

¿CÓMO SE MIDE LA CALIDAD DE LOS SERVICIOS DE CUIDADO INFANTIL?

Guía de herramientas

FEBRERO 2016

Florencia López Boo
María Caridad Araujo
Romina Tomé



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

López Boo, Florencia.

¿Cómo se mide la calidad de los servicios de cuidado infantil? / Florencia López Boo, María Caridad Araujo, Romina Tomé.

p. cm.

Incluye referencias bibliográficas.

978-1-59782-220-6 (Rústica)

978-1-59782-221-3 (Digital)

1. Child development. 2. Child care. 3. Nursery schools. I. Araujo, Maria Caridad. II. Tomé, Romina. III. Banco Interamericano de Desarrollo. División de Protección Social y Salud. IV. Título. V. Serie.

HQ778.5 .L67 2015

IDB-BK-147

Diseño editorial: Círculo Salvo Comunicación | circulosalvo.com

Palabras claves: Desarrollo infantil, guarderías, jardines infantiles, salas cuna, centros de

Códigos JEL: I10, I20, I30, I38, J13.

Copyright © 2016 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.



Banco Interamericano de Desarrollo
División de Protección Social y Salud

¿CÓMO SE
MIDE LA
**CALIDAD DE
LOS SERVICIOS
DE CUIDADO
INFANTIL?**

Guía de herramientas

FEBRERO **2016**

scl-sph@iadb.org
www.iadb.org/ProteccionSocial



TABLA DE CONTENIDO

PREFACIO	06
AGRADECIMIENTOS	08
GUÍA DE USO	09
¿A quién está dirigida esta guía?	09
¿Cómo utilizar esta guía de herramientas?	09
CAPÍTULO 1: ¿POR QUÉ ES FUNDAMENTAL LA CALIDAD DE LOS SERVICIOS DE CUIDADO INFANTIL?	10
1.1 Introducción	11
1.2 ¿A qué nos referimos cuando hablamos de “calidad”? Dimensiones críticas para la atención de bebés y niños pequeños	12
1.3 La importancia de la calidad en los servicios de cuidado y su medición	15
La relación entre los servicios de cuidado de calidad y el desarrollo infantil en países desarrollados	15
La calidad de los servicios de cuidado en América Latina y el Caribe	17
CAPÍTULO 2: INTRODUCCIÓN A LA MEDICIÓN DE LA CALIDAD DE LOS SERVICIOS DE CUIDADO	20
2.1 Introducción	21
2.2 Variables estructurales	21
¿Qué son?	21
¿Cómo se miden?	21
¿Qué dice la literatura?	21
2.3 Variables de proceso	23
¿Qué son?	23
¿Cómo se miden?	23
¿Qué dice la literatura?	24
2.4 ¿Cómo miden los gobiernos la calidad de los servicios de cuidado?	25
CAPÍTULO 3: DEFINICIONES CLAVES PARA LA MEDICIÓN	26
3.1 Introducción	27
3.2 ¿Cuál es el objetivo de la medición de la calidad del centro de cuidado?	27
3.3 ¿Cuál es el grupo de estudio?	27
3.4 ¿Cómo medir?	28

Observación vs. Reporte	29
Filmación	31
3.5 ¿Qué instrumentos existen?	32
3.5.1 ECERS-R e ITERS-R	33
Descripción	33
Experiencias de aplicación	34
3.5.2 Toddler CLASS.....	35
Descripción	35
Experiencias de aplicación	36
3.5.3 ORCE	37
Descripción	37
Experiencias de aplicación	37
3.5.4 CIS	37
Descripción	37
Experiencias de aplicación	38
3.5.5 Perfil de evaluación	38
Descripción	38
Experiencias de aplicación	39
3.5.6 MITRCC	39
Descripción	39
Experiencias de aplicación	39
3.5.7 CC-IT-HOME	40
Descripción	40
Experiencias de aplicación	40
3.5.8 Otros instrumentos	40
3.6 ¿Qué encuestas complementarias se pueden aplicar?	42
3.7 ¿Cómo elegir el instrumento?	43
CAPÍTULO 4: IMPLEMENTACIÓN	46
4.1 Introducción	47
4.2 Capacitación	47
4.3 Piloto y adaptación cultural	48
4.4 Generación de números únicos de identificación entre bases de datos	49
4.5 Variables críticas del contexto	49
4.6 Confiabilidad entre observadores (<i>Inter-rater reliability</i>)	50
4.7 Seguimiento de protocolos	51
Protocolo general	51
Entrevista/observación	52
Filmación	53
Edición	54
Codificación	54
CAPÍTULO 5: CONSIDERACIONES FINALES	56
TABLAS	58
REFERENCIAS BIBLIOGRÁFICAS	79
ANEXO 1. VALIDEZ DEL INSTRUMENTO PRE K CLASS	87

PREFACIO

Las brechas en el desarrollo se hacen evidentes desde los primeros años de vida. Es por esto que los temas de desarrollo infantil se han posicionado recientemente en la agenda de la mayoría de los países de América Latina y el Caribe.

Los gobiernos han enfocado sus esfuerzos principalmente en la expansión de la cobertura de servicios de cuidado para niños pequeños (centros de cuidado —muchas veces llamados guarderías, salas cuna, jardines infantiles y *kínderes*). Actualmente, los países de la región enfrentan el desafío de asegurar la calidad de los servicios de cuidado que ofrecen, dado que los niños que acuden a ellos lo hacen durante sus primeros años de vida, un período crítico para el desarrollo de las capacidades. La inversión en servicios de cuidado de calidad es una política social que favorece la equidad promoviendo oportunidades de desarrollo desde el inicio de la vida. Trabajar enfrentando este reto requiere inicialmente entender qué es un servicio de cuidado de calidad y qué se requiere para ofrecerlo.

Si bien no existe una receta única para asegurar la calidad de los centros de cuidado

infantil, sabemos que ésta abarca una variedad de elementos, desde la provisión de una alimentación nutritiva, el tamaño del grupo de niños a cargo de cada cuidador, el salario y la educación de los cuidadores, el sistema de entrenamiento continuo para los cuidadores, hasta elementos más difíciles de medir como la calidad de las interacciones entre los cuidadores y los niños, por nombrar algunos.

Se podría pensar que un servicio de calidad promedio podría ser suficiente para las familias pobres, pues sus alternativas de cuidado son muy precarias. Por ejemplo, en contextos de familias muy vulnerables es práctica común dejar al niño solo en el hogar, encargarlo al cuidado de otro niño o llevarlo al lugar de trabajo de la madre en donde puede estar expuesto a riesgos y no recibe ninguna atención. En este sentido, la literatura concluye que la calidad de la atención que reciben los niños que asisten a un servicio de cuidado, en particular para los niños de hogares pobres o en situación de riesgo, es fundamental porque se pueden satisfacer sus necesidades de alimentación, de estimulación, de condiciones higiénicas, y de afecto que tal vez no se satisfacen en

el hogar. De esta manera, la inversión en servicios de cuidado infantil de calidad dota a estos niños de mejores cimientos en las dimensiones críticas para su desarrollo durante la primera infancia.

En nuestra región, al igual que en otras, no hay evidencia suficiente acerca de las acciones específicas que son críticas y costo-efectivas para mejorar la calidad de los servicios de cuidado. A diferencia de sectores como el de educación primaria, en donde la región ha invertido esfuerzos importantes para contar con indicadores sobre la calidad de los maestros (por ejemplo la herramienta “Stallings classroom snapshot”), los recursos en las escuelas (censos escolares) y los niveles de aprendizaje en las aulas (por ejemplo, con la implementación del Programa para la Evaluación Internacional de Alumnos, PISA, por sus siglas en inglés), los gobiernos cuentan con poca información sistematizada y confiable sobre los servicios de cuidado.

Contribuye a esta carencia la falta de un uso regular de instrumentos que midan una variedad de aspectos de la calidad de los centros de cuidado infantil y de las experiencias que los niños tienen en ellos. Esta información es clave para identificar los procesos que contribuyen a mejorar la calidad de los centros de forma costo-efectiva y, asimismo, para definir y monitorear el cumplimiento de un conjunto de estándares de calidad en los centros, tanto en aquellos que operan con fondos públicos como en los de provisión privada. Más importante aún, el énfasis en los indicadores de desarrollo de los niños parece estar aumentando en la agenda de desarrollo posterior a la revisión de los avances en los Objetivos del Milenio en 2015. Contar con datos comparables en cuanto a la calidad de centros de cuidado contribuirá entonces a tener un mecanismo cuantificable de rendición de cuenta sobre las inversiones realizadas por los países.

Para contribuir a llenar este vacío, y en un esfuerzo por operacionalizar los resultados principales de la publicación insignia del BID del 2015 “Los primeros años” (Berlinski y Schady, 2015), la División de Protección Social y Salud presenta esta herramienta de fácil uso que aborda en mayor detalle la definición de un servicio de calidad, provee un menú de los instrumentos disponibles para la medición de la calidad de los centros de cuidado infantil que atienden a bebés y niños pequeños de entre 0 y 3 años (36 meses) de edad y también informa sobre el proceso de implementación de dichos instrumentos en la región. En otras palabras, presenta una descripción teórica de los instrumentos y una guía sobre dónde, cómo y cuándo usar cada herramienta, a partir de un abordaje detallado de diferentes dimensiones a considerar para que la medición de la calidad sea exitosa.

Este libro tiene varios aportes originales. Primero, sintetiza en una sola publicación un enfoque tanto teórico como práctico sobre la medición de la calidad de los centros. Segundo, discute en detalle los aspectos importantes a considerar durante la implementación de la medición. Tercero, sistematiza la información sobre los instrumentos para que el lector pueda realizar comparaciones entre sus diversas dimensiones, costos y beneficios de forma sencilla. Por último, se presenta en español con el objeto de facilitar el acceso de los lectores de América Latina.

Esta publicación servirá como una guía para los investigadores y profesionales interesados en traducir la discusión sobre la mejora de la calidad de los servicios de cuidado infantil en acciones y resultados concretos.

Estoy seguro que disfrutarán de esta lectura y que les será de utilidad.

Ferdinando Regalia

Jefe de la División de Protección Social y Salud
Banco Interamericano de Desarrollo

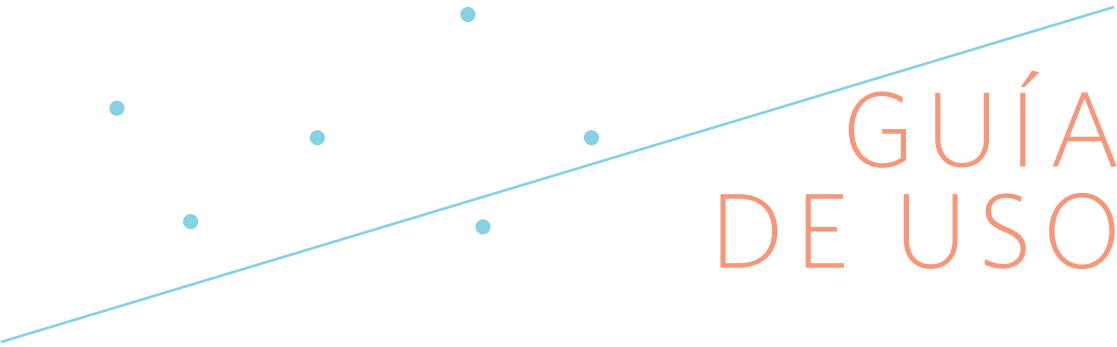


AGRADECIMIENTOS

Primero, queremos destacar los valiosos aportes de los participantes del comité asesor del BID para temas de desarrollo infantil que se reunió en enero de 2012 con el propósito de apoyar a nuestro equipo en el diseño de una agenda aplicada de investigación en el tema de la medición de la calidad en centros de cuidado: Steve Barnett, Raquel Bernal, Margaret Burchinal, Patrice Engle, Sally Grantham McGregor, Jennifer LoCasale Crouch, Milagros Nores, Deborah Phillips, Maureen Samms Vaughan, Sergio Urzúa, Hiro Yoshikawa y Martha Zaslow.

En segundo lugar, queremos agradecer a un revisor anónimo por sus comentarios invaluable, también a Sara Schodt y Milagros Nores por continuos aportes durante el transcurso de la escritura de este documento, así como por leer este documento en su versión final y proveer comentarios. En tercer lugar, queremos agradecer a la empresa Habitus de Ecuador con la cual concretamos nuestro proyecto de medición de calidad en Ecuador. También queremos agradecer la edición de Guillermo Martínez.

Por último, reconocemos con gratitud el sostén institucional recibido de Ferdinando Regalia, Jefe de la División de Protección Social y Salud del BID, Norbert Schady, Asesor Económico Principal para el Sector Social, así como de Héctor Salazar, Gerente del Sector Social, quienes nos apoyaron en sacar adelante este proyecto.



GUÍA DE USO

¿A quién está dirigida esta guía?

Esta guía de herramientas está diseñada para ser un recurso para investigadores y personal técnico de cualquier disciplina, pertenecientes a gobiernos e instituciones interesadas en medir y monitorear la calidad de los servicios de cuidado infantil que atienden a bebés y niños pequeños de entre 0 y 3 años (36 meses) de edad.¹


¿Cómo utilizar esta guía de herramientas?

Esta guía de herramientas está organizada en cinco capítulos que, en forma secuencial, presentan al lector evidencia sobre la importancia de medir la calidad de los servicios de cuidado, y luego lo guían en los pasos a seguir para implementar la medición. Específicamente:

- el **Capítulo 1** explica **qué** es la calidad de los servicios de cuidado y **por qué** debería implementarse su medición;
- el **Capítulo 2** especifica **qué medir** para evaluar la calidad de los servicios de cuidado;
- el **Capítulo 3** discute las **decisiones previas** que deben considerarse al implementar la medición y describe los diferentes tipos de evaluaciones y de instrumentos;
- el **Capítulo 4** reflexiona alrededor de las **dimensiones de la implementación** de la medición;
- el **Capítulo 5** resume algunas **conclusiones y recomendaciones** para quienes vayan a aplicar una medición de la calidad de servicios de cuidado de bebés y niños pequeños.

¹A lo largo de esta guía se utiliza “bebés” y “niños pequeños” para referirse a aquellos menores de 36 meses de edad.

¿POR QUÉ ES FUNDAMENTAL LA CALIDAD DE LOS SERVICIOS DE CUIDADO INFANTIL?



● CAPÍTULO 1



1.1 / Introducción

Cada vez con mayor frecuencia, los países optan por invertir en el desarrollo infantil temprano —focalizando sus esfuerzos hacia los niños de hogares más pobres— como una estrategia para mejorar los resultados escolares de estos niños y su trayectoria durante la vida adulta (Engle et al., 2007; Engle et al., 2011). Concretamente, los niños que crecen en entornos pobres están expuestos a múltiples factores de riesgo, tales como la mala nutrición, condiciones de salubridad precarias, entornos de aprendizaje poco estimulantes, niveles muy altos y constantes de estrés, y otros riesgos ambientales. Todos estos factores afectan su desarrollo cognitivo, socioemocional y físico, y frenan el desarrollo de su potencial tanto en la escuela como en la edad adulta (Fernald, Gertler y Neufeld, 2009; Grantham-McGregor et al., 2007). En América Latina y el Caribe (ALC), se ha documentado que los niños que viven en la pobreza, en promedio, alcanzan niveles más bajos de desarrollo

cognitivo y abandonan la escuela con mayor frecuencia que los demás niños, lo que los condena, de por vida, a trabajos mal pagados y a replicar el ciclo de pobreza en el cual se criaron (Rolnick y Grunewald, 2007; Heckman, 2006; PREAL, 2006).

Las intervenciones de política durante la primera infancia son prometedoras, dado que tienen el potencial de paliar algunos de estos efectos de las privaciones tempranas. Existe evidencia que sugiere que, en el caso de los centros de cuidado infantil, la situación de los niños más vulnerables tiene mucho potencial de mejora, siempre y cuando se provean servicios de calidad. Por esta razón, es fundamental que los gobiernos cuenten con instrumentos que les permitan medir la calidad de la atención que brindan los centros, enfocándose en aquellas dimensiones de la calidad que son críticas en su relación con resultados positivos en el desarrollo de los niños (Martinez-Beck en Zaslow et al., 2011; Peisner-Feinberg et al., 2001; Vandell y Wolfe, 2000; Lamb, 1996, 1998; Phillips, 1987 y Clarke-Stewart y Fein, 1983).

La necesidad de proveer servicios de calidad es crucial en el rango etario de los bebés y



niños pequeños, es decir, aquellos de entre 0 y 3 años (36 meses) de edad. Esta afirmación se sustenta tanto en la evidencia biológica (por ejemplo, la ventana de oportunidad de los primeros 1000 días de vida desde el embarazo para intervenciones nutricionales) como en la literatura de psicología infantil (por ejemplo, los períodos sensibles para el desarrollo del apego o del lenguaje). Es notable la poca información existente sobre las diferencias en los parámetros de calidad que son críticos para el desarrollo adecuado de bebés y niños pequeños, comparados con aquellos que se requieren para niños en edad preescolar (Barros et al., 2011). Las variables de proceso², es decir, aquellas que caracterizan la calidad de las interacciones entre los niños y los adultos a su cargo y las rutinas de cuidado, se vuelven aún más importantes para este grupo etario. Esto es porque si bien los niños pequeños requieren de contenido menos estructurado, ellos demandan atenciones más individualizadas que los niños mayores.

Por otro lado, los instrumentos de medición de la calidad desarrollados a nivel internacional para bebés y niños pequeños cubren distintos rangos de edades, lo que dificulta aún más una medición sistemática y comparable a través del tiempo, incluso dentro de una misma “familia” de instrumentos. Por ejemplo, el *Classroom Assessment Scoring System* (CLASS, por sus siglas en inglés) tiene dos versiones, una para bebés que se aplica hasta los 18 meses de edad y otra para niños pequeños que se utiliza entre los 15 y 36 meses. El *Infant/Toddler Environment Rating Scale – Revised* (ITERS-R, por sus siglas en inglés) se aplica a niños entre los 0 y 30 meses y el *Early Childhood Environment Rating Scale – Revised* (ECERS-R, por sus siglas en inglés) se aplica a niños desde los 30 hasta los 60 meses de edad. El *Child Care Infant/*

Toddler HOME (CC-IT-HOME, por sus siglas en inglés) y el *Missouri Infant/Toddler Responsive Caregiving Checklist* (MITRCC, por sus siglas en inglés) se aplican en niños de entre 0 y 36 meses (y el primero tiene también una versión para niños en edad preescolar). Todos estos instrumentos se explicarán con mayor detalle en la Sección 3.5. Las complicaciones para identificar y también para medir la calidad pueden ser uno de los tantos obstáculos que enfrentan los servicios de cuidado que atienden a bebés y niños pequeños en su afán de garantizarla.

1.2 / ¿A qué nos referimos cuando hablamos de “calidad”?

Dimensiones críticas para la atención de bebés y niños pequeños

Parte del desafío de garantizar una atención de alta calidad en un servicio de cuidado es que aparentemente no hay una receta única para lograrla. Especialmente en la primera infancia, el consenso general es que la calidad debe ser integral, es decir, no centrarse únicamente en la educación o la salud, sino también en el desarrollo cognitivo y socioemocional, la nutrición y la crianza (Halle, Whittaker y Anderson, 2010; Zill et al., 2003; NICHD, 1996). Kagan (2010) identifica que las áreas críticas en las cuales hay que evaluar la calidad de los servicios de cuidado incluye: las relaciones que el niño tiene con otros adultos y con otros niños, el plan de actividades y aprendizaje, la enseñanza, la evaluación del progreso del niño, la atención a la salud, el perfil de los adultos a cargo de la atención de los niños, el contexto familiar, las relaciones comunitarias, el entorno físico, el liderazgo de quien coordina el centro de

² En el Capítulo 2 se analizan detalladamente los dos tipos de dimensiones: de procesos y estructurales. Las primeras observan la calidad de las interacciones entre los niños y sus cuidadores, y entre unos y otros, así como las actividades en las que se involucran mientras están en el centro de cuidado. Las variables estructurales en general identifican los recursos que facilitan las interacciones: tamaño de los grupos, educación, experiencia y salario del cuidador, infraestructura y seguridad, currículo y materiales.

cuidado, y los aspectos relacionados con la gestión del centro.

A pesar de este consenso (NAEYC, 2012; NICHD, 2000a,b), no existe una manera única y universal que describa cómo se deben combinar o evaluar estos elementos —que, en forma conjunta, constituyen la calidad— ni qué ponderación asignarle a cada uno. Por lo tanto, se vuelve complejo medir y definir la calidad de los centros de cuidado infantil. De los estudios internacionales sobre la oferta de cuidado infantil durante la primera infancia se desprende que los gobiernos tienen la capacidad de desempeñar un papel importante para garantizar la calidad a través de su función de control y rectoría. Por ejemplo, tanto en los Estados Unidos como en Europa, aquellos estados o países con estándares de concesión de licencias más estrictos y aquellos que establecen sistemas de acreditación para sus centros además de monitoreo y gestión de la calidad, son aquellos que cuentan con una calidad más alta. En Estados Unidos, se ha encontrado que los centros que logran superar los estándares fijados por el estado también tienden a proporcionar un cuidado de mayor calidad a los niños (Fundación Van Leer, 2011; Helburn, 1995). Howes et al. (1992) encuentran que los centros de cuidado que regulan el coeficiente

de atención (es decir, el número de niños por adulto) así como el tamaño del grupo tienen una mejor calidad del cuidado de los niños. Además observaron que cuando en un aula se encuentra un niño más de lo que determinan las normas estatales de coeficientes de atención óptimos, esto se asocia con una atención de menor calidad.

De acuerdo a la opinión de los expertos consultados por el BID (Barnett, 2012 y BID Advisory Committee, 2012) los seis elementos críticos que deberían definir la calidad en centros de cuidado para el grupo de 0-3 años de edad en ALC son los siguientes:

1. La provisión de una alimentación nutritiva, en condiciones óptimas de higiene, saneamiento y seguridad;
2. La cantidad de niños a cargo de cada cuidador, que no debería ser mayor a seis niños por adulto en el grupo de edad de 12-36 meses y que debería ser aún menor para los bebés;
3. La calidad de las interacciones entre los cuidadores y los niños, su frecuencia e intensidad;
4. La existencia de un sistema que monitorea la calidad con regularidad en todos los centros;



5. Los esfuerzos de capacitación y desarrollo profesional para los cuidadores y maestros;

6. Actividades, materiales de juego y espacios estimulantes.

¿Por qué estas dimensiones críticas podrían diferir de aquellas relevantes a la atención de los niños de más de 3 años? Primero, porque el sistema inmunitario de los bebés y niños pequeños se está recién desarrollando. El cuidado en grupo los expone entonces a mayores riesgos. Además, la literatura especializada documenta que los indicadores de proceso, tales como la capacidad de respuesta y de interacción de los cuidadores hacia los niños, tienen un papel aún más importante en el desarrollo de los bebés y niños pequeños. Esto se debe a que como los bebés y niños pequeños tienen una capacidad limitada para iniciar interacciones con sus cuidadores, con otros niños o con su entorno antes de tener movilidad completa (Howes et al., 1992) dependen más del inicio de la interacción de parte del cuidador. Otra literatura se basa en las teorías de apego en la infancia muy temprana (Bowlby, 1969). Al igual que para los niños mayores, las variables de proceso están más asociadas al desarrollo de los bebés y niños pequeños que las variables estructurales; aunque es importante notar que una mejor infraestructura también permite mejores procesos.

Las variables estructurales se miden con mayor frecuencia en entornos escolares y preescolares. Por ejemplo, el tipo de actividades previstas para que bebés y niños pequeños desarrollen comportamientos que les proporcionen una base cognitiva, social y emocional sólida en la vida y que, con el tiempo, los preparen mejor para la escuela. Sin embargo, dicho desarrollo depende mucho más de la calidad de sus interacciones físicas y verbales con sus cuidadores de lo

que depende en niños mayores (Gonzalez-Mena y Widmeyer Eyer, 2007; Raikes y Edwards, 2009). Esto está también relacionado con el hecho de que los niños más pequeños necesitan una atención más individualizada.

Es así que, si bien los seis elementos críticos mencionados para definir la calidad en centros de cuidado pueden ser relevantes para los niños de entre 0 y 5 años, el segundo y el tercer punto parecen ser mucho más relevantes para el grupo de 0 a 3 años por dos razones. Primero, porque se requieren menores coeficientes de atención (menos niños por cuidador) para los niños más pequeños. Segundo, porque la frecuencia y calidad de las interacciones tienen un papel más importante en este grupo.

¿Cómo podrían diferir estas dimensiones críticas en países de diferentes regiones del mundo? A diferencia de muchos estudios realizados en los Estados Unidos, el análisis de Berlinski, Galiani y McEwan (2008) y de Berlinski, Galiani y Gertler (2009) en ALC sugiere que el mero acceso a la educación preescolar sí tiene un impacto positivo en los resultados en la escuela primaria para los niños (aunque estudios recientes no publicados de Chile parecen demostrar lo contrario). Dado que los estudios no midieron la calidad del preescolar, estos resultados son consistentes con dos posibles explicaciones: (i) que efectivamente los servicios de preescolar fueron de buena calidad o (ii) que para los niños de familias muy pobres, incluso los centros de cuidado que proveen atención de calidad baja pueden producir resultados positivos si el ambiente es mejor que el de su propio hogar.³ No se sugiere que los estándares de los centros de cuidado en ALC sean bajos. Por el contrario, se conoce que muchos de los niños más pobres de la región enfrentan situaciones de riesgo y vulnerabilidad en sus propios hogares. Por lo tanto esta realidad exige

³ El impacto de un servicio de cuidado sobre el desarrollo del niño está relacionado con las características de la alternativa de cuidado a la cual hubiera tenido acceso ese niño en la ausencia del servicio (el llamado contrafactual).

una comprensión precisa de la calidad de atención de los centros de cuidado. En este sentido, Halpern (1986) advierte sobre el peligro de generalizar los estudios de Estados Unidos a otros contextos, ya que un gran número de niños en ALC enfrenta situaciones mucho más sombrías en lo que se refiere a la calidad del cuidado en el hogar en ausencia del servicio del centro de cuidado. Tales niños viven en situación de pobreza y desnutrición, sin las redes de asistencia social que existen en los países desarrollados para ayudar a las familias. Estos desafíos sugieren que los programas de primera infancia en algunos países de ALC tienen la responsabilidad de abordar de manera integral el desarrollo de los niños; es decir, de tomar en cuenta la salud y las necesidades nutricionales, así como el desarrollo cognitivo y social. En consecuencia, los instrumentos para medir la calidad deben considerar también los procesos relacionados con la atención de la salud y la nutrición de los niños en los centros.

