0,05)
Dominio: Apoyo Emocional	0,28*	(0,14)	0,11***	(0,04)	0,02	(0,06)	0,07	(0,08)	0,04	(0,08)	0,08*	(0,05)
Facilitación del Aprendizaje	-0,11	(0,07)	-0,00	(0,06)	0,09*	(0,05)	0,05	(0,10)	0,06	(0,07)	0,05	(0,07)
Calidad de la Retroalimentación	-0,19	(0,22)	0,02	(0,05)	0,12	(0,09)	0,07	(0,10)	0,08	(0,06)	0,08	(0,06)
Modelaje Lingüístico	-0,35***	(0,10)	0,03	(0,07)	0,06	(0,08)	0,19**	(0,09)	0,10	(0,08)	0,13**	(0,06)
Dominio: Apoyo Comprometido	-0,29**	(0,14)	0,03	(0,06)	0,12	(0,10)	0,17	(0,10)	0,12	(0,07)	0,14**	(0,07)
Total CLASS	0,09	(0,18)	0,11*	(0,06)	0,08	(0,08)	0,14	(0,11)	0,09	(0,07)	0,14**	(0,06)
N	182		382		382		382		382		382	
ITERS												
Ayuda a los Niños a Entender el Lenguaje	0,15**	(0,07)	0,07***	(0,02)	0,04	(0,04)	0,05	(0,04)	-0,01	(0,04)	0,05	(0,03)
Ayuda a los Niños a Utilizar el Lenguaje	0,16	(0,12)	0,05	(0,05)	0,11*	(0,06)	0,02	(0,06)	0,09	(0,06)	0,07	(0,05)
Uso de Libros	-0,30***	(0,08)	-0,01	(0,05)	0,05	(0,05)	0,00	(0,07)	-0,02	(0,05)	0,00	(0,04)
Subescala: Escuchar y Hablar	0,11	(0,14)	0,07*	(0,04)	0,11**	(0,05)	0,05	(0,06)	0,02	(0,06)	0,07*	(0,04)
Supervisión del Juego y Aprendizaje	0,17***	(0,06)	0,05	(0,03)	0,05	(0,04)	0,08	(0,05)	0,03	(0,06)	0,07	(0,04)
Interacción entre los Niños	0,03	(0,16)	0,08**	(0,03)	0,01	(0,05)	-0,02	(0,04)	-0,07	(0,05)	0,00	(0,03)
Interacción entre el Personal y Niños	0,06	(0,12)	0,06	(0,04)	-0,04	(0,05)	0,05	(0,05)	-0,08	(0,06)	0,02	(0,04)
Disciplina	0,19**	(0,08)	0,07**	(0,03)	0,07*	(0,04)	0,01	(0,05)	-0,01	(0,06)	0,04	(0,04)
Subescala: Interacción	0,18***	(0,07)	0,08**	(0,03)	0,03	(0,05)	0,04	(0,04)	-0,04	(0,06)	0,04	(0,04)
N	292		657		657		657		657		657	
Lista de Verificación												
Total LV	0,97	(0,64)	0,54**	(0,27)	0,61*	(0,34)	-0,01	(0,46)	0,05	(0,42)	0,33	(0,31)
N	292		657		657		657		657		657	

Nota: Errores estándar en paréntesis (agrupados a nivel EPI). Todas las regresiones incluyen efectos fijos de referente y partido. * p<0.10, ** p<0.05, *** p<0.01.



Tabla E.9: Asociación entre Lista de Verificación y desarrollo del niño

	PRUNAPE						
	Bayley		Personal Social	Motricidad Fina	Lenguaje	Motricidad Gruesa	Total
	Lenguaje Receptivo						
Receptividad (8 ítems)	0,77 (0,84)	0,06 (0,19)	0,09 (0,23)	-0,40 (0,29)	0,08 (0,28)	-0,12 (0,16)	
Organización (1 ítem)	0,85*** (0,24)	0,16 (0,10)	0,10 (0,09)	0,40*** (0,15)	0,61*** (0,16)	0,41*** (0,12)	
Materiales de Aprendizaje (5 ítems)	0,10 (0,38)	-0,03 (0,15)	0,25 (0,16)	0,16 (0,19)	0,08 (0,17)	0,13 (0,15)	
Involucramiento (5 ítems)	0,29 (0,56)	0,15 (0,12)	0,03 (0,12)	-0,04 (0,18)	-0,05 (0,22)	0,03 (0,15)	
Variedad (2 ítems)	-0,15 (0,17)	-0,00 (0,12)	0,16 (0,12)	-0,02 (0,15)	-0,01 (0,12)	0,02 (0,11)	
Rutinas de Cuidado Personal (3 ítems)	-0,78 (0,52)	-0,11 (0,22)	-0,21 (0,23)	-0,33 (0,31)	-0,44 (0,28)	-0,34** (0,16)	
Prácticas de Salud (3 ítems)	-0,48** (0,22)	0,13 (0,12)	0,38*** (0,13)	-0,09 (0,15)	-0,01 (0,16)	0,08 (0,11)	
Escuchar y Hablar (7 ítems)	0,65** (0,31)	0,40** (0,18)	0,41 (0,25)	0,10 (0,26)	0,24 (0,24)	0,33 (0,22)	
Actividades (3 ítems)	-0,74 (0,45)	-0,03 (0,14)	0,39*** (0,11)	-0,26 (0,18)	-0,01 (0,20)	-0,04 (0,12)	
Interacción (18 ítems)	0,88** (0,40)	0,47*** (0,16)	0,12 (0,25)	-0,01 (0,25)	-0,07 (0,30)	0,17 (0,21)	
Estructura del Programa (8 ítems)	0,52 (0,96)	0,02 (0,17)	0,09 (0,18)	0,49* (0,23)	-0,11 (0,44)	0,21 (0,23)	
N	292	657	657	657	657	657	

Nota: Errores estándar en paréntesis (agrupados a nivel EPI). Todas las regresiones incluyen efectos fijos de referente y partido. * p<0.10, ** p<0.05, *** p<0.01.