1.3 / La importancia de la calidad en los servicios de cuidado y su medición

La relación entre los servicios de cuidado de calidad y el desarrollo infantil en países desarrollados⁴

A pesar de que parte de la evidencia empírica demuestra que la asistencia a centros de cuidado tiene algunos beneficios, también hay otra evidencia igualmente convincente que sugiere que es clave asegurar la calidad del cuidado en los centros si se quiere lograr resultados en el desarrollo saludable de los niños (NICHD, 2000b). De hecho, se ha documentado una asociación negativa entre

el cuidado de baja calidad y el desarrollo infantil, en particular en los aspectos socioemocionales, tales como la agresividad del niño (Belsky, 1988; Howes et al., 1992). La pregunta es: ¿cuál es el problema? ¿los centros de cuidado en sí o los servicios de baja calidad que prestan? Si no se provee a los niños de un cuidado mejor o del mismo nivel de calidad que aquel que podrían recibir en su hogar, los efectos podrían ser realmente perjudiciales para su salud y su desarrollo. Varios estudios identifican resultados contrapuestos en la relación entre la calidad del cuidado y el desarrollo socioemocional y del comportamiento del niño. Mientras que Love et al. (2003) y el estudio del NICHD (2000a) encuentran resultados positivos entre el cuidado en centro y el desarrollo infantil, los dos estudios analizados en el próximo párrafo encuentran lo opuesto.

Un buen ejemplo de la importancia de evaluar la calidad de los centros de cuidado es el estudio realizado por Baker, Gruber y Milligan (2008) que estudia la expansión universal de los centros de cuidado subsidiados por el gobierno en Quebec para niños de hasta 5 años. Los autores encontraron que, en el corto plazo, la exposición prolongada a los cuidados en centros era nociva y, particularmente, empeoraba “la hiperactividad, la falta de atención, la agresividad, las habilidades motrices y sociales, la salud mental y la propensión a enfermedades en los niños”. En el seguimiento de largo plazo, los autores demuestran que los déficits no-cognitivos persistieron hasta la adolescencia y que los cohortes expuestos a los centros de Quebec tuvieron peores resultados de salud, menor satisfacción con su vida y mayores tasas de criminalidad (Baker, Gruber y Milligan, 2015). En la misma línea, Yamauchi y Leigh (2011) revisaron la relación entre los centros de cuidado y los efectos en el comportamiento de niños en edad

⁴ Esta sección se basa en la Nota Conceptual del Proyecto (Schodt, 2012).

preescolar de Australia. Encontraron que, en general, asistir a centros de cuidado a tiempo completo se asocia negativamente con el comportamiento de los niños, especialmente para aquellos provenientes de familias con buenos lazos afectivos o de estratos socioeconómicos altos. En ambos casos se observa que el servicio (considerado por los autores de alta calidad) resulta en peores indicadores de comportamiento de los niños pequeños. Sin embargo, ninguna de estas dos últimas referencias aborda explícitamente la medición de la calidad.

Amplia evidencia empírica sugiere que la calidad del cuidado infantil influye directamente sobre los resultados del niño en una serie de dimensiones. Específicamente, la participación en programas de calidad alta o bien tiene claros impactos positivos en el desarrollo, como fue demostrado con las evaluaciones experimentales de los proyectos “Abecedarian” (Campbell y Ramey, 1995), “Early Head Start” (Love et al., 2005) y “Perry Preschool” (Schweinhart et al., 2005); o bien está asociada con mejor salud y desarrollo socioemocional y cognitivo en el corto plazo (Vandell y Wolfe, 2000; Phillips, 1987; NICHD, 2000a,b; Shonkoff y Phillips, 2000; Burchinal et al., 1997). Mientras

que la mayoría de los estudios disponibles sólo toman en cuenta los resultados de los niños en el corto plazo, varios estudios reconocidos de Estados Unidos, como los proyectos “Abecedarian”, “High/Scopes Perry Preschool” y “Early Head Start” también han encontrado efectos positivos, significativos y de largo plazo en servicios de cuidado de calidad, bien diseñados e implementados. Entre los impactos de largo plazo identificados, se destacan menores índices de delincuencia, menos embarazos en adolescentes, menor dependencia de programas de asistencia social y menor uso de drogas, así como mayores salarios y tasas de empleo. Estos estudios evaluaron a niños de hogares vulnerables que recibieron cuidado de muy alta calidad durante la primera infancia (NAEYC, 2012). Si bien los estudios del “Abecedarian” y “High/Scopes Perry Preschool” contaban con muestras muy pequeñas, estos estudios hacen hincapié en la necesidad de contar con cuidado de calidad durante la primera infancia para mejorar los resultados de desarrollo del niño ya que las intervenciones de baja calidad no han demostrado tener una asociación positiva con los resultados del niño (Peisner-Feinberg et al., 2001). La evaluación de largo plazo del programa a escala de Quebec refuerza este importante mensaje.



La “dosis” de atención o la frecuencia, intensidad y duración de la exposición al servicio de cuidado también parece ser un aspecto importante de la calidad que claramente afecta los resultados de los niños. En un metaanálisis de 34 programas de preescolar en los Estados Unidos, Nelson et al. (2003) encontraron que los programas que ofrecen al menos 20 horas de cuidado a la semana, y que proporcionan más años de cuidado en centro, producen mayores impactos sobre el desarrollo cognitivo y el rendimiento académico, lo que sugiere beneficios acumulativos de los cuidados de alta calidad en niños pequeños.

En cualquier parte del mundo, la capacidad de los maestros y cuidadores de interactuar eficazmente con los niños pequeños depende, en primer lugar, del número de niños que tengan a su cuidado. Por lo tanto, se puede suponer razonablemente que los hallazgos de la literatura descritos en esta sección pueden aplicarse en ALC, donde las limitaciones de capital humano y de financiamiento se traducen en coeficientes de atención mayores a los que se observan en los Estados Unidos. Esto significa que, en ALC, es aún más urgente desarrollar medidas de calidad que vayan más allá de los aspectos estructurales y capten realmente la calidad de los procesos entre cuidadores y niños.

La calidad de los servicios de cuidado en América Latina y el Caribe

Aunque existen estudios experimentales en Estados Unidos y en otros países desarrollados, que exponen claramente los efectos positivos y de largo plazo de los programas de primera infancia de calidad sobre el progreso escolar, las aptitudes sociales y el logro profesional, lamentablemente esta evidencia longitudinal

no se ha recopilado en gran parte de ALC.⁵ Otros estudios realizados en ALC evalúan el impacto de la asistencia al preescolar (tales como Berlinski, Galiani y Gertler (2009) en Argentina; Filp y Schiefelbein (1982) en Argentina, Bolivia, Chile y Colombia; McKay y McKay (1983) en Cali, Colombia; y Filp et al. (1983) en Chile). Esta investigación también documenta resultados pequeños y otros medianamente positivos sobre la evolución de los niños en la primaria y sobre su desempeño en pruebas estandarizadas. Más recientemente, en Chile, Urzúa y Veramendi (2011) realizaron un seguimiento de la expansión significativa de la cobertura de los servicios de cuidado y encontraron que la participación en estos mejoró el desarrollo cognitivo de los niños mayores de 2 años.

Por otro lado, Arredondo et al. (2011) evaluaron con una metodología cuasi experimental un programa mexicano que se inició en 2007 llamado *Estancias infantiles para apoyar a madres trabajadoras*, diseñado para promover la participación laboral de las madres de niños de entre 1 y 4 años de edad, mediante subsidios que cubrieran el costo de servicios de cuidado fuera del hogar. El método de evaluación utilizó el proceso natural de registro de los niños en las estancias. De tal forma que una muestra de niños beneficiarios formaron el grupo de tratamiento y una muestra de niños inscritos en la lista de espera conformaron el grupo de control. Los autores encontraron que, mientras que los servicios de cuidado y los subsidios aumentaron la participación de las madres en la fuerza laboral, los niños que ingresaron en los centros no obtuvieron beneficios ni en su desarrollo cognitivo ni en su estado de salud.

En Brasil, Barros et al. (2011) analizaron datos sobre el costo y la calidad de los

⁵ Una excepción, sin embargo, constituye el trabajo de Grantham-McGregor y coautores (1991, 1997) en Jamaica, en el que se evaluó una intervención dirigida a padres que incluía visitas domiciliarias semanales por parte de auxiliares de salud comunitarias. Este programa demostró tener resultados positivos en una amplia gama de dimensiones sociales, de salud y educación hasta la edad de 22 años en aquellos niños cuyos padres se beneficiaron con esta intervención (Walker et al., 2011). Además se observaron resultados similares en los hijos de los niños intervenidos (Walker et al., 2012). Una adaptación de este piloto a Colombia también arroja resultados positivos, aunque no se demuestran efectos de los suplementos nutricionales (Attanasio et al., 2012).

servicios de cuidado, junto con indicadores de desarrollo y características familiares de niños en la ciudad de San Pablo.

Los autores encontraron que fueron las dimensiones de proceso, más que las dimensiones estructurales, las que tuvieron una asociación significativa con los resultados de desarrollo de los niños. También encontraron que la relación de costo-efectividad de mejorar la calidad de los cuidados fue casi de 60 a 1. No obstante, si el aumento en la calidad se centrara en mejorar las dimensiones de proceso que más se relacionan con el desarrollo de los niños, los autores calculan que se podría alcanzar una calidad con una relación de costo-efectividad casi 40 veces mayor.⁶ En cuanto a la “dosis” de atención o la frecuencia, intensidad y duración de la exposición al servicio de cuidado en Bolivia, Behrman, Cheng y Todd (2004) encontraron que los mayores beneficios se dan entre los niños que comienzan a asistir al centro de cuidado a los 2 años de edad y que participan por lo menos entre 12 y 24 meses, con rendimientos decrecientes después de esa edad.

Por último, Rosero y Oosterbeck (2011) comparan el impacto de dos modalidades diferentes de programas de desarrollo infantil que se implementan en Ecuador: las visitas domiciliarias y los centros de cuidado. Los autores del estudio encuentran que, si bien las visitas domiciliarias mejoraron el desarrollo cognitivo del niño y la salud mental de las madres, los centros de cuidado fueron perjudiciales en las dos dimensiones.

⁶ Estos resultados son similares a los del estudio de costos, calidad y resultados en centros de cuidado que determinó que el cuidado infantil es costoso en todo el espectro, pero que los costos de los servicios de calidad, aunque son mayores, lo son sólo marginalmente si se los compara con aquellos de servicios de mala calidad (Helburn, 1995).

•

•

•

•

•

•

•

•

•

•

•

•

INTRODUCCIÓN A LA MEDICIÓN DE LA CALIDAD DE LOS SERVICIOS DE CUIDADO



● CAPÍTULO 2



2.1 / Introducción

La medición de la calidad de los servicios de cuidado para bebés y niños pequeños es un proceso complejo. En las primeras dos secciones de este capítulo se expone explícitamente qué tipo de variables miden esta calidad, qué es exactamente lo que miden esas variables y qué dice la literatura sobre cómo estas variables se relacionan con el desarrollo de los niños pequeños, para poder entender los pros y los contras de cada una de estas variables. Asimismo, se presenta en este capítulo una sección que describe cómo algunos de los servicios de cuidado en ALC miden y monitorean la calidad de sus centros.

seguimiento nutricional. En relación con los niños y el cuidador, algunas características estructurales que se miden son: el tamaño del grupo de niños, el coeficiente de atención (número de niños por adulto), la calificación profesional del personal que se encarga de la atención y el cuidado de los niños (años de educación, experiencia previa y capacitaciones realizadas), la planificación de las actividades diarias, el sistema de supervisión y monitoreo de los cuidadores, y los salarios del personal. Entre las variables de características de la infraestructura se encuentran: acceso a agua potable, sistema de eliminación de residuos, acceso a electricidad y teléfono, cantidad de espacio físico para niños, disponibilidad de materiales, instalaciones protegidas contra riesgos ambientales u otros riesgos, características de las áreas de juego y de los muebles.

2.2 / Variables estructurales

¿Qué son?

Las variables estructurales identifican la presencia o ausencia de aquellos recursos que facilitan las interacciones características de un entorno de cuidado, y son además variables que pueden ser reguladas con mayor facilidad (Thomason y La Paro, 2009). Las autoras han clasificado estas variables en 4 grandes grupos, según se relacionen con la infraestructura del centro, temas de salud y seguridad, el grupo de niños y el cuidador. La Tabla 1 resume estos grupos de variables y provee algunos ejemplos. En cuanto a temas de seguridad y salubridad, algunos ejemplos de variables estructurales son: la existencia de protocolos de aseo, salud y seguridad, el protocolo de acciones para evitar el contagio de enfermedades infecciosas, o el

¿Cómo se miden?

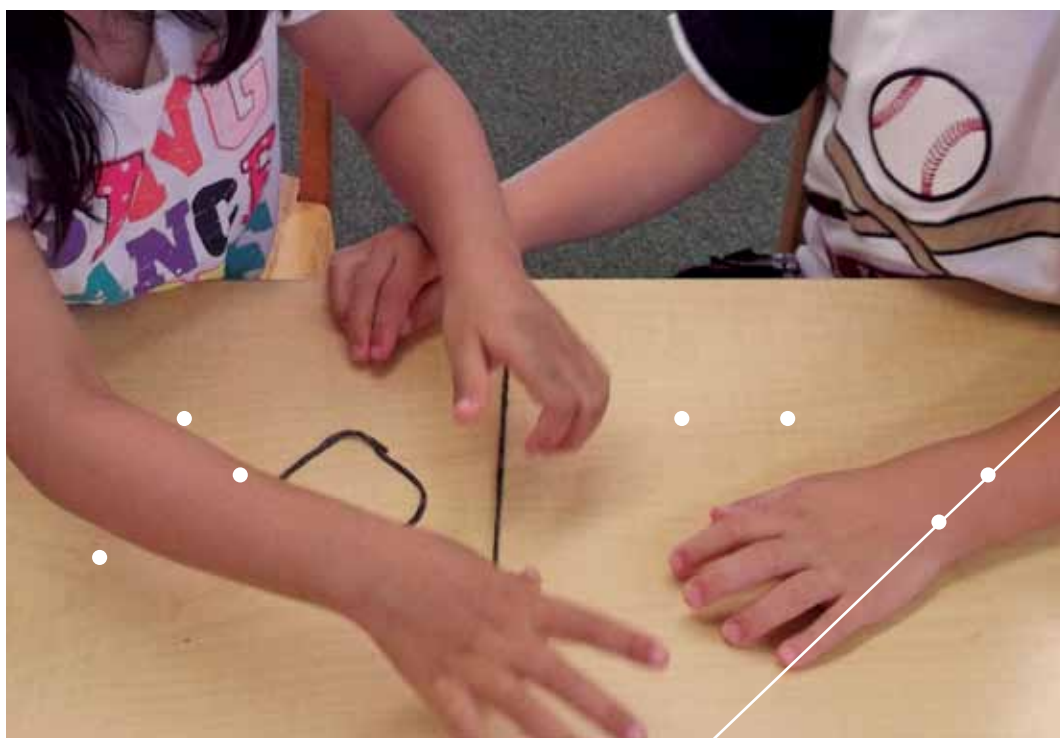
Estas variables se miden a partir de “listas de verificación” por observación (*checklists*, en inglés), entrevistas o cuestionarios. Algunos de los instrumentos que se revisan en esta guía de herramientas y que miden variables estructurales de los centros son: ITERS-R, CC-IT-HOME, MITRCC, ORCE, CIS y Perfil de Evaluación. Las encuestas suplementarias ad hoc pueden también recoger información sobre las variables de infraestructura y sobre las características del cuidador y del coordinador del centro (en la sección 3.6 se presentan detalles de los tres cuestionarios que aplicaron las autoras en Ecuador: Cuestionario estructural, Encuesta de coordinador y Encuesta de cuidador). En algunos programas con sistemas de información sofisticados, se cuenta con datos administrativos que son actualizados regularmente con información para cada centro.

¿Qué dice la literatura?

Se ha encontrado que las variables estructurales están significativamente

correlacionadas con la calidad del cuidado infantil (Howes et al., 1992). Por ejemplo, cuando el coeficiente de atención es más alto, las interacciones positivas y afectivas ocurren con menor frecuencia. Cuando se cuenta con poco o ningún material o infraestructura básica, las rutinas de salud y nutrición tienden a ser de mala calidad o inexistentes y las actividades son menores en frecuencia y calidad (Clarke-Stewart, Gruber y Fitzgerald, 1994; Howes, 1983; NICHD, 1996, 2000a; Phillips en Burchinal et al., 1997; Volling y Feagans, 1995). Además, se ha demostrado que los salarios de los cuidadores están altamente correlacionados con la calidad de los centros de cuidado infantil. Asimismo, los salarios más altos están relacionados con mejores resultados en los niños (Kagan, 2010). En la misma línea, la encuesta nacional de centros de cuidado infantil del Early Childhood Network del NICHD (National Institute of Child Health and Human Development) encontró que los cuidadores que tenían más años de educación y aquellos que se capacitaron en temas de educación de primera infancia tienden a ofrecer actividades más apropiadas y estimulantes para el desarrollo de los niños bajo su cuidado (NICHD, 2000a; Vandell y Wolfe, 2000).

Varios estudios han encontrado que existen relaciones significativas entre variables estructurales de calidad y el desempeño de los niños. En concreto, clases con menos niños, coeficientes de atención bajos (es decir, pocos niños por cuidador) y cuidadores bien calificados se asocian tanto con una mejor calidad del cuidado como con mejores resultados de desarrollo de los niños (NICHD, 1999). Asimismo, las aulas con un mayor número de adultos por niño se asocian con mejores habilidades de lenguaje y conversación más sofisticada del niño, mayor cultura general y mayor participación en interacciones positivas y cooperativas entre los niños (Vandell y Wolf, 2000). Howes et al. (1992) sugieren que esto se debe a que los cuidadores que son responsables de grupos más pequeños tienen la capacidad de brindar más atención y cuidados individualizados a los niños a su cargo. Investigaciones no experimentales han mostrado que los niños en edad preescolar, bajo el cuidado de adultos con una mejor educación y capacitación sobre desarrollo infantil, obtienen resultados significativamente mejores en pruebas cognitivas estandarizadas, se concentran más fácilmente en las tareas y se comunican mejor con los adultos y entre sí (Hausfather et al., 1997; Vernon-Feagans et al., 1997; Howes, 1997; Dunn, 1993; Clarke-Stewart



et al., 2002). El estudio a escala del National Institute of Child Health and Human Development (NICHD) hizo eco de todos estos hallazgos: es más factible brindar servicios de cuidado de calidad cuando la razón de adultos por niño es mayor, cuando los cuidadores han recibido capacitación específica en desarrollo infantil y cuando poseen ideas adecuadas sobre el cuidado de niños pequeños (NICDH, 2000a).

Por otro lado, existe evidencia empírica que sugiere que los elementos estructurales por sí solos no garantizan una atención de calidad y, por lo tanto, utilizarlos como una única representación de las medidas de calidad del cuidado puede ser insuficiente y hasta peligroso (Hamre y Pianta, 2007). Un meta-análisis sobre la efectividad de más de 300 intervenciones que se basan en mejorar la calidad a través de cambios en las características estructurales, tales como incrementar las calificaciones profesionales de los maestros o reducir el tamaño de la clase en niveles desde el preescolar y la primaria hasta la secundaria, encontró muy poca evidencia de que estas mejoras causaron impactos significativos en los logros académicos de los niños o en sus competencias sociales (Hanushek, 2003; NICHD, 2002b). Otra revisión de intervenciones que mejoran las características estructurales sugiere que aumentar los insumos o fortalecer las calificaciones de los cuidadores no tiene ningún impacto sobre el desarrollo de los niños en sus primeros años de escolaridad (Hanushek, 1998). Dicho esto, es importante reiterar que la mayor parte de estos datos proceden de aulas de niños en edad preescolar y primaria en Estados Unidos. La importancia de elementos estructurales como el coeficiente de atención o las condiciones de infraestructura, aseo y saneamiento podría ser mayor para los niños más pequeños o los niños en contextos con altas tasas de desnutrición crónica y de pobreza extrema cuyo sistema inmunológico está todavía en desarrollo (Fernald, Gertler y Neufeld, 2009).

2.3 / Variables de proceso

¿Qué son?

Los indicadores de proceso tienden a focalizarse en aspectos dinámicos (Thomason y La Paro, 2009), como las interacciones que los niños tienen con sus cuidadores, la relación entre los propios niños, el comportamiento de los cuidadores, la implementación del currículo y de los protocolos de salud y seguridad, o la relación entre los cuidadores y las familias (Tabla 1).

¿Cómo se miden?

La medición de los indicadores de proceso requiere de la observación y codificación de las rutinas de cuidado en el centro por una cierta cantidad de horas. Algunos de los instrumentos que se focalizan completamente en las variables de procesos son el Toddler CLASS y los Instrumentos de Rutinas piloteados en español en Ecuador por el BID. Asimismo, instrumentos como el ITERS-R, CC-IT-HOME, MITRCC, ORCE, CIS y el Perfil de Evaluación miden también algunas variables de proceso (ver mayores detalle de la medición en el Capítulo 3).

Si bien los instrumentos de observación que caracterizan los procesos de cuidado infantil son una herramienta vital de diagnóstico para medir la calidad, la mayoría de los instrumentos validados que se enfocan en el proceso agregan puntajes generales que describen las experiencias de los niños en cuanto a interacciones en varias dimensiones y casi siempre también incluyen variables estructurales (NICHD, 2000a; Zaslow et al., 2006). Este es el caso, por ejemplo, del ITERS-R (con algunos subítems sobre materiales usados), el CC-IT-HOME (con una escala sobre materiales de aprendizaje), el Perfil de evaluación

(con medidas de prácticas de seguridad en el aula), y el MITRCC (con algunos ítems sobre materiales). Sin embargo, hay algunas medidas de proceso validadas, como el CLASS, ORCE y CIS, que se centran únicamente en actividades muy específicas del proceso (por ejemplo, la estimulación del lenguaje) efectuadas por los propios cuidadores (Vandell y Wolfe, 2000).

¿Qué dice la literatura?

Los indicadores de proceso que miden la calidad de los centros se encuentran mucho más consistentemente relacionados con la calidad de la atención y con los resultados de desarrollo de los niños, que los indicadores estructurales (La Paro et al., 2004; Hamre y Pianta, 2007). Una cantidad considerable de estudios experimentales y cuasi experimentales han encontrado que los niños en aulas de preescolar o en entornos de atención de alta calidad (medida a través de la calidad de las *interacciones* entre los niños y de ellos con sus cuidadores), no sólo son capaces de iniciar y de participar en actividades de aprendizaje de orden superior con sus pares, sino que además alcanzan niveles más altos en las evaluaciones de preparación académica (Campbell y Ramey, 1995). Estos efectos persisten en el tiempo y pueden durar hasta su vida adulta (Greenberg, Domitrovich y Bumbarger, 2001; Hamre y Pianta, 2007; Shonkoff y Phillips, 2000; Peisner-Feinberg et al., 2001; Burchinal et al., 1997; Helburn, 1995; Howes y Hamilton, 1993; Kisker, Hofferth, Phillips y Farquhar, 1991; Kontos y Wilcox-Herzog, 1997; NICHD, 2000a).

Los indicadores de proceso son los que se ha demostrado que están relacionados con la experiencia diaria, los resultados de aprendizaje y el desarrollo del niño. Pianta (2003) examina una muestra nacionalmente

representativa de 224 aulas de preescolar en seis estados de Estados Unidos y concluye que las características de la calidad del aula más directamente relacionadas con los resultados de aprendizaje de los niños en la escuela o en entornos similares al escolar son justamente el tipo y calidad de las interacciones entre niños y cuidadores.

Las mismas variables de proceso que son vitales para garantizar una atención de calidad a los bebés y niños pequeños son también aquellas más difíciles de medir, ya que requieren la interpretación y el juicio de expertos para su medición. La calidad de la atención en estas dimensiones no se logra ni a través de lineamientos por parte de las autoridades ni por la vía exclusiva de las políticas (Howes y Smith, 1995). La evidencia empírica existente en los Estados Unidos sugiere que, para los bebés y niños pequeños, es mucho más difícil producir cuidado de calidad alta —medida a través de indicadores de proceso— que producirlo para niños en edad preescolar. En Estados Unidos el cuidado que se provee a los niños más pequeños es, en general, de menor calidad (Helburn, 1995; NICDH, 2000a) y mucho menos regulado que el servicio provisto a niños en preescolar. Esto se debe a que ante una provisión pública muy escasa, la regulación suele ser básica y no hay énfasis en una política de calidad. En definitiva, estas características del mercado permiten el desarrollo de un sector privado de baja calidad y sin controles estrictos.

En resumen, para los niños más jóvenes, la calidad está mucho más asociada a las variables de proceso que son justamente las que los gobiernos han descuidado más en términos de medición y de priorización (Howes et al., 1992).⁷

⁷ No obstante, la literatura existente sobre el cuidado de bebés y niños pequeños sugiere que los componentes estructurales de la calidad también pueden ser muy importantes para el rango de edad de 0-3 años más que para otros grupos de edad. Por ejemplo, en un metaanálisis de las últimas tres décadas, Vandell y Wolfe (2000) encuentran que tanto los indicadores de proceso como los indicadores estructurales de calidad se relacionan consistentemente con los resultados del niño en el corto plazo.

2.4 / ¿Cómo miden los gobiernos la calidad de los servicios de cuidado?

En general, los programas y/o gobiernos regulan y miden la calidad de los servicios de cuidado de bebés y niños pequeños a través de indicadores estructurales, como la infraestructura básica y el perfil profesional de los cuidadores y/o educadores (Howes et al., 1995). Sin embargo, esta clase de indicadores no capta información sobre los procesos que afectan el desarrollo cognitivo, emocional y de la salud del niño. Tampoco permiten identificar las áreas de mejora en dichos procesos ni monitorear cambios en el tiempo (Schütz, Ursprung y Woessmann, 2008).

En esta sección se analizan los instrumentos para monitorear la calidad de tres servicios públicos de cuidado —uno nacional de México, otro en la Ciudad Autónoma de Buenos Aires (CABA) en Argentina y un tercero de escala nacional en Ecuador. En México, la Cédula de Supervisión Ordinaria es el instrumento de supervisión que debe aplicar bimestralmente el personal del programa en cada delegación federal de la Secretaría de Desarrollo social (SEDESOL) a la totalidad de Estancias Infantiles en operación en su entidad. En la CABA, el personal encargado de la supervisión del programa recoge información sobre todos los centros cada dos meses; mientras que en Ecuador este instrumento se aplicó por personal del programa dos veces y solo a una muestra de todos los centros.

Los tres países analizados aplican con cierta regularidad a sus programas un cuestionario o lista de chequeo que incluye preguntas sobre los temas mencionados en la Tabla 2. Se puede ver en esta tabla que los temas de seguridad son los que se supervisan con mayor cantidad de ítems (con 70% de estos temas en Ecuador y casi 58% de todos los

temas supervisados en México). En segundo lugar se supervisa la información sobre el personal en la Ciudad de Buenos Aires (con un 30% del total de ítems), y las condiciones del inmueble tanto en México como en Ecuador (12,6% de los ítems en Ecuador y 6,5% en México). Temas de equipamiento, orden e higiene se encuentran en tercer lugar de importancia en la Ciudad de Buenos Aires y en México, mientras que la información del centro ocupa el tercer lugar en Ecuador. Es importante destacar que la Junta Nacional de Jardines Infantiles (JUNJI) de Chile también cuenta con un desarrollado sistema de evaluación de sus jardines.

DEFINICIONES CLAVES PARA LA MEDICIÓN



● CAPÍTULO 3

●

●

●

●

●

●

3.1/ Introducción

Como se ha explicado en el capítulo anterior, la medición de la calidad de los centros de cuidado para bebés y niños pequeños involucra temas claves que van desde la selección de las variables estructurales y de proceso que la miden, hasta cómo estas variables reflejan lo que realmente se quiere medir.

Este capítulo va a los detalles de este proceso y explica cómo implementar estas mediciones. En particular, se discute la importancia de definir el objetivo de la medición y de identificar la unidad de análisis del estudio (niños en un grupo, grupos de niños o centros). Además, este capítulo introduce una batería de instrumentos y propone algunos principios sobre cómo seleccionar el más adecuado.

3.2 / ¿Cuál es el objetivo de la medición de la calidad del centro cuidado?

El primer paso a resolver es definir concretamente cuál es el fin de la medición. Muchos objetivos pueden guiar la medición de la calidad de los servicios de cuidado. Estos son algunos ejemplos: seguimiento y/o monitoreo del estado de los niños en el país, evaluación del estado de los beneficiarios de un programa, investigación del impacto de este servicio sobre ciertos indicadores, rendición de cuentas del uso de los recursos públicos que financian el servicio, seguimiento continuo de la calidad del proceso para su mejora. Según sea el propósito de la medición se definirá a quién, qué y cómo es que se quiere medir.

Es importante destacar que, cuando el propósito es monitorear el cumplimiento de estándares mínimos de calidad y su evolución en el tiempo, la medición debe estar estrechamente vinculada al estándar fijado por el programa. Por ejemplo, un servicio de cuidado diurno para bebés y niños pequeños que no incluye en su modelo de atención la provisión de suplementos nutricionales o micronutrientes a los niños ni un seguimiento a su estado nutricional, no puede ser monitoreado a partir de un indicador que describa el estado nutricional de los niños a los que atiende.

3.3/ ¿Cuál es el grupo de estudio?

Una segunda pregunta de interés es definir la población objetivo de la medición. La medición puede orientarse a todos los niños de un país, región geográfica o etnia, así como al conjunto de niños beneficiarios de un programa específico. Una vez elegida la población de interés se deberá definir si es necesario observar a toda la población o si un estudio muestral responde a los objetivos del estudio. Por ejemplo, si el interés de la medición de la calidad de un servicio es obtener un diagnóstico sobre el estado del servicio que se provee a la población de niños beneficiarios del país, interesa medir una muestra lo suficientemente grande para que sea representativa de los centros que atienden grupos relevantes en ese contexto, por ejemplo, regiones, grupos étnicos o estratos socioeconómicos. Si, por el contrario, se desea monitorear el estado y la trayectoria de los centros que conforman un programa, entonces el interés recae en medir a cada uno de los proveedores del servicio. Esto podría permitir construir un sistema de indicadores de monitoreo y seguimiento. Cuando el objetivo es de diagnóstico, por ejemplo, para identificar

aspectos críticos de mejora de la calidad de un servicio probablemente sea necesario seleccionar una muestra de centros representativa del tipo de servicios que se quiere mejorar e invertir en un conjunto de instrumentos que permita caracterizar a fondo el estado de la calidad del servicio en estos centros. En estudios muestrales, el tamaño muestral deberá determinarse a partir de los cálculos de potencia estadística que permiten determinar el mínimo tamaño que se requiere en una muestra para poder identificar estadísticamente un efecto de una magnitud dada con un nivel de precisión aceptable.

Luego de tener clara la población objetivo de la evaluación, es necesario analizar dentro de cada centro que va a ser investigado, cuál es la unidad de análisis. ¿El centro de cuidado? ¿Uno o más grupos de niños atendidos en ese centro y su(s) cuidador(es)? ¿Cada uno de los niños que asiste a él? Claramente el objetivo del estudio, junto con el presupuesto disponible, van a guiar esta decisión.

Si además de medir la calidad del centro hay que medir resultados en los niños, entonces los instrumentos descritos en esta guía de herramientas deben complementarse con instrumentos de medición del desarrollo infantil. Este es un ejemplo en el cual la unidad de análisis del estudio es el niño y donde la calidad de la atención que recibe en el centro es una variable explicativa del desarrollo del niño. Si, por el contrario, el estudio se enfoca exclusivamente en recolectar medidas de calidad a nivel de los centros, entonces la selección de un grupo y su cuidador en cada centro abriría la posibilidad de aplicar una gama más amplia de instrumentos. En el caso en que los centros cuentan con más de un grupo, la selección de este debería hacerse aleatoriamente, a menos que existan razones válidas para no hacerlo de esta manera (por ejemplo, número de niños en cada clase es menor al requerido para el estudio, rangos de edad fuera del objetivo de estudio, etc.). Cuando el objetivo de la medición de calidad

es analizar las diferencias en la calidad entre los diferentes cuidadores mediante los efectos fijos dentro del centro, entonces se necesita tomar al menos dos grupos por centro.

3.4/ ¿Cómo medir?

La recolección de información sobre la calidad del centro de cuidado puede realizarse a través de dos tipos de mecanismos. El primero consiste en evaluaciones basadas en la observación que realiza una persona entrenada para sistematizar lo que observa en un instrumento. El segundo tipo son evaluaciones basadas en el reporte de un informante a quien se le aplica una encuesta o entrevista (por ejemplo, puede ser la madre del niño, un cuidador del centro o el director del mismo). Algunos instrumentos de medición combinan los dos modos de recoger información.

Las medidas por observación requieren que un observador comparta un tiempo determinado (en general, más de una hora) de la rutina diaria de los niños, para recabar información sobre las dimensiones del cuidado que se evalúan y sistematizarla según el protocolo de uso del instrumento. El observador debe estar capacitado en el uso del instrumento y ser capaz de captar aquellas actividades, vínculos, estímulos o interacciones que deban reportarse en la prueba, sin desviarse en detalles intrascendentes. En el caso de la información recogida por reporte, se realiza una entrevista estructurada o se aplica una encuesta a un informante calificado, según una lista de preguntas sobre las dimensiones de la calidad que se quiere evaluar. También se pueden incluir reportes sobre el comportamiento de los niños, las características de los cuidadores

o del centro de cuidado en sí.

La evaluación puede enriquecerse al combinar los dos mecanismos de recolección de información. La implementación de diferentes formas de medición sobre una misma dimensión brindará mayor información y precisión que la obtenida mediante solo una de ellas.

A su vez, los instrumentos que requieren evaluar la calidad a partir de la observación de la rutina diaria pueden realizarse mediante una observación directa al visitar el centro, o a partir de la filmación de la experiencia para su posterior codificación. De este modo, en lugar de observar y puntuar simultáneamente, una persona se encarga de filmar la rutina del niño o grupo de niños, y en otro momento se analiza el video. Los resultados obtenidos por la opción de filmación dependerán tanto de las características de la filmación y del protocolo que se establezca sobre cómo realizarla, como del observador.

Cabe destacar que, independientemente del tipo de estudio a implementar, al seleccionar qué instrumento se va a aplicar, es necesario explorar la validez del mismo para la población que va a ser estudiada. Esto supone analizar si el instrumento es

apropiado para la población en términos de la pertinencia cultural y la lengua, o si es necesario la traducción o adaptación a la lengua o al contexto local (este tema se analiza detalladamente en el Capítulo 4).

Observación vs. Reporte

La observación demanda más tiempo, capacitación y recursos que la implementación de listas de verificación o de cuestionarios que se basan en el reporte de, por ejemplo, el cuidador o coordinador del centro. La ventaja principal de los instrumentos de observación es que permiten describir las experiencias e interacciones de los niños en el centro. La presencia del observador en el centro podría distorsionar de alguna manera la interacción entre el cuidador y el niño (a pesar de tomar todos los recaudos necesarios para evitar interferir en las actividades, como se detalla en el Capítulo 4). Sin embargo, se considera que cuanto mayor sea la duración de la visita del observador en el centro, será más difícil para el cuidador inhibir tendencias y reacciones que forman parte de su rutina cotidiana, en particular luego de la primera hora de visita.



La evaluación por observación es más compleja porque el observador tiene que ser capaz de documentar, codificar o dar un puntaje en cada dimensión, manteniendo una visión objetiva y precisa al respecto. Por esta razón, se requiere de personal capacitado en el tema, con un conocimiento certero del instrumento a aplicar (características, preguntas o ítems a responder, tiempo necesario de la evaluación). Además, se requiere asegurar que las medidas se realicen con confiabilidad. Es decir, primero, que la evaluación sea parecida, independientemente de cuál sea el observador que la aplique; segundo, que el instrumento otorgue consistentemente a cada centro el mismo resultado sobre su calidad si estos son evaluados en dos momentos cercanos en el tiempo (por ejemplo, que un centro de calidad mediana sea consistentemente de calidad mediana si es evaluado hoy o en dos semanas).

Las mediciones por observación son intensivas en tiempo porque se necesita compartir la experiencia de los niños por un tiempo determinado (en general más de 1 hora), durante el cual se obtiene y codifica la información requerida. Los instrumentos de medición por observación requieren de una capacitación de mayor duración para que los observadores se familiaricen con el instrumento. Además, demandan bastante práctica antes de iniciar las mediciones en campo para asegurar la confiabilidad entre los observadores. Un riesgo es que estas mediciones pueden no ser representativas de un día típico. Por ejemplo, el coeficiente de atención cuidador-niños puede variar de una clase a otra, en el transcurso del día (niños con asistencia a tiempo completo vs. medio tiempo), en diferentes días de la semana o en distintas estaciones (Weber y Kariger, 2011). Estas mediciones también pueden ser más susceptibles a problemas como el error de medición y el sesgo del observador (Zaslow et al., 2006; NICHD, 1996). La capacitación intensiva y constante a lo largo del proceso de codificación y los

ejercicios de confiabilidad son medidas que se toman para mitigar estos riesgos.

Algo importante de mencionar es que, para los instrumentos que requieren de observación (no así para los de reporte), es importante que la hora a la cual se realiza la observación sea siempre la misma en todos los centros observados, dadas las rutinas tan específicas de los niños pequeños. En otras palabras, no se puede comparar una observación realizada durante la mañana con una observación realizada por la tarde, cuando los cuidadores y niños se encuentran más cansados. Otros momentos reveladores de la calidad del servicio y que forman parte de la rutina diaria son la hora de llegada y cómo se recibe a los niños en el centro o la hora de la comida y del aseo. Esto requiere que el observador llegue al centro antes de que este inicie su operación. Para observar el momento de la comida, hay que prever estar presente a la hora del desayuno y/o almuerzo. También se debe considerar la hora de la siesta, ya que durante aproximadamente dos horas no habrá nada que observar; ello justifica que, en general, las observaciones se den entre las 08:00 y las 13:00. Además esto supone que un observador no puede cubrir más de un centro por día.

En los instrumentos de reporte, la información se recoge directamente de las preguntas que se realizan a la madre o el cuidador del niño. Por lo tanto, hay una ventaja en facilidad de aplicación, rapidez, menor costo y menor requerimiento de capacitación. Pero el reporte puede estar sujeto a sesgos. Por ejemplo, una madre tiene información imperfecta sobre lo que sucede en el centro de cuidado en el cual deja a su hijo; un cuidador no está dispuesto a compartir información que pudiera revelar que la atención que brinda a los niños a su cargo no es la óptima; o una persona con bajo nivel educativo puede no ser capaz de reportar con precisión acerca de la frecuencia con la cual realiza actividades con el niño.

La evidencia empírica demuestra que los instrumentos de observación directa de la calidad del servicio predicen los resultados del niño en un grado significativamente mayor del que lo hacen las entrevistas o listas de verificación. Por lo tanto, las medidas de observación deben constituir una parte integral de cualquier esfuerzo por garantizar servicios de calidad (Zaslow et al., 2006). También, como se menciona en detalle más adelante, es necesario mantener confiabilidad en el tiempo, y por ende, cuando se hacen relevamientos, cada tanto se hagan ejercicios de confiabilidad para asegurar la minimización de sesgos durante este proceso.

Filmación

Como se mencionó anteriormente, el tipo de información obtenida por medio de la observación presencial o mediante filmación se caracteriza por un mayor nivel de objetividad que la obtenida por evaluaciones de reporte. Adicionalmente, la filmación cuenta con la ventaja de que el codificador de los videos se concentra en lo que está sucediendo con el niño, aislándose de lo que sucede alrededor (por ejemplo, otros grupos de niños que no corresponden a la muestra o actividades del personal auxiliar) y, por lo tanto, puede producir un trabajo más preciso. Sin embargo, los resultados dependen completamente de la filmación que se realice: la calidad del audio y la imagen y la posición de la cámara que permita una correcta observación de las actividades de los niños del grupo de interés. En otras palabras, la calidad de la información obtenida dependerá de cuán adecuadamente se capta en el video la información visual y auditiva presente en las rutinas. Además, siempre está el riesgo de que quede fuera de la cámara información relevante y valiosa en relación con la medición de la calidad de los procesos.

Una de las ventajas principales de la filmación es que permite alcanzar una codificación de mayor objetividad a través

de procesos de múltiples codificaciones a cargo de observadores diferentes. De esta manera, si los puntajes que otorgan los diferentes observadores no se aproximan entre sí, siempre es posible someter la evaluación de un video a un tercer observador y así identificar la fuente de la discrepancia. Adicionalmente, otra ventaja de la filmación es que se puede volver a ver las veces que sean necesarias y codificar de nuevo apartados que no se habían advertido en la primera pasada del video. También permite realizar el análisis de los datos en un ambiente más controlado y con comodidad, sin que el personal que se dedique a analizarlo tenga que desplazarse hasta los centros. Esto puede ser un elemento importante dado que muchas veces el personal más calificado para realizar este tipo de observaciones puede tener menores incentivos para desplazarse hasta lugares aislados e incurrir en los costos personales que ello representa.

Claramente la filmación puede resultar más costosa. Si bien la filmación puede ser realizada por la misma persona que luego analiza el video, requiere de un mayor número de horas de trabajo (filmación, edición y luego análisis). Una opción que resulta más costo efectiva es utilizar dos personas diferentes: una para la filmación y edición (con perfiles de educación menores) y otra para la codificación (con perfiles de educación mayores). Se necesita contar con la infraestructura mínima necesaria: cámaras, micrófonos, software y computadoras para editar y para reproducir los videos que se analizan. Esto también requiere preparar al personal para una correcta filmación, además de capacitarlo en el instrumento.

Sin embargo, si se cuenta con el presupuesto, observaciones directas y filmaciones pueden utilizarse de forma simultánea para obtener una mayor confiabilidad en los resultados.

3.5/ ¿Qué instrumentos existen?

La evidencia revisada hasta este punto sugiere que es importante medir tanto los elementos estructurales como aquellos de proceso con el fin de tener una medición completa y confiable de la calidad de los servicios de cuidado. Como se dijo antes, varios instrumentos miden ambos elementos de la calidad. También es importante destacar que se han desarrollado instrumentos que permiten evaluar la calidad del cuidado y la atención que reciben los niños en otros entornos, por ejemplo, en el hogar con sus padres o con una niñera.

En las últimas dos décadas, los instrumentos de medición de la calidad de los procesos más utilizados en los Estados Unidos para centros de desarrollo infantil son los siguientes:

- La Escala de calificación del ambiente de la infancia temprana, edición revisada (ECERS-R, por las siglas en inglés de “Early Childhood Environment Rating Scale, Revised edition”) y la Escala de calificación del ambiente para bebés y niños pequeños, edición revisada (ITERS-R, por las siglas en inglés de “Infant and Toddlers Environment Rating Scale, Revised edition”).
- El Sistema de observación y calificación en el aula (CLASS, por las siglas en inglés de “Classroom Assessment Scoring System”) que existe en varias versiones. Las versiones del instrumento CLASS que se han usado desde hace más tiempo son aquellas del Pre-K y tercer grado. Recientemente se han

desarrollado versiones para bebés (“Infant CLASS” que se aplica a niños de entre 0 y 18 meses) y niños pequeños (“Toddler CLASS” que se aplica a niños entre 15 y 36 meses).⁸ El Toddler CLASS es el que más se ajusta al grupo etario considerado en esta guía.

- El Registro de observaciones de los entornos de prestación de cuidado (ORCE, por las siglas en inglés de “Observational Record of the Caregiving Environment”).
- La Escala de interacción del educador (CIS, por las siglas en inglés de “[Arnett] Caregiver Interaction Scale”).
- El Perfil de evaluación (APECP/APFCCCH, por las siglas en inglés de “Assessment Profile for Early Childhood Programs and Assessment Profile for Family Child Care Homes”).
- La Lista de verificación de Missouri para el cuidado sensible de los niños menores de tres años (MITRCC, por las siglas en inglés de “Missouri Infant-Toddler Responsive Caregiving Checklist”).
- HOME para cuidado de bebés y niños pequeños (CC-IT-HOME, por las siglas en inglés de “Child-Care Infant and Toddlers HOME Observation for Measurement of the Environment”).
- Rutinas de comida de un niño (Feeding Scale, NCAST, 2012).⁹

Antes de proseguir con la presentación de cada instrumento, vale la pena mencionar que las autoras de esta guía de herramientas trabajaron con algunos de los instrumentos citados anteriormente.¹⁰ Con el fin de estudiar la calidad del servicio de cuidado en los *Centros Infantiles* del Buen Vivir (CIBV) en Ecuador durante el año 2012 (Araujo et al., 2015), se aplicó una batería de instrumentos que permite

⁸ También se ha desarrollado una versión para aulas de niños mayores de tercer grado, *Upper Elementary CLASS*, y se está desarrollando una versión del instrumento para las aulas de segundo de básica.

⁹ Las autoras adaptaron este instrumento para la construcción del “instrumento de comida” presentado en la subsección 3.5.8. Es importante mencionar que este instrumento difiere de los otros en el sentido que no es una medida “global” de la calidad, sino que se focaliza en un proceso en particular, en este caso, una rutina.

¹⁰ Los cuestionarios aplicados por las autoras en Ecuador no se incluyeron en el documento para no extralimitarse en su extensión. Sin embargo, están disponibles y si algún lector tiene interés en ellos puede contactar a las autoras (mcaraujo@iadb.org, o florencial@iadb.org).

la evaluación de variables de estructura y proceso. Los instrumentos que se aplicaron a los 400 centros de cuidado incluyen la versión en español, tal cual fue diseñada por sus autores, del ITERS-R, y versiones traducidas de Toddler CLASS, MITRCC y CC-IT-HOME. Estas traducciones se encomendaron a especialistas en el tema. En el caso de ITERS-R se capacitó con la instrucción de que se puntuaran todos los ítems (esto es la “Opción de calificación alternativa” que aparece en la versión original en español del manual, página 6, Harms et al., 2006). Asimismo, fue necesario adaptar al contexto del estudio al MITRCC y CC-IT-HOME. Del primero se eliminaron 6 ítems que no eran aplicables al contexto latinoamericano; y se desdobló un ítem en 6 ítems para hacerlo más fácil de aplicar. El CC-IT-HOME fue adaptado para ser aplicado en centros de cuidado y por eso se eliminó un ítem. Además, las autoras diseñaron y pilotearon instrumentos de rutinas que permitan caracterizar algunas de las rutinas de cuidado más importantes que ocurren en los centros.

3.5.1/ ECERS-R e ITERS-R

Descripción

Existe una familia de instrumentos, el ECERS-R, el ITERS-R y el FCCERS-R, que miden la calidad de centros de cuidado en diferentes grupos etarios (ITERS-R para más pequeños y ECERS-R para más grandes) o distintos contextos (ECERS-R/ITERS-R en centros y el FCCERS-R en el hogar).

El ECERS-R está diseñado para las aulas de niños entre 2 años y medio y 5 años, mientras que el ITERS-R es para los niños desde el nacimiento hasta los 2 años y medio. Ambos constan de 7 dimensiones. Una de estas dimensiones se enfoca en la dotación de espacio y muebles. Las otras seis se relacionan con los procesos. De ellas, hay una con énfasis en las interacciones del educador con el niño. Los instrumentos recogen información detallada sobre la infraestructura, los protocolos de salud y nutrición, los factores de seguridad, la disponibilidad y acceso



del material de aprendizaje, el tipo de actividades que se llevan a cabo durante el día, etc. (Cryer et al., 2004a,b). Las pruebas se implementan en tres horas, que se dedican a la observación y calificación. Algunos de los ítems se completan a través de la entrevista a las cuidadoras. La calificación se realiza siguiendo una escala numérica del 1 al 7, para cada uno de los 39 ítems que comprenden las 7 dimensiones presentes en el ITERS-R, donde un 1 significa baja calidad, 3 una calidad mínima, 5 buena calidad y 7 una excelente calidad. Los 43 ítems de ECERS-R se puntúan de la misma forma. En la Tabla 3 se presentan las principales características de ITERS-R.

Experiencias de aplicación

La escala ECERS-R fue ampliamente usada para una muestra de niños de 2 a 4 años en centros de cuidado de los Estados Unidos (Burchinal, Kainz y Cai, 2011). En ALC, cuatro países han aplicado alguno de estos instrumentos (ver Tabla 3).

El programa de Colombia *Hogares Comunitarios* fue evaluado en 2007 utilizando el instrumento FCCERS-R con el fin de medir la calidad de los servicios que el programa ofrecía a los niños. Este programa provee servicios de cuidado con una modalidad comunitaria, en donde madres de la comunidad son las encargadas de la atención de los niños, la misma que se brinda en su propio hogar. El programa tiene una cobertura muy amplia y atiende a los niños más pobres del país desde 1986 (Bernal et al., 2009). Un hallazgo interesante de la evaluación fue que, a pesar del éxito relativo del programa en la mejora de algunos resultados cognitivos y de salud de los niños, la calidad de los servicios medida a través de los puntajes que arrojaron estas pruebas en los aspectos estructurales y de proceso de los entornos del cuidado fue, en promedio, muy baja (un puntaje de alrededor de 2 puntos). Esto dio espacio a una serie de iniciativas para mejorar la calidad del servicio (Bernal et al., 2009). Existen también dos trabajos en curso que están aplicando este instrumento en Colombia.

En 2009, un equipo de investigadores aplicó el ECERS-R y el ITERS-R exitosamente para evaluar la calidad de los centros de cuidado y los servicios de preescolar en Brasil. Dada la heterogeneidad en la calidad de los entornos de cuidado infantil evaluados en este estudio, se adaptó la escala al contexto brasileño, ampliando el rango de la parte más baja de la escala para captar la variabilidad en el segmento de calidad más baja del espectro (Verdisco y Pérez Alfaro, 2010). El análisis permitió concluir que, si bien existe una gran variabilidad en la calidad de los centros de cuidado, por lo general los niños más pobres son los que reciben servicios de cuidado de más baja calidad (Verdisco y Pérez Alfaro, 2010).

El ECERS-R y el ITERS-R se adaptaron de manera similar para su uso en Chile y fueron aplicados en aulas para bebés y preescolares de escuelas públicas y privadas a fin de poner a prueba su validez en el contexto chileno y entender mejor la calidad general de los primeros entornos de aprendizaje de los niños de ese país (Herrera et al., 2005). El equipo de investigación identificó algunas características psicométricas de los instrumentos que eran aplicables al contexto chileno. Los resultados de las aplicaciones de ECERS-R e ITERS-R revelaron que la calidad de los centros evaluados era muy variable: el puntaje promedio fue muy bajo en toda la muestra y especialmente entre los niños de 0 a 3 años. Por último, se encontraron resultados similares a los de la evidencia empírica de otros países: el aumento de la calidad de los programas para la primera infancia en Chile está estrechamente relacionado con la mejora de los resultados de los niños (Herrera et al., 2005). Las autoras de esta guía de herramientas aplicaron el ITERS-R en Ecuador (Araujo et al., 2015) y también se ha aplicado ITERS-R en la línea de base de la evaluación de impacto de la modalidad de cuidado diurno del Programa Nacional Cuna Más en Perú, y se lo está considerando en Uruguay.

3.5.2/ Toddler CLASS

Descripción

De los instrumentos mencionados anteriormente, sólo CLASS mide puramente variables de proceso y no mide en ningún grado los factores estructurales. Todas las versiones del instrumento (pre-K CLASS —para entornos de cuidado de niños de más de 36 meses de edad hasta primer año de educación general básica o *kinder*/prescolar—, Infant CLASS —niños de 0-18 meses de edad—, y Toddler-CLASS —niños de 15-36 meses de edad), fueron diseñadas para enfrentar las limitaciones de otros instrumentos enfocados en variables estructurales. Al hablar de proceso, estos instrumentos evalúan “lo que los maestros hacen con los materiales que tienen y sus interacciones con los niños” (La Paro et al., 2004, p. 412).