Tabla E.10: Asociación entre características del educador, coordinador, condiciones del EPI y desarrollo del niño

	PRUNAPE					
	Bayley			PRUNAPE		
	Lenguaje Receptivo	Personal Social	Motricidad Fina	Lenguaje	Motricidad Gruesa	Total
Coordinadora						
Educación Terciaria o Superior	0,21**	-0,02	-0,15	0,16	-0,11	0,01
Título en Primera Infancia	0,14**	-0,03	0,08	0,21*	0,03	0,10
Salario Mensual Líquido (pesos)	0,00	0,00	0,00*	0,00	0,00	0,00
Aportes a Seguridad Social	0,06	-0,11	0,13	0,34**	0,18	0,19
Vacaciones Remuneradas	-0,16	-0,08	0,04	0,25*	0,08	0,11
Condiciones Laborales						
Ha trabajado antes con niños (0-60)	0,06	0,07	-0,09	0,21	-0,02	0,09
Empleo: con nombramiento	-0,06	-0,02	0,37***	0,01	0,07	0,09
Condiciones del EPI						
Número de Niños Presentes	0,00	0,00**	0,00**	-0,00	0,00	0,00
Niños (as) / Número de Empleados	-0,02	-0,00	0,03**	-0,04**	-0,01	-0,02
Niños (as) / Número de Educadores	-0,01	-0,01	0,01*	-0,02**	-0,01	-0,01
Educadora						
Educación Terciaria o Superior	0,02	0,03	-0,04	0,10	0,16	0,09
Título en Primera Infancia	-0,18	0,09	0,01	-0,07	-0,22	-0,05
Años de Educación	-0,07	0,02	-0,02	0,02	0,00	0,01
Salario Mensual Líquido (pesos)	-0,00	0,00	0,00	-0,00	0,00	0,00
Aportes a Seguridad Social	-0,37**	-0,06	-0,29*	-0,14	-0,49***	-0,27**
Vacaciones Remuneradas	0,21	-0,07	0,19*	0,11	-0,19	0,02
Condiciones Laborales						
Ha trabajado antes con niños (0-60)	0,00	0,35***	0,08	-0,08	0,38***	0,20*
Empleo: con nombramiento	0,35	0,12*	0,00	-0,08	0,02	0,01
Condiciones del Aula						
Número de niños en el aula	-0,02*	-0,00	0,01	0,00	-0,00	0,00
Aula tiene menos de 20 niños	-0,05	-0,04	-0,02	-0,11	-0,11	-0,09
Aula tiene menos de 20 niños	-0,05	-0,04	-0,02	-0,11	-0,11	-0,09
Estructura						
Índice Estructural (Sin imputaciones)	0,03	0,02	0,05**	0,03	0,06**	0,05*
Índice Estructural†	-0,09	0,02	0,02	0,04	0,00	0,03
N*	292	657	657	657	657	657

* El número de observaciones puede variar ligeramente por valores faltantes.

† Como control se incluye una variable dicotómica por cada variable utilizada en la construcción del índice que indica si para esa variable se realizó o no una imputación.

Nota: Errores estándar en paréntesis (agrupados a nivel EPI). Todas las regresiones incluyen efectos fijos de referente y partido. * p<0,10, ** p<0,05, *** p<0,01.



F. Análisis inicial de las propiedades psicométricas de las pruebas de desarrollo infantil en el GBA

La PRUNAPE es un test para la detección de problemas inaparentes del desarrollo que ha sido utilizada en diversos estudios y programas en Argentina (Lejarraga et al., 2014; Lejarraga et al., 2011). Con el propósito de analizar las propiedades psicométricas de la PRUNAPE, utilizando la muestra de este estudio, se llevan a cabo dos ejercicios. Primero, se analiza la consistencia interna del instrumento a través del Alfa de Cronbach. Segundo, se estudia la validez concurrente a través de correlaciones de Pearson entre el instrumento y la subescala de lenguaje receptivo del Bayley, un test de desarrollo mental y psicomotor reconocido internacionalmente por sus excelentes propiedades psicométricas (Pearson, 2018).

Para el análisis de la consistencia interna se excluyen los ítems que no presentan variabilidad. Estos son los ítems del comienzo de la prueba que toman el valor de cero para todos los niños. Para la interpretación de los Alfas de Cronbach es importante recordar que lo deseable es que el alfa asociado a cada dominio sea mayor a 0,8 (considerados como buenos). Alfas entre 0,7 y 0,8 son considerados como aceptables. Por el contrario, para los alfas asociados a cada ítem es deseable que estos sean menores al alfa de la escala, pues los primeros muestran el comportamiento de la escala en ausencia de ese ítem. De esta forma, si el ítem es relevante la escala debe tener un comportamiento menos fiable en su ausencia.

Se observan buenos alfas en los dominios Personal-social y Lenguaje, y aceptables en los dominios de Motricidad Fina y Gruesa (Tablas F.1- F.2). A nivel de ítem, en la mayoría de los casos los alfas son inferiores o iguales a los alfas de los dominios y en todos los casos los ítems se correlacionan de manera positiva y significativa.