Vale la pena destacar que las dimensiones que se aplican a la evaluación de las interacciones entre educadores/docentes (o cuidadores en el caso del Toddler-CLASS) y niños son similares a lo largo de la educación inicial, preescolar y básica. Sin embargo, las diferencias se presentan en las maneras en que estas dimensiones se manifiestan en cada grupo etario (La Paro et al., 2011). Por esta razón, la versión Toddler CLASS adapta las descripciones de los ítems a las especificidades que se encuentran en las aulas de educación inicial y preescolar.

En la versión Toddler CLASS las interacciones se califican a partir de la observación de ocho dimensiones divididas en dos grandes áreas: (a) *apoyo emocional y conductual* y (b) *apoyo motivador del aprendizaje*. El primer grupo incluye cinco dimensiones que se concentran en la expresión de emociones (positivas y

negativas) de los cuidadores y los niños, la capacidad de respuesta, la disponibilidad y la sensibilidad del cuidador, el grado en el que se toman en consideración las perspectivas de los niños y se fomenta su responsabilidad e independencia, y el apoyo al desarrollo de la regulación de la conducta. El segundo grupo se orienta a la observación del cuidador en su rol de facilitador de las actividades orientadas a apoyar el aprendizaje de los niños y a promover su desarrollo lingüístico, así como el grado en que se retroalimenta la respuesta de los niños a partir de la comprensión y participación (La Paro et al., 2011).

En cada dimensión se otorga una calificación numérica en el rango del 1 al 7. La puntuación asignada depende de una serie de valoraciones por parte del observador, guiados por el manual y el aprendizaje obtenido durante la capacitación, de modo de mantener la objetividad y precisión. La calidad de la evaluación se asegura a partir de la capacitación y certificación del personal.

Thomason y La Paro (2009) analizan la validez del Toddler CLASS a partir del estudio de 46 educadores en 30 aulas de educación inicial y preescolar de los Estados Unidos. La validez del constructo del Toddler CLASS se establece a partir de revisiones exhaustivas de medidas existentes, una revisión de la investigación sobre los aspectos únicos del desarrollo infantil temprano, y observaciones sobre entornos de educación inicial y de preescolares. Los instrumentos revisados para establecer su validez incluyen el ITERS-R (Harms et al., 2006), el CIS (Arnett, 1989), y el ORCE usado en el “NICHD Study of Early Child Care” (NICHD, 1996). Los ejemplos e indicadores incluidos en el instrumento adaptado reflejan esta revisión. De forma similar al pre-K CLASS (Pianta et al., 2008), se proveen ejemplos para la codificación de entornos entre el rango bajo (1-2), medio (3-5) o alto (6-7). Adicionalmente, el instrumento fue revisado por un experto

en bebés/niños pequeños para asegurar la validez de los conceptos adaptados.

Fueron significativas las correlaciones con medidas tradicionales de calidad en educación infantil temprana, como así también con el nivel de educación del cuidador, tamaño del grupo, coeficiente de atención niño-educador, puntajes asignados (en función de la inclusión de aspectos de calidad en el programa y el nivel educativo de los cuidadores) y las dimensiones. Ver Anexo 1 para mayor detalle de la validez del instrumento Pre-K CLASS (Pianta et al., 2005). En la Tabla 4 se presentan las características salientes del instrumento.

Experiencias de aplicación

Existe un metaanálisis de 20 estudios publicados en revistas profesionales que pasaron por un proceso de revisión externa por expertos. Estos estudios miden la calidad del cuidado de los niños y sus resultados de desarrollo, usando una muestra de niños de 0 a 4 años en centros de cuidado de los Estados Unidos durante la última década. Así es que Burchinal, Kainz y Cai (2011) encontraron que la escala pre-K CLASS brindaba asociaciones ligeramente más consistentes con los resultados de desarrollo de los niños que el instrumento que ha sido el más utilizado hasta la fecha, el ECERS-R. Los ítems individuales que más predecían los resultados del niño en la escala pre-K CLASS fueron las dimensiones de *manejo del comportamiento, productividad, clima positivo y clima negativo*. Cabe anotar que la versión de CLASS para los niños más pequeños (Infant CLASS) se encuentra actualmente en fase de desarrollo, por lo que la información y la evidencia de su uso para niños menores de 18 meses de edad son limitadas (Teachstone, 2012c).

En la actualidad, el conocido *Programa Nacional Head Start* de los Estados Unidos (NHS, por sus siglas en inglés) utiliza los instrumentos de pre-K CLASS y Toddler CLASS como herramientas de seguimiento

y desarrollo profesional de su personal. Este esfuerzo implica la capacitación de más de 200 especialistas del equipo de monitoreo del NHS en todo el país, quienes utilizan el CLASS para dar seguimiento a la calidad y brindar capacitación sobre la herramienta a los directores locales del programa para que ellos, a su vez, puedan monitorear la calidad y establecer metas de mejora en sus propios programas (Teachstone, 2012a). El pre-K CLASS sirve principalmente para proporcionar un vocabulario común a todos los educadores, directores de programas y responsables de políticas de manera que sean capaces de identificar, discutir y dar seguimiento a los temas de calidad en diversos programas en todo el país. También en los Estados Unidos, en el estado de Virginia, el pre-K CLASS está siendo utilizado como parte de un sistema estatal de acreditación y monitoreo de los entornos de cuidado de la primera infancia. El uso de CLASS permite que los padres y los responsables de políticas puedan distinguir fácilmente aspectos relacionados con la calidad de los establecimientos. Además, hace posible dar seguimiento a la evolución de la calidad en los diferentes programas y guía a los educadores y cuidadores sobre cómo mejorar la calidad de sus servicios (Teachstone, 2012b).

En Chile, el pre-K CLASS está siendo utilizado en Santiago como una herramienta de seguimiento y de desarrollo profesional (Yoshikawa et al., 2008). En 2008, el instrumento fue aplicado por primera vez por el programa piloto *Un buen comienzo* (UBC) como parte de un esfuerzo integral de desarrollo profesional diseñado para mejorar la calidad de la educación preescolar y para mejorar el desarrollo cognitivo, social y de salud de los niños, así como la participación de los padres en la escuela. El pre-K CLASS fue utilizado para monitorear los cambios en la calidad en el transcurso de la intervención y se aplica en la actualidad en una serie de programas piloto adicionales en otras regiones del país (Yoshikawa et al., 2008). En Ecuador, el K-3 CLASS ha sido ampliamente

aplicado a una muestra de alrededor de 600 aulas de *kinder* y de primer grado como parte de un estudio conjunto del Banco Interamericano de Desarrollo y el gobierno de ese país sobre la calidad de la educación inicial. También se está aplicando en Perú (Toddler CLASS). Las autoras de esta guía de herramientas aplicaron el Toddler CLASS en Ecuador (Araujo et al., 2015).

3.5.3/ ORCE

Descripción

El ORCE se ha usado en una amplia variedad de entornos de cuidado diferentes del materno en niños de entre 0 y 60 meses de edad, y tiene un fuerte enfoque en la calidad de las interacciones del cuidador (u otros adultos presentes) con el niño. Este instrumento busca caracterizar las interacciones que experimenta cada uno de los niños de forma individual en el aula (Burchinal, Kainz y Cai, 2011; Vandell y Wolfe, 2000). Es importante destacar que la sección de la calidad de procesos del ORCE se evalúa focalizándose en un niño específico (“target child”), mientras que la parte estructural se evalúa en base a un grupo. La calificación consta de otorgar un puntaje de 1 a 4 a cada ítem. La calificación total surge del promedio entre los tres dominios para todos los ciclos de observación (Vandell y Wolfe, 2000).

En cuanto a la validez del constructo del ORCE, las medidas de calidad del cuidado infantil que este instrumento arroja fueron relacionadas positivamente con variables estructurales como el nivel de educación del cuidador y el coeficiente de atención. En cuanto a la validez predictiva, estas medidas de calidad del cuidado infantil fueron relacionadas con el desempeño cognitivo,

lenguaje y actividad social (Halle, Whittaker y Anderson, 2010). En la Tabla 5 se presentan las características salientes del instrumento.

Experiencias de aplicación

El ORCE se desarrolló para un estudio a gran escala sobre los cuidados en la primera infancia llevado a cabo por el NICHD en los Estados Unidos. Este estudio buscaba entender mejor la relación entre los servicios de cuidado en la primera infancia y los resultados de desarrollo de los niños (en edades entre 0 y 3 años) en diversos tipos de entornos de cuidado diferentes al materno (Burchinal, Kainz y Cai, 2011). El propósito específico del diseño de este instrumento fue hacer frente a dos limitaciones de ECERS-R e ITERS-R. Primero, que esos instrumentos se centraban demasiado en las características estructurales en detrimento de la captación del “minuto a minuto” de la experiencia y proximidad del niño a su cuidador. Y segundo, que las escalas ECERS-R e ITERS-R sólo eran adecuadas para entornos de cuidado institucionales como los centros de cuidado (NICHD, 2009; Vandell y Wolfe, 2000) mientras que el ORCE también podía aplicarse en otros entornos de cuidado tales como el hogar. En la revisión de la literatura presentada, no se ha podido encontrar referencias de aplicación del ORCE fuera de los Estados Unidos y no se ha aplicado en Latinoamérica.

3.5.4/ CIS

Descripción

El CIS combina una escala reducida de tipo “lista de verificación” que evalúa a los cuidadores en su “sensibilidad, severidad y nivel de separación” en las interacciones con los niños de edad preescolar, tanto en el entorno de cuidado del hogar como

en un entorno institucional. Se puntúa la frecuencia con la cual se observan ciertos comportamientos por parte del cuidador, centrándose principalmente en las características estructurales del cuidado y las interacciones entre los cuidadores y los niños (Vandell y Wolfe, 2000; Burchinal, Kainz y Cai, 2011). A cada uno de los 26 ítems de las cuatro dimensiones del instrumento se le otorga una calificación de 1 (ausente) a 4 (muy presente).

La escala no aborda temas como el currículo u otras estrategias de gestión del aula, como el agrupamiento de los niños según su edad o el flujo de actividades (Denham, Ji y Hamre, 2010). En la Tabla 6 se presentan las características salientes del instrumento.

Experiencias de aplicación

El CIS fue aplicado en Bermuda (Arnett, 1989) y en varios estudios de centros de cuidado en Filadelfia (Jaeger y Funk, 2001) y en otras ciudades de Estados Unidos. Este instrumento siempre mostró niveles de cuidado por debajo de los mínimos establecidos por los estándares de los programas.

3.5.5/ Perfil de evaluación

Descripción

El Perfil de Evaluación (PE) es una herramienta que, como las otras, evalúa las prácticas de un centro o un aula. Este instrumento tiene dos versiones¹¹: Medida global de calidad y Medida de autoevaluación. La Medida global de calidad permite evaluar el ambiente de aprendizaje

y las prácticas de enseñanza en las aulas para niños entre 3 y 7 años. Mientras que la Medida de autoevaluación es más bien utilizada para propósitos de mejoramiento de programas de primera infancia. Dicha medida además de evaluar dimensiones tales como las prácticas en el aula, el entorno del aprendizaje, la planificación, los métodos curriculares, las interacciones y la individualización, también evalúa las prácticas administrativas, tales como las instalaciones, los servicios de comida, las políticas y los procedimientos, el personal y el profesionalismo, la evaluación y el desarrollo. Esta medida es más amplia que la herramienta de investigación global, resulta fácil de usar en los procedimientos para la autoevaluación de los programas para niños pequeños (entre 0 y 10 años) ya que los ítems son de simple aplicación, y además genera un análisis extenso y detallado de recomendaciones de mejora para los programas. El instrumento se puntúa otorgando “Sí/No” a cada uno de los ítems de las cinco dimensiones del instrumento y luego se suman los “Sí”.

La validez de criterio del instrumento se estableció mediante el examen de la relación del PE con el ECERS (Harms y Clifford, 1980). En estos estudios, Wilkes (1989) encontró una correlación significativa ($r = 0,64$, $p = 0,001$), y Abbott-Shim (1991) también encontró una correlación significativa ($r = 0,74$, $p = 0,001$).

Para la validez del constructo se realizó un análisis factorial de segundo orden para determinar si las cinco escalas del PE formaban un constructo único de calidad en las aulas. Los resultados indicaron que las mediciones observadas utilizando estos factores sí provienen de un único constructo subyacente de la calidad en el aula (Abbott-Shim, Lambert y McCarty, 2000).

¹¹ Las versiones del Perfil de Evaluación se denominan en inglés “summative” y “formative”. En lugar de usar una traducción literal de estos nombres, en esta guía se denomina a cada versión en función de su contenido y utilización, llamándose Medida global de calidad a la primera y Medida de autoevaluación a la segunda.

A su vez, la validez de contenido se documentó a través de una revisión del instrumento por una amplia gama de profesionales de la primera infancia y una referencia cruzada a los criterios de acreditación de la Asociación Nacional para la Educación de Niños Pequeños (NAEYC, por las siglas en inglés de National Association for the Education of Young Children). La referencia cruzada mostró una consistencia amplia entre las dos medidas con un 100% de coincidencia en los criterios (Abbott-Shim, Neel y Sibley, 2001). En la Tabla 7 se presentan las características salientes del instrumento.

Experiencias de aplicación

Esta escala se aplicó en varios estados de los Estados Unidos. Por enfocarse tanto en los aspectos del entorno en el cual los niños aprenden así como en las interacciones, Denham, Ji y Hamre (2010) identificaron los instrumentos CLASS, ECERS-R, ITERS-R y el Perfil de evaluación como herramientas que evalúan simultáneamente la presencia de elementos críticos para el desarrollo cognitivo y socioemocional saludable de los niños pequeños.

3.5.6/ MITRCC

Descripción

Otro instrumento que utiliza un formato de “lista de verificación” es la Lista de verificación de Missouri para el cuidado sensible de los niños menores de 3 años (MITRCC, por las siglas en inglés de “Missouri Infant/Toddler Responsive Caregiving Checklist”). Se enfoca en niños de hasta 3 años de edad y puede utilizarse en centros de cuidado u hogares.

Este instrumento evalúa por un lado, la labor del cuidador en reconocer y responder a las señales de los niños, enseñar a anticipar rutinas y eventos, compartir interacciones positivas y de empatía con los compañeros, reconocer sus emociones y resolver sus propios problemas, así como estimular el desarrollo de motricidad fina y gruesa, y el desarrollo cognitivo. Por otro lado, el instrumento también incluye la observación de variables de estructura como por ejemplo en el ítem de “si los cuidadores organizan el aula de forma que los niños tengan oportunidad de observar el entorno desde más de un nivel”, el ítem de “si hay objetos similares y diferentes que estén al acceso del niño para comparar y contrastar”, o también “si hay objetos para experimentar verter y vaciar”. Se puntúa “Sí/No” a la presencia de cada ítem en el 75% de los niños que se encuentran alertas y despiertos; y el puntaje total de “Sí” (de un máximo de 20) se lo divide por dos para tener un puntaje final entre 0 y 10. A diferencia de los otros instrumentos analizados, los ítems del MITRCC no se refieren al tipo de características de la calidad que se evalúan sino a las áreas del desarrollo infantil (socioemocional, físico y cognitivo) que se están promoviendo a través de esas atenciones de calidad. Esta es una distinción importante pues subyace un supuesto (o un modelo) según el cual diferentes variables de calidad del cuidado aportan a diferentes dimensiones del desarrollo de los niños. En la Tabla 8 se presentan las características salientes del instrumento.

Experiencias de aplicación

El MITRCC ha sido usado en los Estados Unidos (Missouri Head Start State Collaboration Office, ZERO TO THREE), así como en la evaluación del programa *Aeiotu* en Colombia (NIEER y Bernal, trabajo en proceso). Las autoras de esta guía de herramientas también lo aplicaron en Ecuador y el trabajo está disponible en Araujo et al. (2015) y en Lopez Boo y Tome (2016).

3.5.7/ CC-IT-HOME

Descripción

Cuando el entorno de cuidado es el hogar, el inventario de HOME es una medida de la calidad del medioambiente familiar que se centra en el proceso aunque evalúa también la presencia de variables estructurales.

El inventario HOME original cuenta con una versión (el Child Care Infant/Toddler HOME, CC-IT-HOME) que ha sido adaptada para ser aplicada en entornos de cuidado en el hogar, tales como el cuidado por familiares o niñeras. Asimismo, para el estudio realizado en Ecuador las autoras aplicaron una traducción y adaptación del CC-IT-HOME para ser aplicada a centros realizada por NIEER (National Institutes of Early Education Research). El CC-IT-HOME emplea un enfoque de “lista de verificación”, para evaluar el cuidado e incluye la calificación de elementos como la seguridad del ambiente, prácticas de salud y materiales y actividades provistos.

El instrumento combina la observación con la información provista por los padres (o el cuidador en la versión adaptada a centros). Una característica particular de esta escala es que la unidad de análisis de la observación no es un grupo de niños sino que es la interacción individual entre *un niño y su cuidador*. En otras palabras, la aplicación de esta escala requiere enfocar la observación sobre un niño individual y examinar detalladamente el ambiente de cuidado desde la perspectiva de ese niño. Luego, esta información se generaliza para todos los niños en el entorno. La calificación

de cada ítem se realiza marcando un “Sí/No” y la calificación total máxima es de 42 puntos. El instrumento está dividido en 6 dimensiones. En la Tabla 9 se presentan las características salientes del CC-IT-HOME.

Experiencias de aplicación

El CC-IT-HOME ha sido ampliamente usado en los Estados Unidos y en el Caribe. Clarke-Stewart et al. (2002) cita un estudio que revela que los niños que fueron atendidos en hogares de mayor calidad de acuerdo con el Inventario de CC-IT-HOME, también puntuaron más alto en la prueba de Bayley¹² a los 2 años de edad, y a los 3 años de edad tuvieron significativamente mejores puntajes de aprestamiento escolar, en comparación con sus pares que recibieron cuidados de menor calidad en sus hogares. Las autoras de esta guía de herramientas aplicaron una versión adaptada para centros de cuidado del CC-IT-HOME en Ecuador (Araujo et al., 2015).

3.5.8/ Otros instrumentos

Rutinas del centro

Las autoras de esta guía pilotearon en Ecuador un conjunto de escalas observacionales que evalúan, a través de la observación y codificación, las características de las principales rutinas de cuidado en el centro. El objetivo de este esfuerzo fue revisar si era posible simplificar los instrumentos observacionales que evalúan la calidad de los procesos con el fin de construir

¹² La prueba llamada Bayley, cuyo nombre completo es Bayley Scales of Infant Development consta de tres ediciones en la actualidad: BSID-I, 1st edition; BSID-II, 2nd edition; BSID-III, 3rd edition. El Bayley evalúa de manera directa el desarrollo de niños entre 1 y 42 meses en un amplio espectro de dominios. La prueba sirve para diagnosticar rezagos y planear intervenciones adecuadas y ha sido validada. Esta prueba es considerada la regla de oro o “gold standard” en la medición del desarrollo infantil.

medidas que puedan ser aplicadas a un costo más bajo (en capacitación, tiempo y perfil del personal) y aun así que provean información sobre estas dimensiones de la calidad. Se identificaron momentos claves en los cuales era posible para un cuidador demostrar interacciones individualizadas y sensibles con los niños a su cargo. Los momentos claves de la rutina de cuidado que se eligió evaluar con estos instrumentos son: (a) la comida, (b) el cambio de pañales o ida al baño, según la edad, (c) el manejo de situaciones disruptivas durante el día como el llanto de uno o varios de los niños, experiencias de maltrato verbal, físico o psicológico, o la desconexión del niño con el entorno y las personas que lo rodean.

Las escalas de evaluación de las rutinas de cuidado se han utilizado para medir la calidad de las interacciones entre los padres y los niños o entre estos y los cuidadores en una variedad de contextos. Estas escalas evalúan las interacciones y la comunicación entre los padres o los cuidadores y los niños, ya sea durante la alimentación (para niños entre 0 y 12 meses de edad) o durante la enseñanza (para niños entre 0 y 36 meses de edad). Las escalas presentan la ventaja de que se pueden administrar relativamente rápido; por ejemplo, una escala de

alimentación requiere el tiempo que dura una comida (NCAST, 2012).

El instrumento de comida que fue piloteado en Ecuador es una adaptación realizada por el equipo BID de la “Feeding Scale” (NCAST, 2012) que fue aplicada en los Estados Unidos. Básicamente, se observa la frecuencia y duración de ciertas reacciones (positivas o negativas) del cuidador durante la comida. La medida original posee parámetros de validez interna, externa y predictiva además de una confiabilidad adecuada (NCAST, 2012). El instrumento que las autoras pilotearon para caracterizar las interacciones durante el cambio de pañal o la ida al baño se construye a partir del instrumento de comida, y se enfoca en observar y evaluar las reacciones del cuidador durante este evento particular, tal cual se hace durante una comida.

El tercer instrumento piloteado en Ecuador, enfocado en la observación de momentos de llanto/maltrato/desconexión está basado en varios instrumentos existentes: el *Infant Toddler and Family Instrument* (Provence y Apfel, 2001), el *Family Care Instrument* (FCI), así como el instrumento PRIDI (Engle, Cueto, Ortíz y Verdisco, 2011). El instrumento requiere de observar la frecuencia y duración de alguno de estos



tres eventos (llanto, maltrato entre niños o del cuidador hacia el niño y desconexión) durante 40 minutos consecutivos y ver las acciones del cuidador cuando estos eventos se presentan. Con pocos ítems de observación o preguntas, se busca también evaluar la manera que tiene el cuidador para disciplinar al niño o reaccionar ante momentos de estrés. Una dificultad de este instrumento puede ser que requiera una observación de duración más larga para efectivamente presenciar una situación de esta naturaleza. En entornos de cuidado de calidad se podría pensar que es muy poco frecuente que lleguen a darse circunstancias de desconexión o maltrato y el instrumento se enfocaría entonces en caracterizar únicamente las reacciones de los adultos ante las instancias de llanto.

Todos los instrumentos de rutinas fueron originados y refinados con sugerencias del BID Advisory Committee (2012). Las autoras de esta guía de herramientas pilotearon estos instrumentos en Ecuador en 2012. El análisis de esta aplicación y la funcionalidad de estas escalas se encuentran en proceso al momento de escribir esta guía. En la Tabla 10 se presentan las características salientes de estos instrumentos.

Nuevos instrumentos

Otros instrumentos que miden la calidad y que vale la pena mencionar por su uso relativamente frecuente en las últimas dos décadas son la Escala de clasificación de las evaluaciones de los programas de cuidados para bebés y preescolares (PITC PARS, por las siglas en inglés de “Program for Infant/Toddler Care Program Assessment Rating Scale”) y el Instrumento de evaluación de la calidad de los programas para preescolares (PQA, por las siglas en inglés de “Preschool Program Quality Assessment”). Ambos son instrumentos validados que, en distintos grados, miden los elementos de proceso y estructurales para bebés y niños pequeños y se han aplicado para mejorar el programa,

dar seguimiento o acreditación y con fines de investigación/evaluación en centros de cuidado en Estados Unidos (Denham, Ji y Hamre, 2010; Halle, Whittaker y Anderson, 2010). Una nueva medida —que está en construcción y en fase piloto en 400 centros de cuidado infantil de Estados Unidos— es la Escala de la calidad de la interacción entre el cuidador y los bebés y niños pequeños (Q-CCITT, por las siglas en inglés de “Quality of Caregiver-Child Interaction for Infants and Toddlers”). Financiado por la Oficina de Head Start, el Q-CCITT promete ser un indicador sicométrico lo suficientemente amplio como para hacer comparaciones entre una variedad de entornos de cuidado infantil en los que los padres no están presentes. Con esto, se busca superar las limitaciones de la mayoría de los otros instrumentos que son específicos en términos de entorno, edad y lenguaje (Halle et al., 2011). El Q-CCITT es único por centrarse específicamente en variables de proceso, por aplicarse a entornos de atención diversos (el hogar, centros de cuidado, aulas separadas por edad o de edades mixtas, por enfocarse en niños de 0-3 años), y por cubrir aulas que atienden a niños con discapacidades y a aquellos cuya lengua materna es diferente a la lengua del centro de cuidado (Halle et al., 2011).

3.6/ ¿Qué encuestas complementarias se pueden aplicar?

En esta sección se documenta una experiencia específica de las autoras en el estudio de Ecuador (Araujo et al., 2015). Las autoras de esta guía de herramientas suplementaron su estudio en Ecuador con encuestas que se aplicaron directamente al cuidador del grupo de niños estudiado (entrevista de 40 minutos) y al coordinador

del centro de estudio (entrevista de 1 hora).¹³ En ambos casos se recogió información demográfica del entrevistado y se indagó sobre su formación profesional, experiencia, capacitación continua y salario. Además, tanto los cuidadores como los coordinadores respondieron una escala que evalúa su conocimiento sobre el desarrollo infantil, la KIDI (MacPhee, 1981) por sus siglas en inglés *Knowledge of Infant Development Inventory*, que busca identificar cuánto conocen los adultos sobre distintos comportamientos normales para cada grupo etario en el rango de niños menores de 5 años.

La entrevista al coordinador sirvió también para recoger información general sobre las características del centro, su organización administrativa y su personal, por ejemplo: datos sobre todo el personal que trabaja en el centro (nivel de estudio, experiencia, tiempo en el centro, remuneraciones), información sobre la organización del centro (composición de los grupos, organización, coeficientes de atención), el sistema de información e interacción con los padres o la inclusión de niños de minorías étnicas o con capacidades diferentes.

Por último se aplicó un cuestionario sobre temas estructurales. A partir de la encuesta de estructura del Instituto de la Niñez y la Familia (INFA) de Ecuador y de listas de verificación de distintos países de Latinoamérica, se elaboró un instrumento focalizado íntegramente en la observación de variables de estructura. A partir de una observación de 40 minutos, suplementada con una entrevista de 20 minutos al coordinador del centro, se recogió información sobre el centro de cuidado en términos de los servicios básicos, infraestructura, vulnerabilidades de los espacios, presencia de riesgo para los niños, y capacidades existentes.