Tabla F.1: Consistencia Interna: PRUNAPE (I)

	Correlación	Alfa
Personal-social		0,80
Busca objeto	0,14***	0,81
Come solo	0,34***	0,80
Da un objeto	0,32***	0,80
Ayuda en tareas en el hogar	0,51***	0,80
Juego simbólico	0,57***	0,79
Imita tareas del hogar	0,57***	0,79
Acude al llamado del observador	0,61***	0,79
Control de esfínteres diurno	0,7***	0,77
Se quita ropa o zapatos	0,66***	0,77
Se pone ropa o zapatos	0,68***	0,77
Arma rompecabezas	0,67***	0,77
Aparea colores	0,56***	0,79
Junta dibujos semejantes	0,44***	0,80



Tabla F.2: Consistencia Interna: PRUNAPE (II)

	Correlación	Alfa
Motricidad fina		0,70
Presión pinza superior	0,15***	0,71
Introduce pasa en botella	0,21***	0,70
Garabatea	0,23***	0,69
Verte pasa en botella	0,29***	0,69
Torre de 4 cubos	0,38***	0,69
Torre de 8 cubos	0,62***	0,67
Imita puente	0,66***	0,66
Corrige torre	0,7***	0,65
Dibuja persona (3 partes)	0,61***	0,66
Dobla un papel en diagonal	0,4***	0,69
Copia cruz	0,5***	0,68
Dibuja persona (6 partes)	0,37***	0,69
Copia un triangulo	0,19***	0,70
Lenguaje		0,87
Silabeo da-da-da	0,24***	0,88
Papá mamá, NO específico	0,28***	0,87
Papá mamá, específico	0,38***	0,87
Palabra frase	0,46***	0,87
Señala 2 figuras	0,55***	0,87
Nombra 2 figuras	0,67***	0,86
Frases (sustantivo y verbo)	0,71***	0,86
Tararea en presencia de terceros	0,69***	0,86
Frase completa	0,7***	0,86
Dice su nombre completo	0,63***	0,86
Analogías opuestas	0,55***	0,86
Comprende preposiciones	0,6***	0,86
Cumple 2 indicaciones consecutivas	0,57***	0,86
Reconoce 3 colores	0,57***	0,86
Uso de 2 objetos	0,48***	0,87
Sabe porque es de día y porque es de noche	0,42***	0,87
Motricidad gruesa		0,73
Logra pararse	0,27***	0,73
Camina sujeto a muebles	0,27***	0,73
Camina sujeto de la mano	0,34***	0,73
Camina solo	0,38***	0,73
Se agacha y se levanta sin sostén	0,39***	0,73
Sube a una silla sin ayuda	0,4***	0,72
Patea pelota	0,28***	0,71
Lanza pelota al examinador	0,38***	0,71
Salta con ambos pies	0,34***	0,71
Se para en un pie 5"	0,47***	0,70
Salto amplio	0,48***	0,70
Salta en un pie	0,46***	0,70
Camina talón punta	0,39***	0,71
Retrocede talón punta	0,34***	0,72



En la Tabla F.3 se presentan los resultados del análisis de consistencia interna para la subescala de lenguaje receptivo del Bayley. De nuevo, se excluyen del análisis los ítems que no presentan variabilidad (estos son los primeros 17 ítems, el ítem 19 y el ítem 46). El alfa de la subescala es bueno (0,88) y, a excepción de un caso, el alfa de los ítems es igual o inferior, que es lo deseable. Adicionalmente, todos los ítems se correlacionan de manera positiva y significativa.

Tabla F.3: Consistencia Interna: Subescala de Lenguaje Receptivo, Bayley

	Correlación	Alfa
Lenguaje Receptivo		0,88
Entiende órdenes inhibitorias	0,16***	0,88
Sigue indicaciones simples	0,23***	0,88
Identificación de imágenes: 3 aciertos	0,24***	0,88
Identifica 3 prenda de vestir	0,52***	0,87
Identificación de acciones: 1 acierto	0,3***	0,88
Identifica 5 partes del cuerpo	0,39***	0,87
Sigue instrucciones de 2 consignas	0,51***	0,87
Identificación de acciones: 3 aciertos	0,56***	0,87
Comprende el uso de objetos	0,69***	0,87
Comprende relación parte-todos	0,67***	0,87
Identificación de acciones: 5 aciertos	0,66***	0,87
Entiende ciertos pronombres y determinantes	0,65***	0,87
Discrimina tamaños	0,71***	0,87
Comprensión de conceptos básicos espaciales: 2 aciertos	0,58***	0,87
Entiende la posesión	0,7***	0,87
Comprende estar + gerundio	0,46***	0,87
Identifica colores	0,56***	0,87
Entiende el concepto uno	0,64***	0,87
Entiende los pronombres (ellos, ellas, él)	0,6***	0,87
Entiende la posesión (de él, de ella)	0,65***	0,87
Entiende los plurales	0,32***	0,88
Entiende el concepto más	0,55***	0,87
Entiende el concepto el más	0,42***	0,87
Comprensión de conceptos básicos espaciales: 4 aciertos	0,34***	0,88
Entiende oraciones negativas	0,36***	0,87
Entiende oraciones en pasado	0,27***	0,88
Entiende el concepto del peso	0,19***	0,88
Entiende el concepto el menos	0,06	0,88
Entiende descripciones	0,12**	0,88
Identifica categorías	0,11*	0,88

Finalmente, en la Tabla F.4 se presentan las correlaciones entre los dominios de la prueba PRUNAPE y la subescala de lenguaje receptivo del Bayley, dos pruebas que dan cuenta del desarrollo infantil y se espera estén altamente correlacionadas. En general, las correlaciones son significativas y tienen los signos esperados. Se debe notar que, como se esperaba, la subescala del Bayley tiene correlaciones más altas con los puntajes de los dominios Personal-social y Lenguaje de la PRUNAPE.



Tabla F.4: Correlación entre la PRUNAPE y la subescala de lenguaje receptivo del Bayley

	1	2	3	4	5	6	7	8	9	10	11
Prunape											
1.Puntaje personal-social	1										
2.Puntaje motor fino	0,55***	1									
3.Puntaje lenguaje	0,60***	0,52***	1								
4.Puntaje motor grueso	0,49***	0,51***	0,50***	1							
5.Personal-social: No pasa	-0,39***	-0,11***	-0,23***	-0,08**	1						
6.Motor fino: No pasa	-0,14***	-0,39***	-0,18***	-0,11***	0,06*	1					
7.Lenguaje: No pasa	-0,22***	-0,09***	-0,48***	-0,14***	0,23***	0,22***	1				
8.Motor grueso: No pasa	-0,00	0,01	-0,08**	-0,11***	0,19***	0,12***	0,18***	1			
9.No pasa Prueba	-0,24***	-0,21***	-0,29***	-0,16***	0,43***	0,36***	0,47***	0,40***	1		
10.Número de Áreas con No Pasa	-0,32***	-0,25***	-0,42***	-0,18***	0,60***	0,55***	0,67***	0,61***	0,68***	1	
Bayley											
11.Puntaje Bruto	0,62***	0,47***	0,59***	0,34***	-0,19***	-0,22***	-0,29***	0,06	-0,20***	-0,27***	1

Nota: Los números en las columnas equivalen a la variable con el mismo número en las líneas.

* $p < 0,10$, ** $p < 0,05$, *** $p < 0,01$



G. Análisis adicional de la Lista de Verificación corta

Tabla G.1: Validez de Constructos de la LV (57 ítems)

	Valor Propio	Carga
Receptividad	1,11	
P1.		0,66
P2.		0,34
P3.		0,21
P6.		0,40
P7.		0,51
P8.		0,31
Materiales de Aprendizaje	1,06	
P10.		0,49
P11.		0,57
P12.		0,49
P14.		0,51
Involucramiento	1,68	
P15.		0,60
P16.		0,70
P17.		0,64
P18.		0,58
P19.		0,28
Variedad	0,44	
P20.		0,47
P21.		0,47
Rutinas de Cuidado Personal	1,13	
P22.		0,69
P23.		0,55
P24.		0,60
Prácticas de Salud	1,12	
P28		0,73
P29		0,72
P30		0,25



Tabla G.2: Validez de Constructos de la LV (57 ítems)

	Valor Propio	Carga
Escuchar y Hablar	1,74	
P31		0,39
P32		0,68
P33		0,42
P35		0,41
P36		0,61
P37		0,64
Actividades	0,51	
P38		0,47
P39		0,46
P40		0,27
Interacción	4,94	
P41		0,36
P42		0,13
P43		0,58
P44		0,73
P45		0,60
P46		0,54
P47		0,32
P48		0,63
P49		0,57
P51		0,44
P52		0,69
P53		0,75
P55		0,54
P56		0,58
P57		0,60
P58		0,48
Estructura del Programa	2,13	
P59		0,73
P60		0,33
P61		0,17
P62		0,46
P63		0,53
P64		0,56
P65		0,75
P66		0,29

Tabla G.3: Items descartados de la LV (66 ítems)

	N	Proporción	Loading	Criterio que incumple
Receptividad				
2. Educador/a responde verbalmente a vocalizaciones o verbalizaciones del niño.	95	0,92	0,43	3
3. Educador/a le dice o indica al niño el nombre de un objeto o persona durante la visita.	95	0,75	0,46	3
4. Cuando el educador/a habla, lo que se dice se distingue con facilidad, es claro, se escucha bien.	95	1		2
5. Educador/a conversa libremente y fácilmente.	95	0,96		2
6. Educador/a espontáneamente elogia al niño al menos dos veces.	95	0,46	0,64	3
8. Educador/a acaricia o besa al niño al menos una vez.	95	0,76	0,50	3
Materiales de aprendizaje				
10. Juegos o equipos disponibles que permiten al niño ejercitar sus músculos.	95	0,71	0,49	3
12. Muñeco de peluche o muñeco suave, juegos para imitar roles disponibles.	95	0,69	0,49	3
13. Equipo que facilita el aprendizaje - móviles, mesa y sillas pequeñas, silla de comer alta, corral disponible.	95	0,96		2
Involucramiento				
19. Educador/a mantiene al niño dentro de su rango visual, y lo mira seguido.	95	0,87	0,28	3
Variedad				
20. El educador/a lee cuentos con el niño al menos 3 veces por semana.	94	0,49	0,46	3
21. El niño tiene acceso a 3 o más libros en el jardín.	94	0,40	0,46	3
Siesta				
25. Hay suficiente supervisión para los niños durante la siesta.	46			1
26. Las cunas (o colchonetas o camas infantiles) son para dormir y no para jugar durante largos períodos de tiempo.	47			1
27. La supervisión es agradable, atenta y cariñosa.	73			1
Prácticas de salud				
30. Los niños son atendidos adecuadamente para satisfacer sus necesidades de salud adentro y afuera.	95	0,76	0,25	3
Escuchar y hablar				
31. El contenido de la conversación es generalmente alentador y positivo más que desalentador y negativo.	95	0,93	0,39	3
33. La comunicación verbal es personalizada.	95	0,88	0,42	3
34. Hay una medida moderada de respuesta positiva verbal o no verbal a los intentos de los niños de comunicarse durante el día; se ignora o se dan respuestas negativas a los niños en poca o ninguna medida.	95	0,96		2
35. Hay algunos intentos de interpretar correctamente lo que los niños están intentando comunicar durante el día.	95	0,92	0,41	3
Actividades				
38. Hay como mínimo un juego de bloques (6 o más bloques del mismo tipo) accesibles diariamente.	95	0,65	0,47	3
39. Los materiales son accesibles diariamente, durante gran parte del día	95	0,33	0,46	3
40. Los materiales se encuentran organizados por tipo	95	0,73	0,27	3
Interacción				
41. El personal puede ver, oír y alcanzar a los niños fácilmente, con no más de unos pocos descuidos momentáneos.	95	0,93	0,36	3
42. La atención se centra en las responsabilidades del cuidado, no en otras tareas o intereses.	95	0,95	0,13	3
47. La interacción entre niños es positiva durante la mayor parte del día.	95	0,95	0,32	3
50. Hay sonrisas, conversaciones y muestras de afecto ocasionales para con los niños durante el día.	95	0,98		2
51. El personal por lo general responde de una manera comprensiva para ayudar a los niños dolidos, enojados o molestos.	95	0,78	0,44	3
54. El personal nunca usa un castigo físico ni disciplina severos.	95	0,98		2
58. El personal reacciona de manera consistente al comportamiento de los niños.	95	0,84	0,48	3
Estructura del Programa				
60. El horario de rutinas básicas es flexible e individualizado para satisfacer las necesidades de los niños.	95	0,68	0,33	3
61. El horario proporciona equilibrio entre las actividades en el interior y las actividades al aire libre.	95	0,76	0,17	3
62. No hay períodos largos de espera durante las transiciones entre acontecimientos diarios.	95	0,57	0,46	3
66. Se hacen modificaciones al ambiente, al programa y al horario para que los niños puedan participar en muchas actividades con otros niños.	95	0,46	0,29	3