Es de destacar que fue importante contar con estos tres instrumentos adicionales a la aplicación de instrumentos de calidad, ya que estos contienen información valiosa sobre aspectos críticos para el bienestar de bebés y niños pequeños de la calidad de los centros de cuidado. Por ejemplo, tener información sobre la edad promedio del grupo de niños atendidos (proveniente del cuestionario al coordinador) permitió a las autoras examinar si las medidas de calidad son mejores en los grupos de niños menores o mayores. Además, con el cuestionario del cuidador y el de infraestructura, se pudieron desagregar los datos para subgrupos de centros, por ejemplo según la calidad de su infraestructura física o según el nivel educativo de los cuidadores que trabajan en él. En la Tabla 11 se presentan las características salientes de estos tres cuestionarios adicionales.

3.7/ ¿Cómo elegir el instrumento?

Existen muchos instrumentos para medir la calidad de los servicios de cuidado infantil, cada uno con sus particularidades. Entonces ¿cuál es el apropiado para cada estudio o evaluación? Esto dependerá de diferentes parámetros que van más allá de los niveles de complejidad y costos en la aplicación. El investigador o el formulador de políticas deberá dar prioridad a uno de estos parámetros, dependiendo de si se persigue medir procesos de monitoreo, seguimiento y/o mejora continua de la calidad de sus centros. La Tabla 12 presenta un resumen de los instrumentos presentados que facilita la comparación de los mismos en términos de la medición y su aplicación.

¹³ El coordinador del centro está encargado de liderar, y promover una comunidad de aprendizaje dirigida a los niños. Es la persona responsable de proponer programas de actividades adaptadas a los intereses y necesidades pedagógicas de los menores, así como de la resolución de los conflictos que surjan con el objetivo de lograr un ambiente idóneo.

Asimismo, en la Tabla 13 se proveen ordenamientos para cuatro elementos esenciales: las dimensiones a ser medidas, el tiempo mínimo de administración, la disponibilidad de capacitación oficial y el costo total (que en su mayoría refiere al costo de capacitación). Se ofrece también un ordenamiento que puede ser de utilidad al tomar decisiones en otras cuatro dimensiones que son: si el instrumento fue aplicado/adaptado en ALC, si el instrumento cuenta con traducción al español, si el instrumento ha sido validado y finalmente el perfil del aplicador.

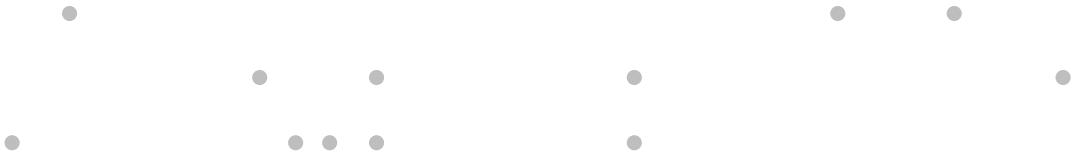
Por ejemplo, en el tiempo de administración está primero en el ordenamiento el CC-IT-HOME que lleva una hora, seguido por los tres instrumentos de rutinas que llevan una hora y media, y el resto de los instrumentos demanda más de una hora y media de observación. Los instrumentos más largos son el ITERS-R, el MITRCC y el PE, todos con más de tres horas.

Muy importante resulta el ordenamiento del costo total, que en parte está relacionado con los costos de licencias de uso y aplicación relacionados con la propiedad intelectual.¹⁴ En esta dimensión, los menos costosos son los instrumentos de rutina y el MITRCC. Es importante notar que no se dispone de datos de estos costos para el CIS y ORCE, y para el PE sólo se reportaron los costos de materiales, que es la razón por la cual resulta más barato en este ordenamiento.¹⁵ A continuación en el ordenamiento se encuentran el ITERS-R (USD 22,9), el CC-IT-HOME (USD 40,3) y por último el Toddler CLASS (USD 902,9). El alto costo de este último instrumento esta explicado porque es el único que requiere una capacitación obligatoria. Vale la pena enfatizar que estos costos son únicamente referentes de unas pocas experiencias de aplicación y que van

a depender mucho del contexto en el cual se apliquen los instrumentos, el tamaño de la lista de centros a evaluar, y la experiencia de quien lleve a cabo la aplicación.

¹⁴ El uso de instrumentos con licencias de propiedad intelectual también conlleva altos costos de transacción.

¹⁵ Es importante notar que al ORCE generalmente lo prestan los que han desarrollado el instrumento por ser este un medio desarrollado para una evaluación federal, pero es una exigencia que se "pida prestado".



IMPLEMENTACIÓN

● CAPÍTULO 4



4.1/ Introducción

Luego de haber reflexionado sobre el instrumento adecuado para la medición de la calidad, la implementación de una evaluación de la calidad conlleva otros interrogantes importantes. ¿Cuál debería ser el perfil de las personas encargadas de recoger información sobre la calidad del servicio de cuidado? ¿Cómo y quién hará la capacitación del equipo que aplicará los instrumentos? ¿Cuánto durará el piloto? ¿Cuáles son las variables del contexto que deberán tenerse en cuenta durante la implementación y los potenciales cambios que necesitarán preverse? ¿Cuáles serán los protocolos a seguirse en campo? ¿Qué protocolo se establecerá para la confiabilidad entre observadores en aquellos instrumentos que requieren de la observación codificada? ¿Con qué frecuencia deben los aplicadores volver a documentar su confiabilidad? Este y otros temas se abordan en esta sección para informar sobre tales decisiones.

4.2/ Capacitación

Los observadores y encuestadores deben ser entrenados adecuadamente por un equipo de capacitadores. Cada uno debe recibir la misma instrucción. Los capacitadores deben ser profesionales, con un amplio conocimiento de desarrollo infantil y el entorno de los servicios de cuidado y con experiencia previa y un profundo conocimiento del instrumento sobre el cual capacitan. Para capacitar en algunos de los instrumentos (como, por ejemplo, CLASS) se requiere que el entrenador haya sido certificado por la empresa que

distribuye el instrumento. Esto supone que ha recibido una capacitación impartida por esa organización y que ha aprobado una serie de evaluaciones que demuestran su conocimiento del instrumento (la certificación adquirida debe actualizarse con cierta periodicidad).

El entrenamiento del equipo a cargo de la recolección de la información sobre calidad comienza con un conocimiento exhaustivo de la prueba a aplicar: ítems y formas de valuación. Además, debe incluir todos los aspectos de la situación de evaluación: contacto y entendimiento con el personal del centro, introducción de la prueba, lectura de las instrucciones, administración de los ítems y registro de las respuestas. Es muy importante que el equipo tenga formación en carreras relacionadas con el desarrollo infantil, psicología o educación y que tenga alguna experiencia con medición en niños o instituciones. Más importante aún es que al equipo se lo familiarice y capacite fuertemente en los instrumentos a aplicar, los ítems que en ellos constan y los conceptos que se presentan antes de salir al campo. De igual manera, el equipo debe tener destreza en el manejo de los instrumentos (formularios, cuestionarios, etc.) que deban emplearse durante el operativo de campo. Así recogerá información precisa y dedicará su atención a la observación del entorno y a la recolección de información, mientras la sistematiza en el cuestionario en forma ágil y sin errores.

Una capacitación efectiva combina las explicaciones de tipo cátedra, con muchos ejemplos, casos y actividades prácticas que simulen situaciones con las cuales los investigadores pueden encontrarse en el campo. Además se discute cómo abordar dichas situaciones. Durante la capacitación, pueden usarse videos y fotos de aulas similares a las del contexto a ser estudiado para ilustrar situaciones que ocurren en los centros de cuidado y motivar la discusión sobre ciertos conceptos o temas. Este tipo de ejercicios deben asegurar que los

investigadores construyan un entendimiento similar sobre conceptos complejos como cuáles elementos caracterizan una interacción cálida y sensible entre el niño y su cuidador y cómo se observan esos elementos en el contexto cultural en el cual se lleva a cabo el estudio en cuestión.

4.3/ **Piloto y adaptación cultural**

La cultura tiene un impacto significativo en la forma en que los conceptos son expresados, así como en la relevancia y dificultad de algunos de ellos. Por esta razón, cuando los instrumentos de medición de calidad son aplicados en un contexto diferente de aquel en el que fueron diseñados o validados, puede ser necesario un esfuerzo de adaptación al lenguaje y la cultura locales. Particularmente, es necesario evaluar cuán equivalente lingüística, funcional, cultural y métricamente es la nueva versión respecto de la original (Peña, 2007).

Existen diversas técnicas para implementar una traducción cultural. Una de ellas consiste en que se traduzca la versión al lenguaje objetivo, luego un segundo traductor (o grupo) se encargue de traducir el texto nuevamente al lenguaje original, y por último se evalúa la precisión de la traducción comparando la versión original con la obtenida por la doble traducción (van de Vijver y Tanzer, 2004). Paralelamente, una persona nativa, bilingüe y bicultural, debería involucrarse desde el comienzo del proceso o al menos en la etapa de revisión (Peña, 2007; van Widenfelt et al., 2005). Idealmente el equipo de traductores debería tener conocimiento, no solo en los dos lenguajes y culturas, sino también en el tema del instrumento, pues si bien se busca una traducción lo más cercana posible al texto original, se necesita de una evaluación crítica y la capacidad de distinguir las situaciones en

las cuales no sería adecuada una traducción literal, ya que se perdería el significado del contenido (van Widenfelt et al., 2005). A pesar de los controles que puedan implementarse al traducir de un lenguaje a otro, la versión obtenida posiblemente sea aún incongruente en su significado y necesite modificarse para obtener un comportamiento objetivo similar en los dos lenguajes (Peña, 2007).

Asimismo, un proceso de adaptación puede involucrar la modificación del orden de los ítems, o incluso el ajuste o eliminación de alguno de ellos (parcial o completamente) por no ser aplicable o apropiado culturalmente (van Widenfelt et al., 2005). No obstante, la adaptación lograda bajo las dos últimas alternativas brinda resultados que no son directamente comparables con la versión original del instrumento (van de Vijver y Tanzer, 2004). Por ejemplo, si se han intercambiado el orden de aplicación de algunos ítems o subescalas en el instrumento a ser aplicado, luego no se podrán hacer comparaciones de los puntajes del instrumento adaptado con los puntajes obtenidos por otros estudios que usan el instrumento original. Es importante tomar en cuenta que algunas de las casas editoriales propietarias de los derechos sobre estos instrumentos contemplan políticas específicas relacionadas con su traducción y adaptación.

Por último, en la sección de encuesta de estos instrumentos es necesario analizar las interpretaciones y respuestas que se obtienen de cada uno de los ítems, así como el grado de dificultad de estos en el nuevo lenguaje: ¿qué piensa el cuidador que se está preguntando? ¿Qué significa la pregunta para el cuidador? La realización de un piloto con un pequeño grupo de personas con características similares a la población de interés es crítica para obtener información útil y evaluar si son necesarias modificaciones adicionales (van Widenfelt et al., 2005). El piloto predice dificultades que podrían originarse en la etapa de recolección

de los datos, brindando la posibilidad de corregir los puntos problemáticos antes de que sea utilizada la prueba (Litwin, 1995). Otro tipo de recursos que pueden ayudar a probar la pertinencia de los instrumentos adaptados son instrumentos de investigación cualitativa como los grupos focales. Se puede convocar a grupos pequeños que tengan características parecidas a la población del estudio y conversar con ellos sobre cómo los ítems en el instrumento responden a la realidad de los centros de cuidado que estos conocen.

En resumen, los pasos apropiados para la adaptación incluyen la traducción exacta de todos los materiales y conceptos subyacentes, la adaptación de los contenidos de las evaluaciones y del proceso de administración de las mismas al contexto local y la realización de pruebas piloto minuciosas que permitan implementar ajustes y reevaluar los instrumentos (Fernald, Kariger et al., 2009). Además, en países donde los recursos financieros y humanos son muy limitados, algunos instrumentos pueden acortarse o simplificarse para reducir costos y tiempos de implementación. Vale la pena recordar que, mientras más se altere un instrumento, las mediciones que se realicen con este serán menos comparables con los resultados de aplicaciones del instrumento original en otros países o contextos.

4.4 / Generación de números únicos de identificación entre bases de datos

Es también muy importante generar números identificadores únicos para cada unidad estudiada (unidad geográfica, centro, cuidador, coordinador, grupo de niños, niño) y que estos indicadores permitan consolidar

los datos de un nivel con los de otro nivel fácilmente. Por ejemplo, es necesario saber si el niño “1” pertenece al grupo “A” o “B” del centro “1” o “2”; de lo contrario al momento del análisis no se podrán identificar las potenciales variables explicativas de la calidad que se quiere medir.

En el mismo sentido, usar identificadores como carnets de identidad, certificados de nacimiento o códigos administrativos de provincia o municipio, o identificadores únicos del centro en relación con su institución madre también facilitarán la consolidación posterior de las diferentes bases con otros datos administrativos. Si bien es necesario ser cuidadosos con respecto a la confidencialidad de los datos y quién tiene acceso a ellos, la literatura reciente ha demostrado que el uso de bases administrativas consolidadas con encuestas “ad hoc” puede dar mayor riqueza al análisis.

4.5/ Variables críticas del contexto

Es extremadamente importante tener un conocimiento profundo del contexto en el que se van a aplicar los instrumentos de medición de la calidad para que estos fluyan y no se encuentren con dificultades en terreno. A continuación se presenta una lista de elementos claves a investigar antes del diseño de cuestionarios y de su prueba piloto:

1. Localización de los centros (urbano, rural, disperso).
2. Frecuencia del servicio de cuidado (diario, o número de días por semana).
3. Horario de atención y número de turnos que brindan al día.
4. Rutinas del centro (hora del recibimiento de los niños; hora del refrigerio, del almuerzo, de la siesta, de la merienda y/o cena; y hora

de la terminación del servicio; ver si los cuidadores tienen tiempo libre durante el día para responder a una encuesta).

5. Estructura organizacional (si los centros son locales, municipales, provinciales, y/o nacionales; cómo se organizan los aspectos operativos, la administración y el financiamiento del centro).

6. Estructura del recurso humano (uno o más coordinadores del centro, uno o más cuidadores por grupo de niños, tamaño de cada grupo, coeficientes de atención según rango de edades, presencia de personal de apoyo que no labora en el centro a tiempo completo).

7. Estructura de los grupos atendidos en el centro (organización de aulas en grupos de edad, la frecuencia de atención a niños de diferentes edades, por ejemplo si hay pocos niños menores a un año o si hay representación de diferentes grupos etarios).

8. Lugares en los cuales transcurren los diferentes momentos de la rutina de los niños (interiores, exteriores, aulas, comedor, etc.).

9. Tamaño de los centros de la población a estudiarse (distribución del número de niños, cuidadores y grupos).

10. Lengua de la población usuaria del servicio y de aquellos que trabajan en los centros.

Además, deberá tomarse en cuenta si las variables antes nombradas varían significativamente entre las diferentes regiones o entre el ámbito urbano y el rural.

4.6/ **Confiabilidad entre observadores** **(*Inter-rater reliability*)**

La confiabilidad entre observadores (*Inter-rater reliability*) cuantifica cuán parecidos son los resultados de la medición de calidad que resulten de la aplicación de un mismo instrumento por diferentes observadores sobre un mismo estudio (Gwet, 2008). En particular, lograr confiabilidad entre observadores significa que diferentes observadores entrenados pueden alcanzar juicios precisos y consistentes sobre el desempeño del cuidador, sobre las dimensiones captadas por las variables de proceso y sobre otras evidencias de la calidad del cuidado que reciben los niños. Este tipo de confiabilidad es importante para asegurar que los distintos observadores están valorando cada situación de la misma manera, y consecuentemente reduce el potencial de sesgo derivado de una valoración personal subjetiva (Fernald, Kariger et al., 2009).

La alta confiabilidad entre observadores se debe alcanzar antes de salir al campo y es necesario mantenerla durante la recolección de la información a partir del análisis de los datos recogidos con una comparación, calibración o recertificación periódica (APPR, 2012). Sin embargo, se debe comenzar a trabajar en obtener confiabilidad entre observadores desde el entrenamiento, con criterios claros y objetivos sobre cómo recoger, analizar y reportar la evidencia observada en los centros. Se pueden tomar videos sobre la rutina diaria en algunos de los centros, y trabajar en grupo sobre los criterios usados, la evidencia y cómo se decide la puntuación o calificación de los ítems de cada prueba.

¿Cómo se prueba si hay confiabilidad entre observadores? Se necesita establecer una persona entrenada y eficiente en la aplicación del instrumento cuya observación se tome como la “regla de oro”. Esta persona valorará la misma situación que un equipo de observadores. Luego se comparan las respuestas de cada uno de los integrantes del equipo con el de la persona considerada como “regla de oro” para asegurar una correlación de al menos 0,8, lo que significa un acuerdo en el 80% de los ítems evaluados (Fernald, Kariger et al., 2009). Este requerimiento de que los observadores demuestren su concordancia con la “regla de oro”, y que el instrumento sea lo suficientemente específico para permitir esta concordancia, es una propiedad fundamental de cualquier prueba de evaluación (Pianta, 2011).

Además, para que el instrumento otorgue un ranking de calidad consistente a un conjunto de centros si estos son evaluados en dos momentos cercanos en el tiempo (por ejemplo, que un centro de calidad mediana sea consistentemente de calidad mediana si es evaluado hoy o en dos semanas) existe la prueba de reprobación de la confiabilidad (*test retest reliability*). La idea es probar que la evaluación de la calidad del centro es consistente en dos momentos cercanos, es decir que el puntaje que se le dio la primera vez no fue el resultado de circunstancias sino que realmente expresa algo sobre la calidad de ese centro. Se esperaría que al menos 80% de los ítems dieran la misma respuesta. Es muy importante documentar este tipo de ejercicio como parte del proceso de preparación de la recolección de información. También debería realizarse “shadow scoring” (o puntaje en la sombra) para que exista consistencia para cada evaluador a través del tiempo.

4.7/ Seguimiento de protocolos

En esta sección se describen algunos de los protocolos que deben definirse para el trabajo de campo durante un levantamiento de información de esta índole.

Protocolo general

Este protocolo contiene la información indispensable que se da al centro a ser estudiado al momento de la introducción del estudio. Por ejemplo, es necesario informar sobre: quién es el investigador, para quién trabaja, para qué se realiza el estudio y quién lo comisiona, qué actividades va a realizar, con quién, cuánto tiempo va a permanecer en el centro, qué recursos de apoyo necesita durante su visita, y qué documentación o espacios necesitará verificar o visitar. También es necesario transmitir que se cuenta con un protocolo preciso sobre cómo trabajar con niños. Además, una vez que se hayan aclarado todas las preguntas que pudieran surgir con respecto a la evaluación, el representante del centro a ser estudiado deberá dar su consentimiento informado para participar en el estudio. Cuando la evaluación de la calidad está encargada a un tercero (por ejemplo, una empresa recolectora de encuestas), es común que la entidad de la cual dependen los servicios provea una carta explicando a los proveedores del servicio de cuidado la naturaleza, el objetivo y las características de estudio y solicitando su participación y colaboración con los investigadores. También es importante explicar si la información será anónima o no, quien tendrá acceso a la misma y que impacto puede tener esta información sobre el funcionamiento del centro.

El segundo tema importante al llegar a un centro es la selección del grupo de niños a ser estudiado. Dado que los grupos en cada centro están organizados de acuerdo a la edad de los niños, existen muchos compromisos cuando uno toma esta decisión. Se puede seleccionar a *todos los grupos* del centro que conforman la edad objetivo (0-36 meses por ejemplo), lo cual puede ser costoso, o seleccionar *un solo grupo* de todos los que quedan totalmente contenidos en el rango de la edad objetivo. Cuando hay más de un grupo que cumple con las características que requiere el estudio, la selección de aquel que se va a estudiar debería hacerse de manera aleatoria si se busca tener validez externa en los niños/cuidador/grupos/aulas que se estará observando. Por ejemplo, se puede asignar un número a cada grupo que queda totalmente contenido en el rango de edad, luego tirar un dado, y ver cuál es el grupo que coincide con el número del dado. Lo óptimo es que el centro no sepa qué aula o clase o grupo de niños se va a observar ese día.

También puede ser importante recoger información demográfica (nombre, edad, etnicidad) de cada uno de los niños del grupo a ser estudiado así como variables relevantes sobre el grupo (su tamaño, el número de cuidadores a su cargo, el rango de edad de niños que se atiende, etc.), aunque algunos instrumentos ya contienen algunas de estas preguntas como parte del mismo. Esta información es importante no solo para la selección del grupo, sino también para la desagregación del análisis e interpretación de los datos.

Entrevista/observación

Cuando se visita un centro de cuidado infantil con el objetivo de entrevistar u observar, es necesario recordar que interesa conocer cómo es la rutina diaria que se desarrolla en el lugar, por lo cual se busca minimizar los efectos que la presencia de observadores/ filmadores/ encuestadores pueda tener sobre las actividades que se

desarrollan habitualmente en el centro de cuidado y cómo el personal y niños se comportan a diario. Se debe actuar con discreción durante la visita, entablando relaciones de confianza y respeto con el personal de los centros. Se recomienda invertir el tiempo que sea necesario al inicio del trabajo para entablar esas relaciones, y conocer al personal, de tal manera que el flujo del día pueda avanzar sin inconvenientes.

El observador deberá ubicarse en un lugar del aula que no perturbe el ambiente, de modo de no interferir en la rutina de actividades mientras observa. Asimismo, se recomienda que no tenga interacción con los niños, ni haga interrupciones al cuidador u otro miembro del personal en el aula. Aunque es necesario que el observador se ubique en una posición en la cual pueda registrar sus observaciones con comodidad, es preferible que no utilice el mobiliario destinado a las actividades de los niños o su cuidadora.

En el caso de realizar una entrevista, se recomienda que la persona use un tono y estilo adecuado y respetuoso. El entrevistador debe sentirse cómodo en su rol, ser capaz de improvisar y adaptarse a diferentes situaciones. Se busca generar una charla fluida en la cual sea capaz de permitir desviar la conversación del tema central y luego retomar el hilo cordialmente, reconocer cuándo un tema requiere mayor indagación (por ejemplo, si detecta inconsistencias entre dos respuestas obtenidas), indagar con mayor profundidad sobre temas que así lo requieran, o interrumpir por unos minutos la entrevista (por ejemplo para prestar atención a los niños). El objetivo fundamental es evitar que las preguntas sean percibidas como amenazadoras o sentenciosas, puesto que esto llevaría a una actitud defensiva y sesgada del personal cuyas respuestas terminarían siendo las que se perciben correctas en lugar de las reales.

Es primordial recordar que ante cualquier contexto o respuesta tanto el entrevistador como el observador deben ser objetivos y comprensivos, y no mostrar una actitud que apruebe o desapruebe la situación. Es decir, el entrevistador no tiene que expresar opiniones o hacer sugerencias y debe mantener un tono neutral a lo largo de su visita e incluso cuando el entrevistado le pregunte su opinión sobre una situación específica. Esta actitud debe ser transmitida al personal del centro con el fin de evitar que las respuestas que brinde el personal sean aquellas que ellos consideran las “correctas”. Un ejemplo de esto es cuando el cuidador consulta al encuestador u observador su opinión sobre la información que está transmitiendo. En este caso, la respuesta del entrevistador no debe dejar dudas respecto de cuál es su rol (obtener y sistematizar información) y sobre la importancia de las respuestas que brinde el cuidador no para calificar su desempeño o el de ese centro sino para entender el estado de la calidad en la muestra o población total que está siendo estudiada.

Si, a pesar de aplicar las recomendaciones mencionadas, la actitud del personal del centro es de disconformidad o incomodidad, la opción más prudente es ofrecer realizar la visita en otro momento.

Filmación

El objetivo principal de realizar una evaluación a partir de la filmación de videos, es obtener información de las actividades y del tipo de acciones e interacciones que se llevan a cabo en un centro de cuidado infantil. Por tanto, es fundamental que todo el proceso de registro audiovisual se realice de la manera más precisa posible. A continuación se enumeran aspectos claves para lograr que la filmación capture adecuadamente la información visual y auditiva presente en un “típico día”.

Recomendaciones generales para una adecuada labor del camarógrafo¹⁶

- Llegar al centro de cuidado antes del inicio de la jornada, con tiempo suficiente para contactar al cuidador, entender en qué espacios se desarrollará la rutina de los niños ese día e instalar los equipos en la sala donde se realizará la filmación.
- Explicar al cuidador que la idea es que la grabación sea lo más natural posible y que no se preocupe de presentar a los niños ni de dar mayores explicaciones sobre su presencia.
- Tener una actitud amable y abierta para no poner tensos/nerviosos ni a los niños ni a los cuidadores.
- No intervenir, hablar con los niños ni ayudar en las actividades. El camarógrafo puede sentarse en una silla para no interferir demasiado en la clase, pero no debe utilizar las mesas de trabajo ni las sillas de los niños. Hay que evitar utilizar el espacio propio de la interacción entre cuidador y niños.
- Guardar las pertenencias del camarógrafo en un lugar donde no queden al alcance de los niños y no obstruyan el espacio de la clase. En lo posible, asistir al aula sólo con los equipos.
- Una vez presente en el centro de cuidado infantil, el camarógrafo debe apagar todo aparato electrónico, incluyendo teléfonos celulares por la duración de su visita.

Recomendaciones sobre la grabación del video

- Utilizar la opción en la cámara que permita grabar durante el mayor número de horas posible (calidad de video: económica), en lugar de utilizar la grabación de alta resolución. Esto minimiza el riesgo de que se tenga que cortar el video sin haber concluido la filmación por un tema de almacenamiento en la memoria de la cámara.

¹⁶ Este protocolo se basa en el protocolo de filmación del Toddler CLASS (Reyes Ugalde y Schodt, 2011). En el estudio realizado por las autoras en Ecuador se usó este mismo protocolo.

- Asegurar que la batería recargable para la cámara esté totalmente cargada. La cámara debe instalarse lo más cerca de una pared al fondo del salón para que se pueda tener una visión general de la sala (por ejemplo, cuando se usa CLASS) o del grupo de niños completo (si se aplicará el MITRCC).¹⁷

Evitar armar estructuras complejas o inestables para apoyar la cámara y que puedan resultar peligrosas. Utilizar siempre la cámara con el trípode para que las tomas sean fijas, lo que facilitará la observación del video y análisis posterior de las imágenes. Preocuparse de que el lugar seleccionado para grabar no reciba luz solar directa de frente. Es muy importante no grabar a contraluz pues esto repercute gravemente sobre la calidad y claridad de la imagen captada.