Por otro lado, con el objetivo de estudiar si la LV corta logra discriminar niveles de calidad medios y altos, y no solamente inadecuados, se lleva a cabo un análisis de umbrales. Para esto se estiman efectos heterogéneos (ver ecuación 1), primero por el puntaje obtenido en la subescala de *Interacción* del ITERS y, después, de acuerdo con el puntaje obtenido en el dominio de *Apoyo emocional y conductual* del CLASS.³⁸

$$PuntajeLV_i = \beta_0 + \beta_1 Puntaje_i + \beta_2 Puntaje \times Alto_i + \epsilon_i \quad (1)$$

Donde, *PuntajeLV* es el puntaje obtenido por el EPI *i* en la LV corta. *Puntaje*: es el puntaje obtenido por el EPI en *Interacción* o *Apoyo emocional y conductual*. *Alto*: es una variable dicotómica igual a 1 si *Puntaje* es mayor o igual a un valor *x*, para $x \in [1, 7)$.

En las Figuras G.1–G.4 se presentan los resultados de este análisis. En general, las correlaciones que son estadísticamente significativas entre los constructos de la LV y la subescala de *Interacción* del ITERS utilizando la muestra completa, se mantienen para aulas con niveles de calidad de mínimo a bueno.³⁹ Con relación al CLASS, la significancia de las correlaciones no se mantiene para aulas con niveles de calidad medio-alto o superior. En particular, para los constructos *Involucramiento* e *Interacción* de la LV se observa que, si bien la magnitud de las correlaciones se mantiene, los intervalos de confianza se amplían. Para los constructos de *Organización*, *Escuchar y hablar* y *Estructura del programa* se observa tanto un aumento en el tamaño de los intervalos de confianza como una disminución en la magnitud de las correlaciones. Luego, no está claro si la significancia de las correlaciones desaparece porque la LV corta no está logrando discriminar aulas con niveles de calidad altos, o si es por que se pierde precisión por falta de observaciones para este rango de la distribución. Es importante repetir el análisis en otros estudios similares para tener evidencia más sólida al respecto.

³⁸No se realiza el análisis para la otra subescala del ITERS ni el otro dominio del CLASS pues en estos casos los puntajes están concentrados en los niveles bajos de la distribución.

³⁹*Organización* es el único constructo de la LV Corta donde no se observa este comportamiento.



Figura G.1: Asociación entre los constructos de la LV Corta e ITERS: Interacción. I

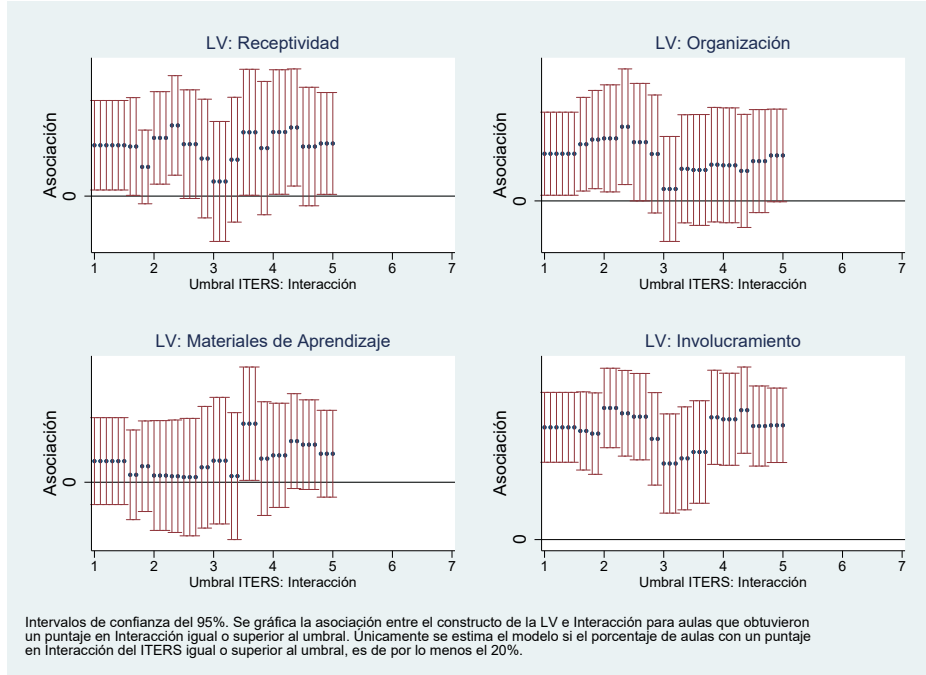


Figura G.2: Asociación entre los constructos de la LV Corta e ITERS: Interacción. II