- Asegurar que la cámara esté prendida y se haya sacado la tapa del lente. Es recomendable hacer una prueba de sonido e imagen antes de comenzar las grabaciones.
- Asegurar constantemente que la cámara de video esté grabando correctamente.
- No cortar las grabaciones por periodos de recesos, ni generar nuevos archivos, pues el trabajo posterior de análisis y codificación se vuelve desordenado y complejo si toda la información no está en un solo archivo.
- Una vez iniciada la filmación, el camarógrafo no debe conversar con el cuidador ni con los niños. Si le hablan, gentilmente puede señalar que al final de la jornada pueden retomar la conversación.

Edición

La edición es una etapa importante y costosa del proceso de trabajar con videos. Los elementos metodológicos fundamentales de este proceso es que los editores estén familiarizados con el protocolo de edición¹⁸, que se vuelva a editar lo que no cumpla el protocolo, que se guarde respaldo de todo

el material para reeditar siempre que sea necesario y que se organice bien el material en forma ordenada (nombres de archivos, identificador único del centro, etc.).

Codificación

El equipo de codificadores trabaja en un espacio fijo de oficina. Para el análisis de los videos (es decir, para la codificación), cada codificador necesita una computadora que tenga capacidad de leer archivos .mov, .mp4, .avi, y DVD; un par de auriculares y una estación de trabajo fija y tranquila.

Los codificadores pueden trabajar en jornada flexible. Se recomienda que no trabajen en codificar videos más de cuatro horas por día (más un descanso) pues la codificación es una tarea que requiere de un nivel de concentración muy alto. Para asegurar la calidad y validez de la evaluación, el trabajo de codificación en video requiere la observación a cargo de dos personas distintas de cada uno de los videos y, al menos en una submuestra, alcanzar una adecuada confiabilidad entre observadores (*inter rater reliability*). Los videos se deben asignar a los codificadores de una manera aleatoria, para evitar que los codificadores estén evaluando los mismos videos al mismo tiempo o en colusión. En los casos de muchas discrepancias entre los dos codificadores en el mismo segmento de video, se puede realizar una tercera codificación individual o de grupo, o encargarla al capacitador o supervisor de los codificadores.

¹⁷ Si se quisiera filmar, por ejemplo, un instrumento cuyo objeto de estudio es un niño (CC-IT-HOME), entonces la cámara debería solo focalizarse en el cuidador y el niño.

¹⁸ El protocolo de edición no se incluyó en el documento para no extralimitarse en su extensión. Sin embargo, está disponible y si algún lector tiene interés en él puede contactar a las autoras (mcaraujo@iadb.org, florencial@iadb.org).



CONSIDERACIONES FINALES

● CAPÍTULO 5



Esta guía de herramientas ha mostrado que la medición de la calidad de los centros de cuidado para bebés y niños pequeños involucra temas claves que van desde la selección de las variables que la miden, hasta cómo elegir un instrumento que refleje lo que realmente se quiere medir, como así también los elementos críticos del contexto y las calificaciones profesionales de los investigadores, por nombrar algunos.

A manera de una reflexión final, esta guía de herramientas propone la identificación de 3 etapas a tener en cuenta cuando se quiere medir la calidad en centros de cuidado para bebés y niños pequeños.

1. Determinar cuál es el propósito de la medición. Por ejemplo, se busca medir la calidad para:

- a. ¿Mejorar calidad de los servicios para expandirlos?
- b. ¿Monitorear los programas?
- c. ¿Realizar evaluaciones de impacto?
- d. ¿Diagnosticar la calidad del servicio?
- e. ¿Rediseñar un currículo?

2. Determinar el tipo de medición a realizarse:

a. Variables estructurales. Son las que identifican la presencia o ausencia de aquellos recursos que facilitan las interacciones características de un entorno de cuidado. Estas variables pueden ser reguladas y se clasifican en cuatro grandes grupos, según se relacionen con la infraestructura del centro, los temas de salud y seguridad, el grupo de niños, y el cuidador. Estas variables son útiles para mediciones que tengan los diferentes propósitos arriba mencionados.

b. Variables de proceso. Los indicadores de proceso tienden a focalizarse en aspectos dinámicos, como las interacciones que los niños tienen con sus cuidadores, la relación entre los propios niños, el comportamiento de los cuidadores, o la implementación

del currículo y de los protocolos de salud y seguridad. Estas variables son útiles cuando el objetivo de la medición se ajusta a situaciones como las descritas arriba en los puntos (a), (c), (d) y (e). Se recomienda que cuando el propósito de la medición sea el monitoreo, se considere el uso de medidas más sencillas de la calidad.

3. Seleccionar el o los instrumentos a aplicar de acuerdo a:

- a. propósito de la medición
- b. validez del instrumento
- c. edad y características del grupo de bebés y niños pequeños sujeto del estudio
- d. si se quiere realizar un diagnóstico con i) observación directa (o filmación) ii) reporte o iii) una combinación de (i) y (ii)
- e. dimensiones de la calidad que se quieren medir
- f. contexto (centro institucional o comunitario o cuidado en hogar; grupos pequeños/ grandes)
- g. adaptación al contexto del país donde se va a realizar la medición
- h. presupuesto para la medición
- i. derechos de autor para el uso de instrumentos
- j. tiempo de administración
- k. tiempo (y costo) de capacitación
- l. perfil requerido del aplicador del instrumento
- m. disponibilidad de personal profesional que pueda conducir las capacitaciones necesarias (lo que implicará distinta duración y costos de capacitación del personal a aplicar el instrumento)

Tabla 1: Variables de estructura y de proceso, algunos ejemplos

	Variables de estructura	Variables de proceso
Salud y seguridad	<ul style="list-style-type: none"> • Medidas de salud pública, procedimientos de salud y seguridad, documentos. 	<ul style="list-style-type: none"> • Prácticas observadas de salud y seguridad. • El cuidador ayuda a los niños a seguir las reglas de seguridad y les explica las razones en que se basan estas reglas.
Grupos de niños	<ul style="list-style-type: none"> • Tamaño del grupo. • Razón niño por adulto. 	<ul style="list-style-type: none"> • Existe interacción entre los niños durante gran parte del día. • El cuidador ayuda a los niños a tener empatía con sus compañeros; explica las acciones, intenciones y sentimientos de los niños a otros niños. • El cuidador interrumpe la interacción negativa entre los niños y los ayuda a entender los efectos de sus propias acciones sobre los demás.
Cuidador	<ul style="list-style-type: none"> • Calificación: años de educación, formación en desarrollo infantil, experiencia previa y capacitaciones. • Planificación de las actividades. • Sistemas de supervisión de los cuidadores. • Salario. 	<p>Comportamiento del cuidador:</p> <ul style="list-style-type: none"> • Los cuidadores se muestran atentos con todos los niños, incluso cuando trabajan sólo con uno. • Cómo responde el cuidador ante el llanto de un niño; el cuidador no expresa fastidio u hostilidad hacia el niño. • Cuántas veces el cuidador usa movimientos bruscos al alimentar a un niño, se queja de su comportamiento o presenta una actitud amenazante. • Los cuidadores saludan a cada niño y a su madre o padre al recibirlos y despedirlos. • Los cuidadores reaccionan rápidamente para solucionar los problemas. <p>Interacciones niño-cuidador:</p> <ul style="list-style-type: none"> • El cuidador usa una variedad de palabras sencillas para comunicarse con los niños. • El cuidador habla acerca de muchos temas distintos con los niños, les hace preguntas simples y/o agrega palabras e ideas a las ideas de los niños. • El cuidador no regaña, critica o castiga al niño. • El cuidador anima a los niños a bailar, aplaudir o cantar juntos. • El cuidador acaricia o besa al niño al menos una vez. <p>Implementación del currículo:</p> <ul style="list-style-type: none"> • El cuidador es flexible en la planificación y actividades, la mayoría de las actividades en el aula están dirigidas por el educador teniendo en cuenta las preferencias de los niños. • Los cuidadores introducen conceptos de correspondencia individual, más-menos-igual o causa y efecto, durante momentos de enseñanza. • La siesta es personalizada y hay actividades para los niños que no duermen. • Hay juego libre durante gran parte del día. • Los cuidadores hacen juegos de bloques sencillos con los niños.
Infraestructura	<ul style="list-style-type: none"> • Acceso al agua potable. • Eliminación de residuos. • Electricidad. • Teléfono. • Espacio físico por niño. • Materiales. • Instalaciones protegidas, áreas de juego. 	

Tabla 2: Comparación de instrumentos usados para monitorear la calidad de los servicios de cuidado en Ecuador, Ciudad de Buenos Aires y México (peso relativo de cada ítem, %)

Temas	Ecuador	Ciudad de Buenos Aires	México	México - Anexo	Ejemplos de la información recabada en cada categoría
Datos de la Supervisión.	0,0%	1,2%	1,8%	0,0%	Fecha de la supervisión, supervisor responsable, estado en que se encontró el centro (operando o sin actividad).
Información del centro.	7,2%	3,9%	2,5%	4,5%	Ubicación geográfica (provincia, cantón, localidad, municipio, barrio), domicilio, teléfono, responsable del centro, fecha de apertura del centro, modalidad, horario de funcionamiento.
Condiciones del inmueble.	12,6%	3,1%	6,5%	15,0%	Estado y material de los techos, paredes, pisos, baños, instalaciones eléctricas y mobiliario; estado y otros usos del inmueble.
Información sobre el personal.	4,0%	29,8%	14,0%	0,0%	Género, antigüedad en el centro, nivel de estudio, capacitaciones, si posee libreta sanitaria del personal (responsable del centro, maestras jardineras, auxiliares, personal de cocina, etc.).
Población atendida.	5,4%	8,1%	2,5%	8,5%	Número de niños y niñas atendidos, por salas o grupos de edad, con o sin apoyo económico.
Seguridad.	70,4%	5,0%	57,5%	4,5%	Condiciones de la iluminación, ventilación, manejo de residuos, presencia de químicos, escaleras y muebles; presencia de elementos de seguridad como matafuegos y botiquín; realización de simulacros; información para casos de emergencia.
Equipamiento, orden e higiene.	0,0%	5,8%	5,5%	5,5%	Nivel de limpieza de baños y cocinas, presencia de baño exclusivo para los niños, inodoros y lavabos fijos, mobiliario libre de elementos que pueden presentar riesgo para los niños (puntas afiladas, astillas, partes oxidadas o sueltas).
Registro de asistencia.	0,0%	0,0%	2,3%	0,0%	Existencia de un registro de asistencia e información registrada (fecha, nombres, hora de entrada y salida).
Documentación del centro y de los niños.	0,0%	5,4%	4,5%	15,5%	Existencia de un reglamento interno, acta de reuniones y planificación de actividades; existencia de copia del acta de nacimiento, foto actualizada, certificado médico, cartilla de vacunación, domicilio particular y teléfono de contacto con la familia de cada uno de los niños.
Atención alimentaria, actividades educativas y monitoreo de crecimiento.	0,4%	36,0%	1,3%	0,0%	Los niños reciben desayuno, almuerzo y/o merienda, cumplimiento del menú y tamaño de la porción; contenidos que se trabajan en actividades de estimulación y educativas; realización de talleres con las familias; control antropométrico.
Gastos de electricidad y agua.	0,0%	0,0%	0,0%	5,0%	Tarifa de los servicios de electricidad y agua, modificaciones en la tarifa desde que funciona el centro.
Visitas de inspección de personal externo.	0,0%	0,0%	0,0%	33,0%	Supervisión por parte de personal de Secretarías o Ministerios que no llevan adelante el programa.
Otros aspectos.	0,4%	1,6%	1,8%	8,5%	Observación de maltrato verbal, físico o psicológico, personal fumando y/o ingiriendo bebidas alcohólicas, supervisión de los niños en todo momento (incluso cuando duermen); condicionamiento del servicio a un pago.
Total de ítems	223	258	400	200	

Tabla 3: Características del instrumento ITERS-R

Instrumento		Escala de calificación del ambiente para bebés y niños pequeños, edición revisada (ITERS-R, por las siglas en inglés de "Infant and Toddlers Environment Rating Scale, revised edition").
Diseño	Autores/Fuente	Harms, Cryer y Clifford (2006).
	País donde fue diseñado	E.E.U.U.
	País donde se aplicó	Brasil - Verdisco y Pérez Alfaro (2010). Chile - Herrera et al. (2005). Colombia - Bernal et al. (2009). Ecuador - Araujo et al. (2015). EEUU - entre otros, Harms, Cryer y Clifford (2006).
	Confiabilidad y validez	Cryer et al. (2004a) analizan la validez y confiabilidad con una prueba de campo sobre grupos de programas de Carolina del Norte, a partir de 90 observaciones con 2 observaciones emparejadas cada una en 45 ambientes grupales. Medidas de confiabilidad: <ul style="list-style-type: none"> • Confiabilidad del observador: <ul style="list-style-type: none"> ▪ Confiabilidad del indicador: de un total de 39 ítems hubo concordancia en el 91,6% de las calificaciones asignadas por los evaluadores. Si sólo se consideran las primeras 6 dimensiones (32 ítems), la concordancia fue del 90,3%. El ítem de Prácticas de Seguridad fue el único con una concordancia menor al 80,0%; ▪ Confiabilidad del ítem: entre pares de observadores, se obtuvo concordancia en un intervalo de 1 punto -en la escala de 7 puntos-. Para los 32 ítems relacionados con niños, se obtuvo concordancia en el 83,0% de los casos. Mientras que para el total de 39 ítems, se obtuvo concordancia en el 85,0% de los casos. La medida de confiabilidad Kappa Cohen ponderada promedio en los primeros 32 ítems fue de 0,55 y en la escala completa de 39 ítems fue de 0,58. Sólo dos ítems presentaron una Kappa ponderada menor a 0,40; en consecuencia los autores realizaron cambios menores en los ítems con una Kappa ponderada menor a 0,50 para mejorar su confiabilidad sin cambiar el contenido básico; • Confiabilidad general: la correlación dentro de cada clase fue de 0,92; • Coherencia interna: en general, la escala tiene un alto nivel de coherencia interna con un alfa de Cronbach de 0,93. Sin embargo, los autores recomiendan tener cuidado al usar las subescalas de Espacio y muebles y Rutinas de cuidado personal. Además, sugieren usar la subescala Estructura del programa sin incluir el ítem 32, a menos que la mayoría de los programas que se evalúen incluyan niños con discapacidades (Cryer et al., 2004a).
Medición	Tipo de evaluación	Observación presencial y reporte.
	Dimensiones	1. Espacio y muebles (5 ítems). 2. Rutinas de cuidado personal (6 ítems). 3. Escuchar y hablar (3 ítems). 4. Actividades (10 ítems). 5. Interacción (4 ítems). 6. Estructura del programa (4 ítems). 7. Padres y personal (7 ítems).
	Resultados	Calificación de cada ítem de 1 a 7. La calificación de la dimensión surge del promedio de las calificaciones de los ítems.
	Tiempo de administración	3 horas y 30 minutos (como mínimo). Esto es, al menos 3 horas de observación y codificación, y de 20 a 30 minutos de reporte.
	Requerimiento del aplicador	Requerimientos particulares: <ul style="list-style-type: none"> • Realizar la capacitación y lograr confiabilidad en la codificación. Requerimientos generales deseados: <ul style="list-style-type: none"> • Buenas relaciones personales. • Facilidad de comunicación verbal (claridad y articulación). • Afinidad con niños. • Estudios de nivel superior relevantes (por ejemplo, en psicología, educación preescolar, docencia primaria). • Experiencia de trabajo con niños. • Experiencia de campo en trabajo cualitativo. • Buena capacidad de negociación y resolución de problemas. • Habilidad para resolver situaciones imprevistas.

Tabla 3: Características del instrumento ITERS-R (cont.)		
Instrumento	Escala de calificación del ambiente para bebés y niños pequeños, edición revisada (ITERS-R, por las siglas en inglés de "Infant and Toddlers Environment Rating Scale, revised edition").	
Aplicación	¿Dónde?	Centro.
	¿A quién?	Ambiente (por ejemplo: salita, aula), grupo de niños, cuidador principal y cuidador asistente cuando este está presente.
	Rango de edad	0-30 meses.
Requerimientos	Capacitación	Para utilizar la escala formalmente Harms, Cryer y Clifford (2006) recomienda la participación en una capacitación dirigida por un capacitador con experiencia en ITERS-R; no es necesario participar en una capacitación oficial. Los codificadores/observadores que usarán la escala deben participar por lo menos en 2 observaciones de clase práctica con un grupo pequeño de observadores y realizar comparaciones de confiabilidad entre ellos. ¹
	Materiales	Manual y hojas de codificación. Mesa y silla para la parte de reporte.
	Otros	El alcance de la confiabilidad deseada dentro de los grupos de codificadores/observadores puede requerir la realización de observaciones prácticas de campo adicionales. Para esto, se pueden adquirir como material adicional videos de capacitación, disponibles en Teacher College Press. ² Además, Cryer et al. (2004a) presenta información y fotos de utilidad para el aprendizaje de la escala o la interpretación de la observación de cara a mejorar la confiabilidad de los observadores.
Costo	Capacitación	No se requiere una capacitación formal. ³
	Material	USD 22,9 el manual. ⁴ Las hojas de codificación son gratuitas y de libre acceso. ⁵
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.
Otros	<p>1. Una persona con nivel de calificación alta (terciario) y especialización en desarrollo infantil podría capacitarse en ITERS-R y CC-IT-HOME al mismo tiempo con lo que la duración de la capacitación podría reducirse.</p> <p>2. Existe una versión de este instrumento aplicable a programas de cuidado infantil en familia, que fue aplicada en Colombia (Bernal et al., 2009): Escala estandarizada de calidad en centro (FDCRS, por las siglas en inglés de "Family Day Care Rating Scale").</p>	

¹En el estudio del BID en Ecuador (Araujo et al., 2015) participaron un capacitador y 15 codificadores/observadores. La capacitación consistió en 5 días en NIEER (New Jersey) de capacitación a la entrenadora ("train the trainer"), divididos en 1 día de teoría y 4 días de prácticas para alcanzar confiabilidad. Además se implementaron 2 días de capacitación a codificadores/observadores en campo.

²<http://www.tcpress.com/ERS.html>.

³Los autores ofrecen una capacitación en la Universidad de Carolina del Norte, Chapel Hill. Para más información ver http://ersi.info/training_live.html.

⁴Los precios fueron recuperados el 28 de diciembre de 2015, <http://www.tcpress.com/ERS.html>.

⁵Disponibles en <http://www.tcpress.com/pdfs/iterss.pdf>.

Tabla 4: Características del instrumento Toddler CLASS

Instrumento	Sistema de Observación y Calificación en el Aula, versión para niños pequeños (Toddler CLASS, por las siglas en inglés de "Classroom Assessment Scoring System").
Diseño	Autores/Fuente Pianta, La Paro y Hamre (2008).
	País donde fue diseñado E.E.U.U.
	País donde se aplicó E.E.U.U. Ecuador (versión traducida al español) - Araujo et al. (2015).
	Confiabilidad y validez Thomason y La Paro (2009) analizan la validez del constructo de Toddler CLASS a partir del estudio de 46 educadores en 30 aulas preescolares diferentes de EEUU. La validez del constructo de Toddler CLASS se establece a partir de revisiones exhaustivas de medidas existentes, una revisión de la investigación sobre los aspectos únicos del desarrollo infantil temprano, y observaciones sobre entornos preescolares. Los instrumentos revisados incluyen el ITERS (Harms et al., 2006), el CIS (Arnett, 1989), y el ORCE (NICHD, 1996). Los ejemplos e indicadores incluidos en el instrumento adaptado reflejan esta revisión. De forma similar al pre-K CLASS (Pianta et al., 2005), se proveen ejemplos para la codificación de entornos entre el rango bajo (1-2), medio (3-5) o alto (6-7). Adicionalmente, el instrumento fue revisado por un experto en bebés y niños pequeños para asegurar la validez de los conceptos adaptados. No se encontró información adicional sobre la validez (ver Apéndice 1 de esta Guía de Herramientas para información de validez del Pre-K CLASS en el que se basa el Toddler CLASS). Las correlaciones entre medidas tradicionales de calidad en educación infantil temprana, como el nivel de educación del cuidador, tamaño del grupo, coeficiente de atención, puntajes asignados (en función de la inclusión de aspectos de calidad en el programa y el nivel educativo de los cuidadores) y las dimensiones fueron significativas.
Medición	Tipo de evaluación Observación (presencial o video).
	Dominios y dimensiones 1. Apoyo emocional y conductual. 1.1. Clima positivo (1 ítem). 1.2. Clima negativo (1 ítem). 1.3. Sensibilidad del educador (1 ítem). 1.4. Consideración por la perspectiva de los niños (1 ítem). 1.5. Orientación de la conducta (1 ítem). 2. Apoyo motivador del aprendizaje. 2.1. Facilitación del aprendizaje y el desarrollo (1 ítem). 2.2. Calidad de retroalimentación (1 ítem). 2.3. Modelaje lingüístico (1 ítem).
	Resultados Calificación de cada dimensión de 1 a 7.
Tiempo de administración	2 horas (como mínimo). Esto es, 4 ciclos de 20 minutos de observación y 10 minutos de codificación.
Requerimiento del aplicador	Requerimientos particulares: <ul style="list-style-type: none"> • Realizar la capacitación y pasar un examen de confiabilidad. • Realizar una prueba diaria de validez. Requerimientos generales deseados: <ul style="list-style-type: none"> • Buenas relaciones personales. • Facilidad de comunicación verbal (claridad y articulación)- Afinidad con niños. • Estudios de nivel superior relevantes (por ejemplo, en psicología, educación preescolar, docencia primaria). • Experiencia de trabajo con niños. • Experiencia de campo en trabajo cualitativo. • Buena capacidad de negociación y resolución de problemas . • Habilidad para resolver situaciones imprevistas.

Tabla 4: Características del instrumento Toddler CLASS (cont.)		
Instrumento	Sistema de Observación y Calificación en el Aula, versión para niños pequeños (Toddler CLASS, por las siglas en inglés de "Classroom Observation Scoring System").	
Aplicación	¿Dónde?	Centro.
	¿A quién?	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.
	Rango de edad	15-36 meses.
Requerimientos	Capacitación	Se requiere que cada codificador/observador obtenga la certificación en el instrumento, a partir de la participación en la capacitación de 2 días y de aprobar el examen de confiabilidad. Para aprobar este examen se debe alcanzar un 80% de acuerdo (se acepta como máximo un punto de divergencia en los puntajes) en 5 casos modelo, y por lo menos 2 puntajes por dimensión deben estar correctos para los 5 casos (sin divergencia alguna en el puntaje). La capacitación puede ser realizada por Teachstone o por un capacitador certificado por Teachstone, quien debe haber obtenido su certificación, en una capacitación de 3 días, hasta un año antes de la capacitación. ¹ Luego, los codificadores/observadores mantienen su validez a través de una prueba diaria de confiabilidad antes de empezar su trabajo. ²
	Materiales	Manual, hojas de codificación, videos modelo, pantalla para proyectar videos.
Costo	Capacitación	USD 4.500 cuesta la capacitación de Teachstone para capacitadores ("train the trainer"). Alternativamente, se puede capacitar en Teachstone a cada codificador/observador por USD 850. ³
	Material	USD 49,9 el manual y USD 15 las hojas de codificación. ⁴
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.
Otros	El instrumento puede codificarse en campo o en video. Para más detalles de la codificación en video ver la Sección 3.4 de esta Guía de Herramientas.	

¹ Para más información sobre las certificaciones en el instrumento de los observadores y del capacitador ver www.classobservation.com.

² En el estudio realizado por el BID en Ecuador (Araujo et al., 2015) participaron un capacitador y 15 codificadores/observadores. El tiempo de capacitación total fue de aproximadamente 3 semanas. Consistió en 3 días de capacitación oficial de Teachstone al capacitador ("train the trainer"), 5 días de preparación de videos modelo y otro material por parte del capacitador, y 5 días de entrenamiento a codificadores/observadores de campo.

³ Los precios fueron recuperados el 28 de diciembre de 2015, corresponden a capacitaciones en EEUU, y no incluyen los gastos de viaje o viáticos, <http://store.teachstone.org/>.

⁴ Los precios fueron recuperados el 28 de diciembre de 2015, <http://store.teachstone.org/>. Las hojas de codificación están disponibles en inglés, en paquetes de 5 cuadernillos. Cada cuadernillo contiene 6 hojas de observación y una hoja de resumen de puntuación. El cuestionario traducido al español que las autoras aplicaron en Ecuador están disponibles, y si algún lector tiene interés puede contactar a las autoras de esta Guía Herramientas.

Tabla 5: Características del instrumento ORCE

Instrumento	Registro de Observaciones de los Entornos de Prestación de Cuidado (ORCE, por las siglas en inglés de "Observational Record of the Caregiving Environment").
Autores/Fuente	NICHD Early Child Care Research Network (1996).
País donde fue diseñado	E.E.U.U.
País donde se aplicó	E.E.U.U. (2002, 2010).
Confiabilidad y validez	<p>En base a una muestra de 1.364 familias, entre 10 sitios de Estados Unidos (NICHD Study of Early Child Care and Youth Development), con niños nacidos en 24 hospitales, la confiabilidad de la información es la siguiente:</p> <ul style="list-style-type: none"> Las escalas de comportamiento presentaron correlaciones Pearson entre 0,41 y 0,99 (la mayoría de las estimaciones por encima de 0,80), en el rango de 0,08 y 0,99, y entre 0,34 y 0,97 para las escalas de 6, 36 y 54 meses de edad, respectivamente. No fue estimada la confiabilidad de cada variable para la escala de 15 meses de edad y se obtuvieron niveles aceptables de confiabilidad en la escala de 24 meses de edad (con algunas excepciones). Las calificaciones cualitativas revelaron correlaciones Pearson por encima de 0,80 para las escalas de 6 meses de edad, en el rango de 0,20 y 0,85 para escalas de 15 meses, entre 0,47 y 0,76 en las escalas de 24 meses, alrededor de 0,57 y 0,93 en escalas de 36 meses, y entre 0,20 y 0,83 en el resto. <p>(Halle, Whittaker y Anderson, 2010).</p>
Tipo de evaluación	Observación (presencial o video).
Diseño	
Dominios y dimensiones	<p>El instrumento tiene cuatro versiones, una para cada edad del periodo en que se recolectan los datos: 6, 15, 24, 36 y 54 meses (las versiones para 24 y 36 meses son las mismas). Las cuatro versiones se dividen en los mismos 3 dominios; sin embargo, hay algunas diferencias en sus dimensiones. A continuación se presenta la versión para 6 meses (ver Halle, Whittaker y Anderson (2010) para una derivación de las otras versiones a partir de esta).</p> <ol style="list-style-type: none"> 1. Escala de comportamiento. <ol style="list-style-type: none"> 1.1 Afecto positivo y negativo (3 ítems). 1.2 Interacción focalizada en el lenguaje (3 ítems). 1.3 Estimulación (2 ítems). 1.4 Orientación del comportamiento (5 ítems). 1.5 Actividades (6 ítems). 1.6 Interacción entre los niños (2 ítems). 2. Calificaciones cualitativas. <ol style="list-style-type: none"> 2.1 Calificaciones de los educadores (8 ítems). 2.2 Calificaciones de los niños (5 ítems). 3. Variables estructurales observadas. <ol style="list-style-type: none"> 3.1 Coeficiente de atención (1 ítem). 3.2 Tamaño de los grupos (1 ítem). 3.3 Número de niños (1 ítem). 3.4 Número de educadores (1 ítem). 3.5 Proporción de la observación que se realiza al aire libre (1 ítem). 3.6 Tiempo que el educador comparte con el niño (1 ítem). 3.7 Edades de los niños en el grupo (1 ítem).
Medición	
Resultados	Calificación de cada ítem de 1 a 4. La calificación total surge del promedio entre dominios para todos los ciclos en un periodo de edad (Vandell y Wolfe, 2000).
Tiempo de administración	90 minutos (como mínimo) siguiendo NICHD Study of Early Child Care Phase I Instrument Document (2004) en Halle, Whittaker y Anderson (2010). Esto es, de 2 a 4 ciclos de 44 minutos. Cada ciclo incluye 3 intervalos de 10 minutos de observación, intercalados por intervalos de 2 minutos para tomar notas, seguidos por 10 minutos para observación y notas focalizados en características globales de comportamiento. Sin embargo, Vandell y Wolfe (2000) mencionan que son un mínimo de 4 ciclos de observación distribuidos en un periodo de 2 días.

Tabla 5: Características del instrumento ORCE (cont.)

Medición	Requerimiento del aplicador	<p>Requerimientos particulares:</p> <ul style="list-style-type: none"> Realizar la capacitación y lograr confiabilidad en la codificación. <p>Requerimientos generales deseados:</p> <ul style="list-style-type: none"> Buenas relaciones personales. Facilidad de comunicación verbal (claridad y articulación). Afinidad con niños. Estudios de nivel superior relevantes (por ejemplo, en psicología, educación preescolar, docencia primaria). Experiencia de trabajo con niños. Experiencia de campo en trabajo cualitativo. Buena capacidad de negociación y resolución de problemas. Habilidad para resolver situaciones imprevistas.
	Aplicación	
	<p>¿Dónde?</p> <p>¿A quién?</p> <p>Rango de edad</p>	<p>Contexto de cuidado no materno (centro, hogar).</p> <p>Un niño y un cuidador u otro adulto presente.</p> <p>6, 15, 24, 36 y 54 meses.</p>
Requerimientos	Capacitación	Es necesario realizar una capacitación y prácticas de administración del instrumento para lograr la confiabilidad de los codificadores/observadores. ¹
	Materiales	Manual y hojas de codificación.
Costo	Capacitación	N/D
	Material	N/D
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.
	Otros	El instrumento puede codificarse en campo o en video. Para más detalles de la codificación en video ver la Sección 3.4 de esta Guía de Herramientas.

¹ En el NICHD Study of Early Child Care se requirió que los codificadores/observadores alcanzaran al menos un 90% de confiabilidad para ser certificados, y se verificó que se mantuviera la confiabilidad cada 4 meses, Bradley et al. (2003) en Halle, Whittaker y Anderson (2010).

Tabla 6: Características del instrumento CIS

Instrumento	Escala de Interacción del Educador (CIS, por las siglas en inglés de "(Arnett) Caregiver Interaction Scale").	
Diseño	Autores/Fuente	Arnett (1989).
	País donde fue diseñado	E.E.U.U.
	País donde se aplicó	Bermudas - Arnett (1989). E.E.U.U. - Jaeger y Funk (2001).
	Confiabilidad y validez	Jaeger y Funk (2001) reportan: <ul style="list-style-type: none"> • Confiabilidad del observador: en un rango de 0,75-0,97 entre observadores certificados y entrenados. Además, se reportan coeficientes de 0,81 y mayores para las escalas de Sensibilidad, Indiferencia y Severidad. • Coherencia interna: se encontró que los alphas Cronbach (Layzer et al., 1993) eran: Sensibilidad= 0,91 y Severidad= 0,90.
Medición	Tipo de evaluación	Observación (presencial o video).
	Dimensiones	1. Sensibilidad (10 ítems). 2. Severidad (8 ítems). 3. Indiferencia (4 ítems). 4. Permisividad (4 ítems).
	Resultados	Calificación de cada ítem de 1 a 4.
	Tiempo de administración	90 minutos. Esto es, 2 observaciones de 45 minutos, en ocasiones separadas (Vandell y Wolfe, 2000).
	Requerimiento del aplicador	Requerimientos particulares: <ul style="list-style-type: none"> • Realizar la capacitación y lograr confiabilidad en la codificación. Requerimientos generales deseados: <ul style="list-style-type: none"> • Buenas relaciones personales. • Facilidad de comunicación verbal (claridad y articulación). • Afinidad con niños. • Estudios de nivel superior relevantes (por ejemplo, en psicología, educación preescolar, docencia primaria). • Experiencia de trabajo con niños. • Experiencia de campo en trabajo cualitativo. • Buena capacidad de negociación y resolución de problemas. • Habilidad para resolver situaciones imprevistas.
Aplicación	¿Dónde?	Contexto de cuidado no materno (centro, hogar).
	¿A quién?	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.
	Rango de edad	36-60 meses.
Requerimientos	Capacitación	Se requiere realizar una capacitación teórica y práctica. Previo a realizar observaciones independientes, los observadores tienen que alcanzar confiabilidad con el capacitador. Para ser un observador certificado en la Escala de Interacción del Educador se debe alcanzar una confiabilidad de 70% en 2 visitas consecutivas (Jaeger y Funk, 2001).
	Materiales	Manual y hojas de codificación.
	Otros	Se requieren 2 observadores diferentes (Vandell y Wolfe, 2000).
Costo	Capacitación	N/D
	Material	El manual y las hojas de codificación son gratuitos y de libre acceso. ¹
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.
Otros	El instrumento puede codificarse en campo o en video. Para más detalles de la codificación en video ver la Sección 3.4 de esta Guía de Herramientas.	

¹ Disponibles en http://www.eec.state.ma.us/docs1/qris/20110121_arnett_scale.pdf

Tabla 7 : Características del instrumento Perfil de Evaluación

Instrumento		Perfil de Evaluación (APECP por las siglas en inglés de "Assessment Profile for Early Childhood Programs").
Diseño	Autores/Fuente	Abbott-Shim, Neel y Sibley (2001).
	País donde fue diseñado	E.E.U.U.
	País donde se aplicó	E.E.U.U. (NICHD, 2000a).
	Confiabilidad y validez	<ul style="list-style-type: none"> Confiabilidad del observador: para las dos versiones (Medida global de calidad y Medida de autoevaluación), la confiabilidad entre el capacitador y observador es consistentemente reportada con una media de 93,0% a 95,0% de concordancia, con un rango de 83,0% a 99,0% de concordancia (Abbott-Shim, Lambert y McCarty, 2000). Otros estudios han reportado niveles similares de confiabilidad entre observadores. Coherencia Interna: los coeficientes de confiabilidad para las 5 dimensiones (Entorno para el aprendizaje, Planificación, Métodos curriculares, Interacciones e Individualización) se encuentran en el rango de 0,79 a 0,98 para Kuder-Richardson 20 y desde 0,81 a 0,98 para el coeficiente de Spearman-Brown. La confiabilidad basada en IRT para las 5 escalas se encuentra entre 0,83 y 0,91 (Abbott-Shim, Neel y Sibley, 2001).
Tipo de evaluación		Observación presencial y reporte.
Medición	Dominios y dimensiones	<p>Medida global de calidad:</p> <ol style="list-style-type: none"> Prácticas en el aula. <ol style="list-style-type: none"> Entorno para el aprendizaje (12 ítems). Planificación (12 ítems). Métodos curriculares (12 ítems). Interacciones (12 ítems). Individualización (12 ítems). <p>Medida de autoevaluación²:</p> <ol style="list-style-type: none"> Prácticas en el aula. <ol style="list-style-type: none"> Prácticas de seguridad en el aula (109 ítems). Entorno de aprendizaje (73 ítems). Planificación (34 ítems). Métodos curriculares (49 ítems). Interacciones (61 ítems). Individualización (25 ítems). Prácticas administrativas. <ol style="list-style-type: none"> Instalaciones (68 ítems). Servicios de comida (45 ítems). Políticas y procedimientos (63 ítems). Personal (38 ítems). Profesionalismo, evaluación y desarrollo (31 ítems).
	Resultados	Calificación de cada ítem Sí/No.
	Tiempo de administración	En un día se pueden evaluar 2 o 3 aulas/grupos. Se debe observar el aula durante la mañana, verificar registros y entrevistar a cuidadores en la tarde. Para el componente administrativo de la Medida de autoevaluación se requiere de 4 a 6 horas. Por lo tanto, si se supone una jornada laboral de 8 horas, se estima que la administración del instrumento en un aula en la Medida global de calidad requiere 3 horas, mientras que la Medida de autoevaluación requiere de 7 a 9 horas.
	Requerimiento del aplicador	<p>Requerimientos particulares:</p> <ul style="list-style-type: none"> Realizar la capacitación y lograr confiabilidad en la codificación. <p>Requerimientos generales deseados:</p> <ul style="list-style-type: none"> Buenas relaciones personales. Facilidad de comunicación verbal (claridad y articulación). Afinidad con niños. Estudios de nivel superior relevantes (por ejemplo, en psicología, educación preescolar, docencia primaria). Experiencia de trabajo con niños. Experiencia de campo en trabajo cualitativo. Buena capacidad de negociación y resolución de problemas. Habilidad para resolver situaciones imprevistas.

Tabla 7: Características del instrumento Perfil de Evaluación (cont.)

Aplicación	¿Dónde?	Centro.
	¿A quién?	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.
	Rango de edad	Medida global de calidad: aulas/grupos con niños de 3 a 7 años. Medida de autoevaluación: bebés (0-12 meses), niños pequeños (12-26 meses), preescolar (3-5 años) y edad escolar (5-10 años) en cada aula/grupo.
Requerimientos	Capacitación	La formación es necesaria para establecer confiabilidad entre los evaluadores. Implica una revisión de los criterios y métodos de recolección de datos y en el lugar de la observación práctica, la revisión de registros y reportes. La formación general lleva de 2 a 3 días (Halle, Whittaker y Anderson, 2010).
	Materiales	Manual y hojas de codificación. Mesa y silla para la parte de reporte.
Costo	Capacitación	N/D
	Material	Medida global de calidad: USD 50 el manual y USD 25 las hojas de codificación ³ . Medida de autoevaluación: USD 50 el manual y USD 120 el paquete de hojas de codificación para cada dominio y rango de edades (Prácticas administrativas y Prácticas en el aula para bebés, niños pequeños, preescolares y escolares) ⁴ .
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.

¹ Las versiones del Perfil de evaluación se denominan en inglés "summative" y "formative".

² El número de ítems en cada dimensión varía dependiendo del grupo de edad observado. Entre paréntesis se detalla el número máximo posible.

³ Todos los precios fueron recuperados el 28 de diciembre de 2015, <http://www.qassist.com/pages/research-and-evaluation>.

⁴ Se pueden adquirir las hojas de codificación por separado para cada grupo de edad (USD 25).

Tabla 8: Características del instrumento MITRCC		
Instrumento	Lista de verificación Missouri para el cuidado sensible de los niños menores de tres años (MITRCC, por las siglas en inglés de "Missouri Infant-Toddler Responsive Caregiving Checklist").	
Diseño	Autores/Fuente	MU Center for Family Policy & Research (2003).
	País donde fue diseñado	E.E.U.U.
	País donde se aplicó	E.E.U.U. Colombia (versión traducida al español y adaptada). Ecuador (versión traducida al español y adaptada) - Araujo et al. (2015).
	Confiabilidad y validez	Thomburg (2009) usa una muestra de 99 niños con la versión 2009 de la lista de chequeo del MITRCC, y reporta un Coeficiente alpha para la lista de chequeo del MITRCC = 0.85, y coeficiente de correlación entre el ITERS y el MITRCC de 0.69.
Medición	Tipo de evaluación	Observación (presencial o video).
	Ítems	20 ítems.
	Resultados	Calificación de cada ítem Sí/No. Calificación total entre 0 y 20, que luego se divide por dos para tener un puntaje final entre 0 y 10.
	Tiempo de administración	3 horas y 30 minutos (como mínimo). Esto es, al menos 3 horas de observación y 30 minutos de codificación.
	Requerimiento del aplicador	<p>Requerimientos particulares:</p> <ul style="list-style-type: none"> Realizar la capacitación y lograr confiabilidad en la codificación. <p>Requerimientos generales deseados:</p> <ul style="list-style-type: none"> Buenas relaciones personales. Facilidad de comunicación verbal (claridad y articulación). Afinidad con niños. Estudios de nivel superior relevantes (por ejemplo, en psicología, educación preescolar, docencia primaria). Experiencia de trabajo con niños. Experiencia de campo en trabajo cualitativo. Buena capacidad de negociación y resolución de problemas. Habilidad para resolver situaciones imprevistas.
	¿Dónde?	Centro.
Aplicación	¿A quién?	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.
	Rango de edad	0-36 meses.
Requerimientos	Capacitación	No es un requisito la realización de una capacitación oficial, ni con un capacitador entrenado. Sin embargo, se recomienda que los codificadores/observadores participen en una capacitación y logren confiabilidad en las codificaciones. Si bien en el manual se proveen algunos ejemplos, puede ser necesario la revisión de ejemplos adicionales, así como es útil la revisión de videos. ¹
	Materiales	Manual, hojas de codificación, pantalla para proyectar videos (opcional).
Costo	Capacitación	No se requiere una capacitación formal.
	Material	El manual y las hojas de codificación son gratuitos y de libre acceso. ²
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.
Otros	El instrumento puede codificarse en campo o en video. Para más detalles de la codificación en video ver la Sección 3.4 de esta Guía de Herramientas.	

¹ En el estudio realizado por el BID en Ecuador (Araujo et al., 2015) participaron un capacitador y 15 codificadores/observadores. El tiempo de capacitación total fue de 2 días (en campo).

² Disponibles en https://www.openinitiative.org/content/pdfs/MoNotes/IT_Checklist_Notes.pdf

El cuestionario traducido al español y adaptado que las autoras aplicaron en Ecuador están disponibles, y si algún lector tiene interés puede contactar a las autoras de esta Guía Herramientas.

Tabla 9: Características del instrumento CC-IT-HOME

Instrumento	HOME para cuidado de bebés y niños pequeños (CC-IT-HOME, por las siglas en inglés de "Child-Care Infant and Toddlers HOME Observation for Measurement of the Environment").	
Diseño	Autores/Fuente	Bradley, Caldwell y Corwyn (2003).
	País donde fue diseñado	E.E.U.U.
	País donde se aplicó	EEUU - NICHD (1996, 2000b) ¹ , Clarke-Stewart et al. (2002). Colombia (versión traducida al español y adaptada por NIEER). Ecuador (versión traducida al español y adaptada por NIEER) - Araujo et al. (2015).
	Confiabilidad y validez	En el estudio NICHD Study of Early Child Care realizaron el análisis de una muestra de 53 observaciones emparejadas. Medidas de confiabilidad: <ul style="list-style-type: none"> Confiabilidad del observador: en los ítems hubo un nivel de concordancia muy alta. Después de 24 meses de recopilación de datos un análisis de las calificaciones emparejadas reveló correlaciones Pearson ($r = 0,94$) y Winer ($r = 0,97$) muy altos. Después de 54 meses las correlaciones Pearson ($r = 0,98$) y Winer ($r = 0,99$) todavía fueron altas (Halle, Whittaker y Anderson, 2010; Bradley et al., 2003, p. 301). Coherencia interna: en general, la escala tiene un alto nivel de coherencia interna con un alfa de Cronbach de 0,81 (NICHD, 1996).
Medición	Tipo de evaluación	Observación presencial y reporte.
	Dimensiones	1. Capacidad de respuesta (11 ítems). 2. Aceptación (7 ítems). 3. Organización (6 ítems). 4. Materiales de aprendizaje (9 ítems). 5. Involucramiento (6 ítems). 6. Variedad (4 ítems).
	Resultados	Calificación de cada ítem Sí/No. Calificación total entre 0 y 43.
	Tiempo de administración	1 hora de observación, reporte y codificación.
	Requerimiento del aplicador	Requerimientos particulares: <ul style="list-style-type: none"> Realizar la capacitación o el entrenamiento sugerido por los autores del instrumento, y lograr confiabilidad en la codificación. Requerimientos generales deseados: <ul style="list-style-type: none"> Buenas relaciones personales. Facilidad de comunicación verbal (claridad y articulación). Afinidad con niños. Estudios de nivel superior relevantes (por ejemplo, en psicología, educación parvularia, docencia primaria). Experiencia de trabajo con niños. Experiencia de campo en trabajo cualitativo. Buena capacidad de negociación y resolución de problemas. Habilidad para resolver situaciones imprevistas.
Aplicación	¿Dónde?	Contexto de cuidado no materno (babysitter o familiares). En Colombia y Ecuador el instrumento fue adaptado para ser usado en centros.
	¿A quién?	Un niño y un cuidador.
	Rango de edad	0-36 meses.

Tabla 9: Características del instrumento CC-IT-HOME (cont.)

Instrumento		HOME para cuidado de bebés y niños pequeños (CC-IT-HOME, por las siglas en inglés de “Child-Care Infant and Toddlers HOME Observation for Measurement of the Environment”).
Requerimientos	Capacitación	La participación en una capacitación a cargo de capacitadores entrenados no es una condición para usar el instrumento. Sin embargo, los autores consideran que es una buena idea su realización; incluso sostienen que es lo que se hace habitualmente antes de usar el instrumento. Si el observador no puede realizar una capacitación, los siguientes pasos son esenciales: 1) leer y releer cuidadosamente el manual; 2) un capacitador acompaña a un observador entrenado en una o dos visitas, codificando de forma independiente y comparando los resultados; 3) realizar 5 visitas más, analizando de forma crítica la técnica utilizada; 4) repetir el paso 2. Se espera alcanzar un 90% de confiabilidad entre los observadores. Los autores del instrumento recomiendan la codificación de videos como práctica adicional para lograr el nivel de confiabilidad deseado. (Caldwell y Bradley, 2003) ² .
	Materiales	Manual y hojas de codificación. Mesa y silla para la parte de reporte.
Costo	Capacitación	No se requiere una capacitación formal.
	Material	USD 40 el manual y USD 15 las hojas de codificación. ³
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.
Otros		Una persona con nivel de calificación alta y especialización en desarrollo infantil, podría capacitarse en ITERS-R y CC-IT-HOME al mismo tiempo, con lo que la duración de la capacitación podría reducirse si se usan ambos instrumentos.

¹ Estos trabajos refieren al IT-HOME, una versión anterior al CC-IT-HOME.

² En el estudio realizado por el BID en Ecuador (Araujo et al., 2015) participaron un capacitador y 15 codificadores/observadores. La capacitación consistió en 4 días en NIEER (National Institute of Early Education Research) de entrenamiento a la capacitadora (“train the trainer”), divididos en medio día de teoría, y tres días y medio para lograr confiabilidad. Luego se capacitó en 4 días a los codificadores/observadores en campo. En el estudio NICHD Study of Early Child Care se realizó una capacitación de medio día, seguida de prácticas con el instrumento para alcanzar una confiabilidad del 90% en la codificación de videos. En relación a este estudio Bradley et al. (2003) sugiere que para lograr la confiabilidad deseada no es generalmente necesario una capacitación tan intensa.

³ Los precios fueron recuperados el 28 de diciembre de 2015, <http://fhdri.clas.asu.edu/home/contact.html>. Las hojas de codificación están disponibles en paquetes de 50 cuadernillos. El cuestionario traducido al español y adaptado que las autoras aplicaron en Ecuador están disponibles, y si algún lector tiene interés puede contactar a las autoras de esta Guía Herramientas.

Tabla 10: Características del instrumento Rutinas

Instrumento		Rutinas.
Diseño	Quien lo diseñó	BID (2012).
	Pais donde se aplicó	Ecuador - Araujo et al. (2015).
	Confiabilidad y validez	N/A - Sin embargo, la Feeding Scale (NCAST) posee medidas de validez interna, externa y predictiva, además de una confiabilidad adecuada.
Medición	Tipo de evaluación	Observación presencial.
	Dominios y dimensiones	<p>1. Comida, basado en Feeding Scale (NCAST, 2012) que fue aplicada en E.E.U.U.</p> <p>1.1 Detalles de la rutina tales como duración y la presencia de otras personas (8 ítems).</p> <p>1.2 Interacción entre un niño y el cuidador a la hora de la comida (20 ítems).</p> <p>2. Pañal/Baño.</p> <p>2.1 Interacción entre un niño y el cuidador durante el cambio de pañales o la ida al baño (22 ítems).</p> <p>3. Llanto/Maltrato/Desconexión, basado en Infant Toddler and Family Instrument (Provence y Apfel, 2001), Family Care Instrument (FCI), Instrumento PRIDI (Engle, Cueto, Ortiz y Verdisco (2011) y Advisory committee BID, Enero 2012).</p> <p>3.1 Estado del niño antes, durante y al terminar el llanto, una situación de maltrato o de desconexión (16 ítems).</p> <p>3.2 Reacción del cuidador ante el llanto, situación de maltrato o de desconexión de un niño (5 ítems).</p>
	Resultados	Calificación de cada ítem Sí/No.
	Tiempo de administración	90 minutos. Esto es, 40 minutos para la dimensión de Comida, 10 minutos para la dimensión Pañal/Baño y 40 minutos en la dimensión Llanto/Maltrato/Desconexión.
	Requerimiento del aplicador	<p>Requerimientos generales deseados:</p> <ul style="list-style-type: none"> • Buenas relaciones personales. • Facilidad de comunicación verbal (claridad y articulación). • Afinidad con niños. • Estudios de nivel superior relevantes (por ejemplo, en psicología, educación preescolar, docencia primaria). • Experiencia de trabajo con niños. • Experiencia de campo en trabajo cualitativo. • Buena capacidad de negociación y resolución de problemas. • Habilidad para resolver situaciones imprevistas.
	¿Dónde?	Centro.
	¿A quién?	Un cuidador y un niño según la situación de rutina observada (un niño que NO come solo; un niño que NO va al baño solo; un niño que llora, ha sido maltratado o se desconecta de la realidad del centro, del cuidador o de los otros niños).
Aplicación	Rango de edad	Según la dimensión: 0-24 meses para Comida, 0-30 meses para Pañal/Baño, 0-60 meses para Llanto/Maltrato/Desconexión.
	Capacitación	Se requieren 2 días de capacitación.
	Materiales	Manual y hojas de codificación.
Requerimientos	Capacitación	No se requiere una capacitación formal.
	Material	Manual y hojas de codificación son gratuitos y de libre acceso (consultar a autoras de esta Guía de Herramientas).
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.
Costo	Capacitación	No se requiere una capacitación formal.
	Material	Manual y hojas de codificación son gratuitos y de libre acceso (consultar a autoras de esta Guía de Herramientas).
	Personal necesario	Entrenador de los codificadores/observadores de campo y codificadores/observadores de campo.

Tabla 11: Características del cuestionario de Cuidador/Coordinador/Estructural

Encuesta		Cuidador, Coordinador y Estructural.	
Diseño	Quien lo diseñó	BID (2012).	
	País donde se aplicó	Ecuador - Araujo et al. (2015).	
	Confiabilidad y validez	N/A	
Medición	Tipo de evaluación	Observación presencial y reporte.	
	Secciones	<p>1. Cuidador.</p> <p>1.1 Datos del cuidador (28 ítems).</p> <p>1.2 Escala de conocimiento sobre desarrollo infantil -KIDI (58 ítems).</p> <p>1.3 Escala sobre prácticas, basada en un instrumento diseñado en EEUU para operacionalizar los lineamientos del NAEYC (National Association for the Education of Young Children) (23 ítems).</p> <p>1.4 Relación del cuidador con UN niño, basada en ITFI (Infant-Toddler and Family Instrument, 2001) (11 ítems).</p> <p>2. Coordinador.</p> <p>2.1 Información sobre todo el personal del centro (12 ítems).</p> <p>2.2 Información sobre el coordinador (10 ítems).</p> <p>2.3 Información sobre el centro (20 ítems).</p> <p>2.4 Escala de conocimiento sobre el desarrollo infantil - KIDI (58 ítems).</p> <p>2.5 Escala de Disciplina (2 ítems).</p> <p>3. Estructural.</p> <p>3.1 Información general sobre el centro (9 ítems).</p> <p>3.2 Información sobre servicios básicos (7 ítems).</p> <p>3.3 Información sobre la infraestructura (4 ítems).</p> <p>3.4 Información sobre vulnerabilidades de los espacios (68 ítems).</p> <p>3.5 Información sobre presencia de riesgos para los niños (11 ítems).</p> <p>3.6 Información sobre capacidades existentes (6 ítems).</p> <p>3.7 Otros (1 ítem).</p>	
	Resultados	Respuestas abiertas o precodificadas.	
	Tiempo de administración	2 horas y 40 minutos. Esto es, 40 minutos de reporte sobre el cuidador, 1 hora de reporte sobre el coordinador, 40 minutos de observación del centro y 20 minutos de reporte sobre el centro.	
	Requerimiento del aplicador	<p>Requerimientos generales deseados:</p> <ul style="list-style-type: none"> • Buenas relaciones personales • Facilidad de comunicación verbal (claridad y articulación) • Afinidad con niños • Estudios de nivel superior relevantes (por ejemplo, en psicología, educación parvularia, docencia primaria) • Experiencia de trabajo con niños • Experiencia de campo en trabajo cualitativo • Buena capacidad de negociación y resolución de problemas • Habilidad para resolver situaciones imprevistas 	
	Aplicación	¿Dónde?	Centro.
		¿A quién?	Cuidador, coordinador, centro, según la dimensión que se aplique.
Requerimientos	Duración de la Capacitación	2 días.	
	Materiales	Manual y hojas de codificación. Mesa y silla.	
Costo	Capacitación	No se requiere una capacitación formal.	
	Material	Manual y hojas de codificación son gratuitos y de libre acceso (consultar a autoras de esta Guía de Herramientas).	
	Personal necesario	Entrenador de los entrevistadores de campo y entrevistadores de campo.	