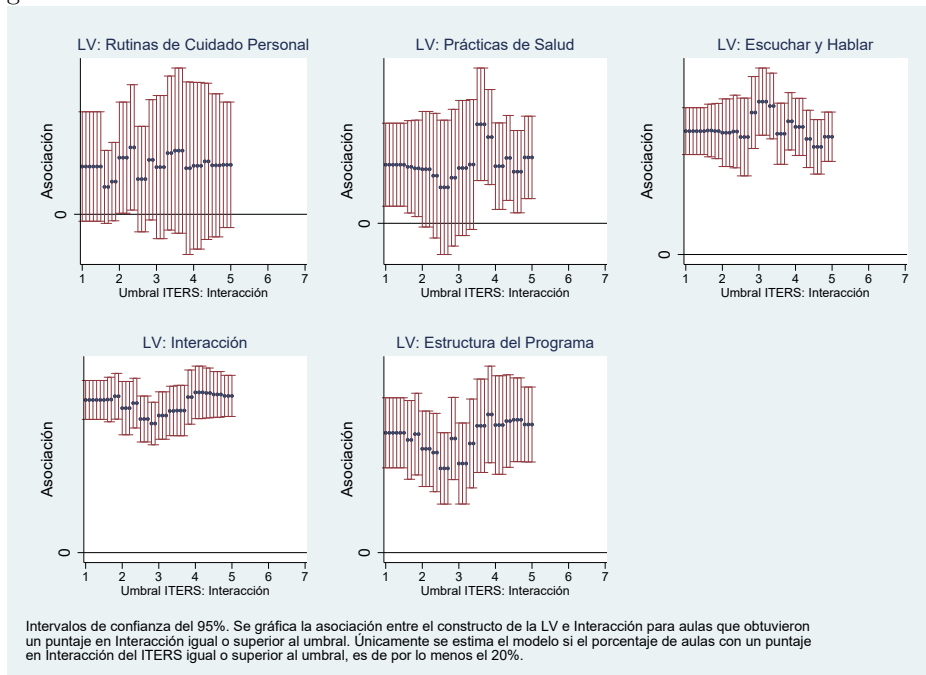




Figura G.3: Asociación entre los constructos de la LV Corta y CLASS: Apoyo Emocional. I

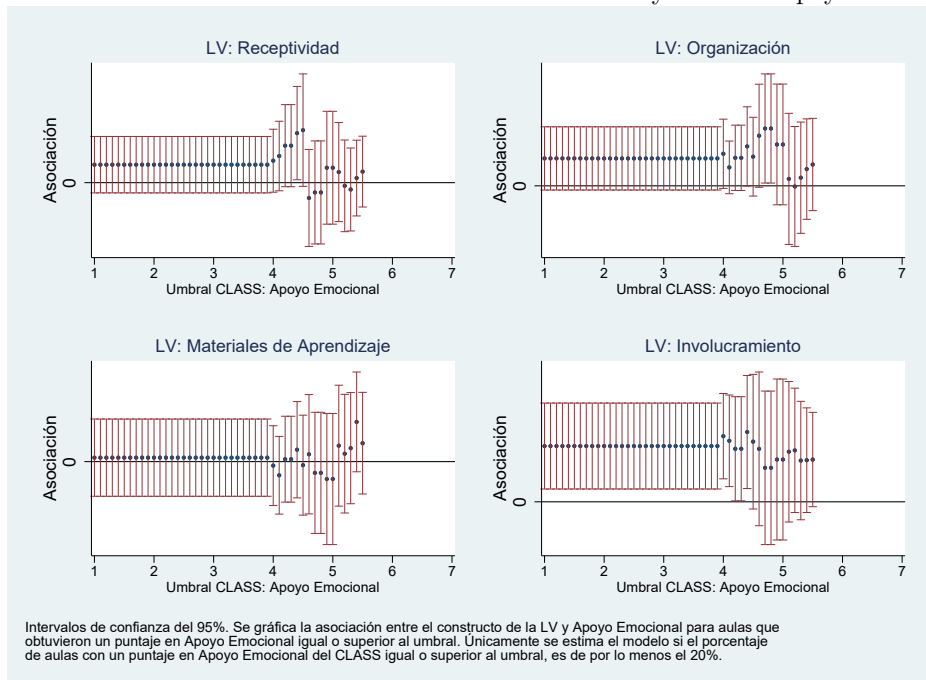


Figura G.4: Asociación entre los constructos de la LV Corta y CLASS: Apoyo Emocional. II

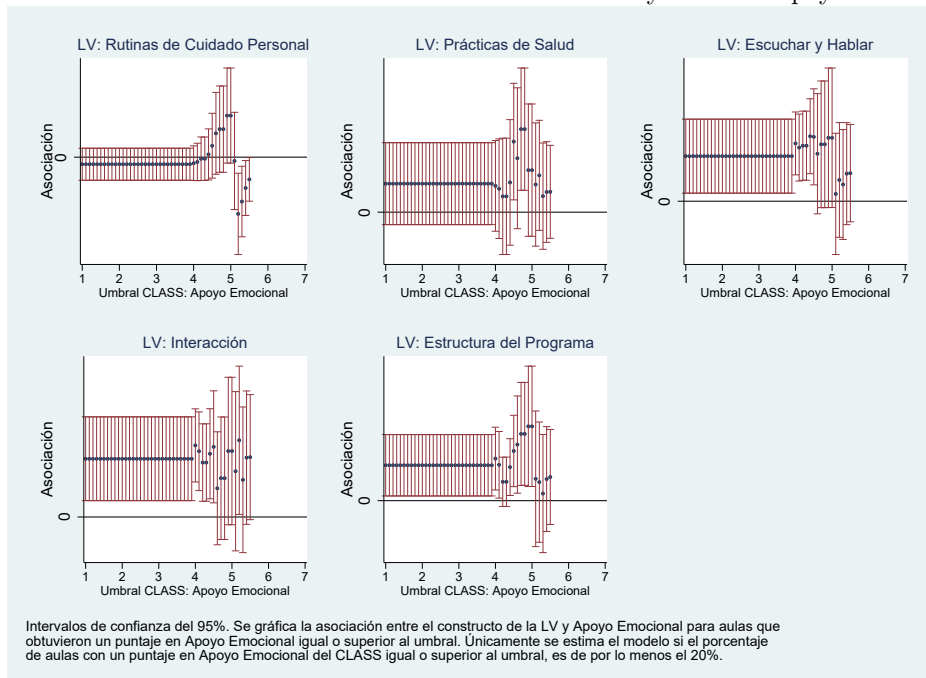


Tabla G.4: Correlación Bivariada entre Constructos de la LV corta y subescalas CLASS e ITERS, utilizando datos de Ecuador

	1. Dominio: Apoyo emocional	2. Dominio: Apoyo comprometido	3. Total CLASS	4. Subescala: Escuchar y hablar	5. Subescala: Interacción	6. Total ITERS
1. Receptividad (2 ítems)	0,21***	0,16***	0,20***	0,24***	0,40***	0,36***
2. Organización (1 ítem)	0,04	0,03	0,04	0,17***	0,32***	0,28***
3. Materiales de Aprendizaje (2 ítems)	0,13**	0,10**	0,12**	0,34***	0,37***	0,39***
4. Involucramiento (4 ítems)	0,24***	0,25***	0,26***	0,40***	0,51***	0,50***
5. Rutinas de Cuidado Personal (3 ítems)	0,23***	0,17***	0,22***	0,38***	0,46***	0,47***
6. Prácticas de Salud (2 ítems)	0,35***	0,29***	0,35***	0,43***	0,38***	0,43***
7. Escuchar y Hablar (3 ítems)	0,31***	0,30***	0,33***	0,77***	0,57***	0,70***
8. Interacción (11 ítems)	0,37***	0,31***	0,36***	0,64***	0,79***	0,79***
9. Estructura del Programa (4 ítems)	0,34***	0,27***	0,33***	0,54***	0,63***	0,64***
10. Total LV corta (32 ítems)	0,41***	0,35***	0,41***	0,73***	0,83***	0,86***

Nota: N=403 (columnas 1-3); 404 (columnas 4-6). Total ITERS es igual al promedio de los ítems de las subescalas Escuchar y hablar e Interacción. * p<0.10, ** p<0.05, *** p<0.01.



Bibliografía

- [1] Andrew, A., Attanasio, O., Bernal, R., Cordona, L., Krutikova, S., Martínez, D., Medina, C., Peña, X., Rubio-Codina, M., y Vera-Hernandez, M. (2018). *Evaluation of infant development centres: An early years intervention in Colombia*, 3ie Impact Evaluation Report 76. Nueva Delhi: International Initiative for Impact Evaluation (3ie).
- [2] Araujo, M.C., Dormal, M., y Schady, N. (2017). *Child Care Quality and Child Development*. J. Human Resources; publicado antes de su impresión en March 2, 2018, doi:10.3368/jhr.54.3.0217.8572R1.
- [3] Araujo, M.C., López Boo, F., Novella, R., Schodt, S., y Tomé, R. (2015). *La calidad de los Centros Infantiles del Buen Vivir en Ecuador*. Washington D.C.: Banco Interamericano de Desarrollo.
- [4] Araujo, M.C., López Boo, F., y Puyana, J.M. (2013). *Overview of Early Childhood Development Services in Latin America and the Caribbean*. Washington D.C.: Banco Interamericano de Desarrollo.
- [5] Baker M., Gruber J., y Milligan K. (2008). *Universal Childcare, Maternal Labor Supply, and Family Well-Being*. Journal of Political Economy. 116(4):709–45.
- [6] Baker M., Gruber J., y Milligan K. (2015). *Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program*. NBER Working Paper No. 21571.
- [7] Bayley, N. (2006). *Bayley Scales of Infant and Toddler Development: Bayley III*. Estados Unidos: Harcourt Assessment, Psych. Corporation.
- [8] Berlinski, S., y Schady, N. (2015). *The Early Years: Child Well-being and the Role of Public Policy*. Development in the Americas (DIA) 2015. Washington D.C.: IDB Flagship Publication.
- [9] Bernal, R., Attanasio, O., Peña, X. y Vera-Hernández, M. (2017). *The Effects of Transition from Home-based Childcare to Center-based Childcare in Colombia*. Manuscrito Universidad de los Andes y University College London.
- [10] Bradley, R. H., Caldwell, B. M., y Corwyn, R. F. (2003). *The Child Care HOME Inventories: Assessing the quality of family child care homes*. Early Childhood Research Quarterly, 18(3), 294-309.
- [11] Bryant D., Clifford, R.M., Early, D., Burchinal, M., Barbarin, O., Saluja, G., Pianta, B., y Howes, C.H. (2003). *Diversity and directions in state pre-kindergarten programs*. Recuperado de: <http://www.ncedl.org/>
- [12] Calmels, D. (2011). *Espacio Habitado. En la vida cotidiana y la práctica profesional*. Rosario, Argentina: HomoSapiens ediciones.
- [13] Clarke-Stewart, K., Gruber, C., y Fitzgerald, L. (1994). *Children at home and in Day Care*. Hillsdale, NJ: Erlbaum.
- [14] Dormal, M., y López Bóo, F. (2018). *Hacia el desarrollo de un instrumento de medición de calidad de procesos para centros de cuidado en América Latina y el Caribe*. Manuscrito sin publicar.

- [15] FLACSO/UNICEF/CIDECC. (2019). *Informe sobre evaluación de efectos del Programa CPI: un estudio muestral*. Recuperado de: <https://www.unicef.org/argentina/informes/informe-evaluaci%C3%B3n-efectos-programa-cpi>
- [16] Furr, R.M., y Bacharach, V.R. (2013). *Psychometrics: An introduction*. Thousand Oaks, CA: Sage Publications, Inc.
- [17] Guttman, L. (1954). *Some necessary conditions for common factor analysis*. *Psychometrika*, XIX, 149-61.
- [18] Hamre, B., y Pianta, R. (2007). *Learning opportunities in preschool and early elementary classrooms*. En R.C. Pianta, M.J. Cox y K.L. Snows, Eds. *School readiness & the transition to kindergarten in the era of accountability*. 49-83. Baltimore, MD: Paul H. Brookes Publishing Co.
- [19] Harms, T., Cryer, D., y Clifford, R.M. (2006). *Infant/Toddler Environment Rating Scale-Revised Edition Manual*. New York, NY: Teachers College Press.
- [20] Howes, C., Phillips, D., y Whitebook, M. (1992). *Thresholds of quality: Implications for the social development of children in center-based child care*. *Child Development*. 63: 449-460.
- [21] Kagan, S.L. (2010). Presentación *Promoting ECD Quality: Policy Lessons for Central America*, en Banco Interamericano de Desarrollo, noviembre de 2010, El Salvador.
- [22] La Paro, K., Pianta, R., y Stuhlman, M. (2004). *The classroom assessment scoring system: Findings from the prekindergarten year*. *Elementary School Journal*. 104: 409-426.
- [23] Law, J., Rush, R., Schoon, I., y Parsons, S. (2009). *Modeling Developmental Language Difficulties From School Entry Into Adulthood: Literacy, Mental Health, and Employment Outcomes*. *Journal of Speech, Language, and Hearing Research*, 52(6), 1401-1416.
- [24] Lejarraga, H., Pascucci, M.C., Masautis, A., Kelmansky, D., Lejarraga, C., Charrúa, G., Insua, I., y Nunes, F. (2014). *Desarrollo psicomotor infantil en la Cuenca Matanza-Riachuelo. Pesquisa de problemas inaparentes del desarrollo*. *Rev Argent Salud Pública*; 5(19):17-24.
- [25] Lejarraga, H., Berardi, C., Ortale, S., Contreras, M.M., Sanjurjo, A., Lejarraga, C., Martínez Cáceres, M.J., y Rodríguez, L. (2011). *Crecimiento, desarrollo, integración social y prácticas de crianza en niños que viven con sus madres en prisión*. *Archivos Argentinos de Pediatría*, 109(6), 485-491.
- [26] Lemay, L., Lehrer, J., y Naud, M. (2017). *Le CLASS pour mesurer la qualité des interactions en contextes culturels variés*. *Les Dossiers des Sciences de l'Éducation*, 37, 15-34.
- [27] López Bóo, F., Dormal, M., y Weber, A. (2019). *Validity of four measures of child care quality in a national sample of centers in Ecuador*. *PLoS ONE*, 14(2).
- [28] López Bóo, F., Araujo, M. C., y Tomé, R. (2016) *¿Cómo se mide la calidad de los servicios de cuidado infantil? Guía de herramientas*. Washington D.C.: Banco Interamericano de Desarrollo.
- [29] Mashburn, A.J., Pianta, R.C., Hamre, B.K., Downer, J.T., Barbarin, O.A., Bryant, D., Burchinal, M., Early, D.M., Howes, C. (2008). *Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills*. *Child Dev*; May-Jun 79(3):732-49.

- [30] Molina, E., Farwa, S., Ho, A., Melo, C., Wilichowski, T., y Pushparatnam, A. (2018). *Measuring Teaching Practices at Scale: Results from the Development and Validation of the Teach Classroom Observational Tool*. Policy Research Working Paper 8653.
- [31] NICHD Early Child Care Research Network. (2000). *Characteristics and quality of child care for toddlers and preschoolers*. Applied Developmental Science. 4: 116-135.
- [32] NIEER. (2012). *The aeioTU Early Childhood Longitudinal Study. Report I. Baseline Data Collection*. New Brunswick, NJ: National Institute for Early Education.
- [33] Ortiz, J., Corbo, G., Barcia, A., Russi, M., Hiriart, V., Estevez, G., Telles, A., Bruzzone, C., Rubio, I., Pérez, C., Bordoli, C., Ravazzani, S., y Sánchez, C. (2018). *Instrumento para la evaluación de la calidad de los servicios para niños de cero a tres años*. IECS-03. No publicado.
- [34] Pearson. (2018). *Bayley-III, Escalas Bayley de desarrollo infantil-III*. Recuperado de <http://www.pearsonclinical.es/producto/120/bayley-iii-escalas-bayley-de-desarrollo-infantil-iii>
- [35] Rosero, J., y Osterbeek, H. (2011). *Trade-offs between different early childhood interventions: Evidence from Ecuador*. Tinbergen Institute Discussion Paper 11-102/3.
- [36] Rubio-Codina, M., Araujo, M.C., Attanasio, O., y Grantham-McGregor, S. (2016). *Concurrent Validity and Feasibility of Short Tests Currently Used to Measure Early Childhood Development in Large Scale Studies: Methodology and Results*. Washington D.C.: IDB WORKING PAPER SERIES N° IDB-WP-723.
- [37] Schady, N. (2012). *El desarrollo infantil temprano en América Latina y el Caribe: acceso, resultados y evidencia longitudinal de Ecuador*. En M. Cabrol y M Székely.(Ed.), Educación para la transformación (pp. 53-92). Banco Interamericano de Desarrollo.
- [38] Schady, N., Behrman, J., Araujo, M.C., Azuero, R., Bernal, R., Bravo, D., López-Boo, F., Mancours, K., Marshall, D., Paxson, C., y Vakis, R. (2014). *Wealth Gradients in Early Childhood Cognitive Development in Five Latin American Countries*. Washington D.C.: IDB WORKING PAPER SERIES N° IDB-WP-482.
- [39] Schoon, I., Parsons, S., Rush, R., y Law, R. (2010). *Children's Language Ability and Psychosocial Development: A 29-Year Follow-up Study*. Pediatrics, 126(1), e73-e80.
- [40] SENAF/BID. (2019). *Manual de Administración MIRAr*, mimeo Ministerio de Desarrollo Social, Buenos Aires, Argentina.
- [41] SIEMPRO/INDEC. (2010). *Cálculos de EPIs por provincia*, SIEMPRO, sobre la base INDEC, Censo Nacional de Población, Hogares y Vivienda.
- [42] Slot, P. (2018). *Structural characteristics and process quality in early childhood education and care: A literature review*. OECD Education Working Paper No. 176
- [43] Vásquez, A., Andrade, V., Schiappacasse, P., y Díaz-Simón, N. (2018). *Inventario de caracterización de ambientes para educación inicial (INCA-EI): Manual de Aplicación*. No publicado.