Tabla 12: Resumen de los instrumentos

Instrumentos	Medición				
	Dominios ¹	Número de dimensiones	Dimensiones/Subescalas ¹	Puntuación de cada ítem	Tiempo mínimo
Toddler CLASS	1. Apoyo emocional y conductual (P). 2. Apoyo motivador del aprendizaje (P).	8	1.1 Clima positivo (P). 1.2 Clima negativo (P). 1.3 Sensibilidad del educador (P). 1.4 Consideración por la perspectiva de los niños (P). 1.5 Orientación de la conducta (P). 2.1 Facilitación del aprendizaje y el desarrollo (P). 2.2 Calidad de retroalimentación (P). 2.3 Modelaje lingüístico (P).	1-7	2 horas.
ITERS-R	N/A	7	1. Espacio y muebles (E). 2. Rutinas de cuidado personal (P). 3. Escuchar y hablar (P). 4. Actividades (P). 5. Interacción (P). 6. Estructura del programa (P). 7. Padres y personal (P).	1-7	3 horas y 30 minutos.
CC-IT-HOME	N/A	6	1. Capacidad de respuesta (P). 2. Aceptación (P). 3. Organización (P). 4. Materiales de aprendizaje (E). 5. Involucramiento (P). 6. Variedad (P).	Sí/No	1 hora.
MITRCC	N/A	N/A	N/A	Sí/No	3 horas y 30 minutos.
Rutinas	1. Comida (P). 2. Pañal/Baño (P). 3. Llanto/ Maltrato/ Desconexión (P).	5	1.1 Detalles de la rutina (P). 1.2 Interacción entre un niño y el cuidador a la hora de la comida (P). 2.1 Interacción entre un niño y el cuidador durante el cambio de pañales/ida al baño (P). 3.1 Estado del niño (P). 3.2 Reacción del cuidador (P).	Sí/No	90 minutos.

Tabla 12 (continuación)

Instrumentos	¿Dónde se puede aplicar?	¿A quién?	Aplicación							Capital Humano			Costo Total (en USD) ³	
			¿A que edades? ²				¿Cómo?			Mínimo requerimiento del aplicador				
			Niños de 0-15 meses	Niños de 15-36 meses	Niños de 36-60 meses	Niños de 60 meses y más	Observación presencial	Reporte	Observación por videos	Ninguno	Secundario	Terciario		
Toddler CLASS	Centro.	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.		x				x		x			x	902,9
ITERS-R	Centro.	Ambiente, grupo de niños, cuidador principal y cuidador asistente cuando este está presente.	x	x				x	x				x	22,9
CC-IT-HOME	Contexto de cuidado no materno.	Un niño y un cuidador.	x	x				x	x				x	40,3
MITRCC	Centro.	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.	x	x				x		x			x	0
Rutinas	Centro.	Un niño y un cuidador.	x	x	x			x					x	0

Tabla 12 (continuación)

Instrumentos	Medición				
	Dominios ¹	Número de dimensiones	Dimensiones/Subescalas ¹	Puntuación de cada ítem	Tiempo mínimo
ORCE	1. Escalas de comportamiento (P). 2. Calificaciones cualitativas (P). 3. Variables estructurales observadas (E).	15 ⁴	1.1 Afecto positivo y negativo (P). ⁵ 1.2 Interacción focalizada en el lenguaje (P). 1.3 Estimulación (P). 1.4 Orientación del comportamiento (P). 1.5 Actividades (P). 1.6 Interacción entre los niños (P). 2.1 Calificaciones de los educadores (P). 2.2 Calificaciones de los niños (P). 3.1 Coeficiente de atención (E). 3.2 Tamaño de los grupos (E). 3.3 Número de niños (E). 3.4 Número de educadores (E). 3.5 Proporción de la observación que se realiza al aire libre (E). 3.6 Tiempo que el educador comparte con el niño. 3.7 Edades de los niños en el grupo (E).	1-4	90 minutos
CIS	N/A	4	1. Sensibilidad (P). 2. Severidad (P). 3. Indiferencia (P). 4. Permisividad (P).	1-4	90 minutos ⁶
Perfil de evaluación - Medida Global de Calidad	1. Prácticas en el aula (P)(E).	5	1.1 Entorno para el aprendizaje (E). 1.2 Planificación (E). 1.3 Métodos curriculares (E). 1.4 Interacciones (P). 1.5 Individualización (P).	Si/No	3 horas ⁷
Perfil de evaluación - Medida de autoevaluación	1. Prácticas en el aula (P)(E). 2. Prácticas administrativas (E).	11	1.1 Prácticas de seguridad en el aula (E). 1.2 Entorno de aprendizaje (E). 1.3 Planificación (E). 1.4 Métodos curriculares (E). 1.5 Interacciones (P). 1.6 Individualización (P). 2.1 Instalaciones (E). 2.2 Servicios de comida (E). 2.3 Políticas y procedimientos (E). 2.4 Personal (E). 2.5 Profesionalismo, evaluación y desarrollo (E).	Si/No	7 horas

¹ En esta tabla se considera el constructo de dominio o dimensión según lo definido por los manuales de cada instrumento. Entre paréntesis, el tipo de variables: de proceso (P) y estructurales (E).

² Los rangos de edades propuestos no representan los rangos exactos para los cuales pueden aplicarse los siguientes instrumentos: ITERS-R, Comida, Pañal/Baño, ORCE y Perfil de evaluación - Formative. Para un mayor detalle ver las tablas resumen de cada uno de los instrumentos.

³ Corresponde al costo de los materiales. La única excepción es Toddler CLASS, en el cual se incluye el costo de la capacitación oficial que es obligatoria.

⁴ Las dimensiones asociadas a los dominios "Escalas de comportamiento" y "Calificaciones cualitativas" corresponden a la versión para niños de 6 meses. Las versiones para 15, 24, 36 y 54 meses adicionan dimensiones o presentan diferencias.

Tabla 12 (continuación)

Instrumentos	¿Dónde se puede aplicar?	¿A quién?	Aplicación						Capital Humano			Costo Total (en USD) ³	
			¿A qué edades? ²				¿Cómo?		Mínimo requerimiento del aplicador				
			Niños de 0-15 meses	Niños de 15-36 meses	Niños de 36-60 meses	Niños de 60 meses y más	Observación presencial	Reporte	Observación por videos	Ninguno	Secundario		Terciario
ORCE	Contexto de cuidado no materno.	Un niño y un cuidador u otro adulto presente.	x	x	x		x					x	N/D
CIS	Contexto de cuidado no materno.	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.			x		x					x	N/D
Perfil de evaluación -Medida Global de Calidad	Centro.	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.				x	x	x				x	43,0
Perfil de evaluación - Medida de autoevaluación	Centro.	Grupo de niños, cuidador principal y cuidador asistente cuando este está presente.	x	x	x	x	x	x				x	N/D

¹ El instrumento tiene 4 versiones, una para cada edad del período en que se recolectan los datos: 6, 15, 24, 36 y 54 meses (las versiones para 24 y 36 meses son las mismas). Las 4 versiones se dividen en los mismos 3 dominios; sin embargo, las dimensiones presentan diferencias. Aquí se presentan la versión para 6 meses (ver Halle, Whittaker y Anderson (2010) para una derivación de las otras versiones a partir de esta).

² Estos 90 minutos tienen que implementarse en 2 observaciones separadas de 45 minutos cada una (Vandell y Wolfe, 2010).

³ En un día se observan/reportan 2 o 3 grupos. Se consideran 8 horas de trabajo en el día.

Tabla 13: Comparaciones de los instrumentos

Instrumentos	Número de dimensiones	Tiempo mínimo de administración		Costo total ¹		Capacitación oficial	Aplicado / adaptado en algún país de ALC	Traducción al español	Validez ²	Nivel educativo del aplicador
		Tiempo	Orden (de menor a mayor tiempo)	Costo (en USD)	Orden (de menor a mayor costo)					
Toddler CLASS	8	2 horas	3	902,9	6	Sí	Sí	Sí	Sí	Terciario
ITERS-R	7	3 horas y 30 minutos	5	22,9	2	No	Sí	Sí ³	Sí	Terciario
CC-IT-HOME	6	1 hora	1	40,3	3	No	Sí	Sí	Sí	Terciario
MITRCC	N/A	3 horas y 30 minutos	5	0	1	No	Sí	Sí	Sí	Terciario
Rutinas	3	1 hora y 30 minutos	2	0	1	No	Sí	Sí	N/A	Secundario
ORCE	3	1 hora y 30 minutos	2	N/D	N/D	No	No	No	Sí	Terciario
CIS	4	1 hora y 30 minutos	2	N/D	N/D	No	No	No	Sí	Terciario
Perfil de evaluación - Medida global de calidad.	5	3 horas ⁴	4	75,0	4	No	No	No	Sí	Terciario
Perfil de evaluación - Medida de autoevaluación.	11	7 horas	6	170,0	5	No	No	No	Sí	Terciario

¹ Corresponde al costo de los materiales por observador-codificador. La única excepción es Toddler CLASS, en el cual se incluye el costo de la capacitación oficial para un observador que es obligatoria.

² Corresponde a la existencia o no de estudios sobre la validez y confiabilidad del instrumento.

³ Versión publicada por Teachers College Press.

⁴ En un día se observan/reportan 2 o 3 grupos. Se consideran 8 horas de trabajo en el día.

Referencias bibliográficas

- Abbott-Shim, M. (1991). *Quality care: A global assessment*. Unpublished manuscript. Georgia State University.
- Abbott-Shim, M., Lambert, R., y McCarty, F. (2000). Structural model of Head Start classroom quality. *Early Childhood Research Quarterly*. 15(1): 115-134.
- Abbott-Shim, M., Neel, J., y Sibley, A. (2001). *Assessment Profile for Early Childhood Programs- Research Edition II: Technical Manual*. Atlanta, GA: Quality Counts, Inc.
- Annual Professional Performance Review (APPR). (2012). <http://goo.gl/orIqwB>
- Araujo, M.C., López Boo, F. y Puyana, J.M. (2012). “Panorama sobre los servicios de desarrollo infantil temprano en América Latina y el Caribe”, IDB-MG-149, *División de Protección Social y Salud, Banco Interamericano de Desarrollo*.
- Araujo, M.C. López Boo, F., Novella R., Schodt S. y Romina Tomé (2015). “La calidad de los Centros Infantiles del Buen Vivir en Ecuador”. Resumen de políticas DB-PB-248.
- Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology*. 10: 541-522.
- Arredondo, S.A., Del Pilar Torres Pereda, M., Kenefick, S., Neufeld Lynnette, M., De la Peza, P.G. y Valencia, A. (2011). “Evaluación de impacto del programa de estancias infantiles para apoyar a madres trabajadoras”. CIEE.
- Attanasio, O., Duque, J. F., Fitzsimmons, E., Grantham-McGregor, S., Meghir, C. y Rubio-Codina, M. (2012). Early Childhood Development Pilot in Colombia: Home Visiting & Nutrition. Center for the Evaluation of Development Policies @ IFS and UCL, London. Consultado el 25 de abril de 2012. <http://www.theigc.org/wp-content/uploads/2014/08/ecd-pilot-in-colombia-for-india-igc.10.pdf>
- 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.
- Barnett, W. (2012). Presentación “What is quality” en IDB Advisory Committee, enero de 2012, Washington DC.
- Barros, R., de Carvalho, M., Franco, S., Mendonca, R. y Rosalem, A. (2011). “A Short-Term Cost Effectiveness Evaluation of Better-Quality Daycare Centers”. Inter-American Development Bank.
- Behrman, J., Cheng, Y. y P. Todd. (2004). “Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach.” *Review of Economics and Statistics*. 86(1): 108–132.
- Belsky, J. (1988). “The effects of infant day care reconsidered”. *Early Childhood Research Quarterly*. 3: 237-272.
- Berlinski, S., Galiani, S. y Gertler, P. (2009). “The effect of pre-primary education on primary school performance.” *Journal of Public Economics*. 93:219-234.
- Berlinski, S., Galiani, S. y MacEwan, P. (2008). “Pre-school and maternal labor outcomes: evidence from a regression discontinuity design”. *IFS Working Paper*. W09/05.

- Berlinski, S. y Schady, N. eds. 2015. *The Early Years: Child Well-being and the Role of Public Policy*. Development in the Americas Series. New York: Palgrave MacMillan, and Washington, DC: Inter-American Development Bank.
- Bernal, R., Fernández, C., Flórez, C., Gaviria, A., Ocampo, P., Samper, B. y Sánchez, F. (2009). *Evaluación de impacto del Programa hogares comunitarios de bienestar del ICBF*. Documentos CEDE, ISSN 1657-7191.
- BID- Advisory Committee on Quality of Center-Based Care for Infants and Toddlers (2012). *Measuring Quality of Center-Based Care for Infants and Toddlers in Latin America and The Caribbean, Summary of Main Agreements*, Banco Interamericano de Desarrollo, Washington, DC, 20 de enero de 2012.
- Bowlby, J. (1969). *Attachment. Attachment and loss: Vol. 1. Loss*. New York: Basic Books.
- Bradley, R., Caldwell, B. y Corwyn, R. (2003). The Child Care HOME Inventories: Assessing the quality of family child care homes. *Early Childhood Research Quarterly*. 18: 294-309.
- Burchinal, M., Campbell, F., Bryant, D., Wasik, B. y Ramey, C. (1997). "Early intervention and mediating processes in cognitive performance of children of low- income African American families". *Child Development*. 68: 935-954.
- Burchinal, M., Kainz, K. y Cai, Y. (2011). "How well do our measures of quality predict child outcomes?". En Zaslow et al. Eds. *Quality Measurement in Early Childhood Settings*. Brookes Publishing, Washington, DC.
- Caldwell, B., y Bradley, R. (2003). *Home Observation for Measurement of the Environment: Administration manual*. Little Rock, AR: University of Arkansas.
- Campbell, F. y Ramey, C. (1995). "Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention". *American Educational Research Journal*. 32: 743-772.
- Clarke-Stewart, K., y Fein, G. (1983). "Early Childhood Programs." En P. H. Mussen, Ed. (M. M. Haith y J. J. Campos, series Eds.). *Handbook of Child Psychology, Volume II: Infancy and Developmental Psychobiology*. New York: John Wiley and Sons.
- Clarke-Stewart, K., Gruber, C. y Fitzgerald, L. (1994). *Children at Home and in Day Care*. Hillsdale, NJ: Erlbaum.
- Clarke-Stewart, K., Vandell, D., Burchinal, M., O'Brien, M. y McCartney, K. (2002). "Do regulable features of child-care homes affect children's development?" *Early Childhood Research Quarterly*. 17:52-86.
- Cryer, T., Harms, T. y Riley, C. (2004a). "All About the ITERS-R". Lewisville, NC: PACT House Publishing.
- Cryer, T., Harms, T. y Riley, C. (2004b). "All About the ECERS-R". Lewisville, NC: PACT House Publishing.
- Denham, S., Ji, P. y Hamre, B. (2010). *Compendium of Preschool Through Elementary School Social-Emotional Learning and Associated Assessment Measures*. University of Illinois, Chicago.
- Dunn, L. (1993). "Proximal and Distal Features of Day Care Quality and Children's Development". *Early Childhood Research Quarterly*. 8: 167-192.
- Engle, P., Black, M., Behrman, J., Cabral de Mello, M., Gertler, P., Kapirini, L., Martrell, R. y Eming Young, M. (2007). "Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world". *The Lancet*. 369:229-242.

- Engle, P., Cueto, S., Ortíz, M.E. y Verdisco, A. (2011). “Programa Regional de Indicadores de Desarrollo Infantil (PRIDI). Marco Conceptual”. Banco Interamericano de Desarrollo.
- Engle, P., Fernald, L., Alderman, H., Behrman, J., O’Gara, C., Yousafzai, A., et al. (2011). “Strategies for reducing inequalities and improving developmental outcomes for young children in low and middle income countries”. *The Lancet*. 378(9799):1339-1353.
- Fernald, L., Gertler, P. y Neufeld, L. (2009). “10 year effect of Oportunidades, Mexico’s conditional cash transfer programme, on child growth, cognition, language, and behavior: a longitudinal follow-up study”. *The Lancet*. 374:1997-2005.
- Fernald, L., Kariger, P., Engle, P. y Raikes, A. (2009). “Examining Early Child Development in Low-Income Countries: A toolkit for the Assessment of Children in the First Five Years of life”. The World Bank.
- Filp, J. y Scheifelbein, E. (1982). “Efecto de la educación preescolar en el rendimiento de primer grado de primaria: El estudio UMBRAL en Argentina, Bolivia, Colombia y Chile”. *Revista Latinoamericana de Estadística Educación*. 12: 9-41.
- Filp, J., Donoso, S., Cardemil, C., Dieguez, E., Torres, J. y Sheifelbein, E. (1983). “Relationship between preprimary and grade one primary education in state schools in Chile”. En K. King y R. Meyers, Eds. *Preventing School Failure: The Relationship Between Preschool and Primary Education*. Ottawa: IDRC.
- Fundación Van Leer. (2012). Early Childhood Matters.
- Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L. y Strupp, B. (2007). “Developmental potential in the first 5 years for children in developing countries”. *The Lancet*. 369:60-70.
- Grantham-McGregor, S., Powell, C., Walker, S., et al. 1991. “Nutritional Supplementation, Psychosocial Stimulation, and Mental Development of Stunted Children: The Jamaican Study.” *Lancet*. 338(8758): 1-5.
- Grantham-McGregor, S., Walker, S., Chang, S., et al. 1997. “Effects of Early Childhood Supplementation With and Without Stimulation on Later Development in Stunted Jamaican Children.” *American Journal of Clinical Nutrition*. 66: 247-253.
- Greenberg, M., Domitrovich, C. y Bumbarger, B. (2001). “The prevention of mental disorders in school-aged children: Current state of the field”. *Prevention & Treatment*, 4, Article 1.
- Greenberg, M., Weissberg, R., O’Brien, M., Zins, J., Fredericks, L., Resnik, H. y Elias, M. (2003). “School-based prevention: Promoting positive social development through social and emotional learning”. *American Psychologist*. 58(6/7): 466-474.
- Gonzalez-Mena, J. y Widmeyer Eyer, D. (2007). *Infants, Toddlers, and Caregivers: A Curriculum of Respectful, Responsive Care and Education*. McGraw-Hill Education.
- Gwet, K. (2008). “Computing inter-rater reliability and its variance in the presence of high agreement”. *British Journal of Mathematical and Statistical Psychology*. 61: 29-48: The British Psychological Society.
- Halle, T., Anderson, R., Blasberg, A., Chrisler, A. y Simkin, S. (2011). *Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIT): A Review of the Literature*, OPRE 2011 – 25. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

- Halle, T., Vick Whittaker, J.E. y Anderson, R. (2010). *Quality in Early Childhood Care and Education Settings: A Compendium of Measures, Second Edition*. Washington, DC: Child Trends. Elaborado por Child Trends for the Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Halpern, R. (1986). "Effects of Early Childhood Intervention on Primary School Progress in Latin America". *Comparative Education Review*. 30 (2): 193-215. The University of Chicago Press on behalf of the Comparative and International Education Society. Recuperado el 12 de marzo de 2012 de <http://www.jstor.org/stable/1188528>.
- Hamre, B. y Pianta, R. (2007). Learning opportunities in preschool and early elementary classrooms. En R.C. Pianta, M.J. Cox y K.L. Snow, Eds. *School readiness & the transition to kindergarten in the era of accountability*. 49-83. Baltimore, MD: Paul H. Brookes Publishing Co.
- Hanushek, E. (1998). "Conclusion and controversies about the effectiveness of school resources". *Economic Policy Review*, Federal Reserve Bank of New York. 4(1): 11-28.
- Hanushek, E. (2003). "The failure of input based schooling policies". *Economic Journal*. 113: 64-98.
- Harms, T., y Clifford, R. (1980). *The early childhood environment rating scale*. New York: Teachers College Press.
- Harms, T., Cryer, T. y Clifford, R. (2006). *Infant/Toddler Environment Rating Scale Revised Edition – ITERS-R*, Teachers College Press.
- Hausfather, A., Toharia, A., LaRoche, C. y Engelsmann, F. (1997). "Effects of Age of Entry, Day-Care Quality, and Family Characteristics on Preschool Behavior." *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 38: 441-448.
- Heckman, J. (2006). "Skill Formation and the Economics of Investing in Disadvantaged Children". *Science*. 312(5782): 1900-1902.
- Helburn, S., editor. (1995). *Cost, Quality and Child Outcomes in Child Care Centers: Technical Report*. Denver, CO: Dept of Economics, Center for Research in Economic and Social Policy, University of Colorado.
- Herrera, M., Mathiesen, M., Merino, J. y Recart, I. (2005). "Learning contexts for young children in Chile: process quality assessment in preschool centres". *International Journal of Early Years Education*. 13 (1): 13-27.
- Howes, C. (1983). "Caregiver Behavior in Center and Family Day Care." *Journal of Applied Developmental Psychology*. 4: 99-107.
- Howes, C. (1997). "Children's Experiences in Center-Based Child Care as a Function of Teacher background and adult:child ratio". *Merrill-Palmer Quarterly: Journal of Developmental Psychology*. 43(3): 404-425.
- Howes, C. y Hamilton, C. (1993). "The changing experience of child care: Changes in teachers and in teacher-child relationships and children's social competence with peers". *Early Childhood Research Quarterly*. 8: 15-32.
- 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.
- Howes, C. y Smith, E. (1995). "Relations among Child Care Quality, Teacher Behavior, Children's Play Activities, Emotional Security, and Cognitive Activity in Child Care." *Early Childhood Research Quarterly*. 10: 381-404.

- Jaeger, E. y Funk, S. (2001). *The Philadelphia Child Care Quality Study: An Examination of Quality in Selected Early Education and Care Settings*. Philadelphia, PA: Saint Joseph's University.
- Kagan, S.L. (2010). Presentación "Promoting ECD Quality: Policy Lessons for Central America", en IDB, noviembre de 2010, El Salvador.
- Kisker, E., Hofferth, S., Phillips, D. y Farquhar, E. (1991). *A Profile of Child Care Settings: Early Education and Care in 1990, Volume 1*. Princeton, NJ: Mathematica Policy Research.
- Kontos, S. y Wilcox-Herzog, A. (1997). "Influences on Children's Competence in Early Childhood Classrooms." *Early Childhood Research Quarterly*. 12: 247-262.
- 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.
- La Paro, K., Hamre, B. y Pianta, R. (2011). Classroom Assessment Scoring System. Toddler Manual. Pre-publication advance review draft. Teachstone.
- Lamb, M. (1996). "Effects of Non-Parental Child Care on Child Development: An Update." *The Canadian Journal of Psychiatry*. 41: 330-342.
- Lamb, M. (1998). "Nonparental Child Care: Context, Quality, Correlates, and Consequences." En W. Damon, I. E. Sigel, and K. A. Renninger, Eds. *Handbook of Child Psychology. Vol. 4: Child Psychology in Practice*. New York: John Wiley.
- Layzer, J., Goodson, B., y Moss, M. (1993). *Observational Study of Early Childhood Programs, Final Report, Volume I: Life in Preschool*. Cambridge, MA: Abt Associates, Inc.
- Litwin, M. (1995). *How to measure survey reliability and validity*. The survey Kit, Vol 7. SAGE Publications.
- Love, J., Harrison, L., Sagi-Schwartz, A., van Ijzendoorn, M., Ross, C. y Ungerer, J. (2003). "Child care quality matters: How conclusions may vary with context". *Child Development*. 74: 969-1226.
- Love, J., Kisker, E., Ross, C., Raikes H., Constantine, J., Boller, K., Brooks-Gunn, J., Chazan-Cohen, R., Tarullo, L., Brady-Smith, C., Fuligni, A., Schochet, P., Paulsell, D. y Vogel, C. (2005). "The effectiveness of early head start for 3-year-old children and their parents: lessons for policy and programs". *Dev Psychol*. 41(6):885-901
- MacPhee, David (1981) Knowledge of Infant Development Inventory. Survey of child care experiences - Manual.
- Martinez-Beck, I. (2011). "Introduction: Why Strengthening the Measurement of quality in Early Childhood Settings Has Taken on New Importance". En Zaslow et al., Eds. *Quality Measurement in Early Childhood Settings*. Paul H. Brookes Publishing Co.
- McKay, A. y McKay, H. (1983). "Primary School Progress after Preschool Experience: Troublesome Issues in the Conduct of Follow-Up Research and Findings from the Cali, Colombia, Study". En M. Herrera y C. Super. *School Performance and Physical Growth of Underprivileged Children: Results of the Bogota Project at Seven Years*, Report to the World Bank, Cambridge, MA: Harvard School of Public Health.
- MU Center for Family Policy & Research. (2003). Missouri Infant/Toddler Responsive Caregiving Checklist. Disponible en: <http://goo.gl/RYjAqy>
- National Association for the Education of Young Children (NAEYC). (2012). Recuperado el 22 de marzo de 2012 de: <http://www.naeyc.org/>

- NCAST. (2012). Parent-Child Interaction (PCI) Feeding & Teaching Scales. Recuperado el 18 de abril de 2012 de <http://ncast.org/index.cfm?category=2>.
- Nelson, G., Westhues, A. y MacLeod, J. (2003). "A meta-analysis of longitudinal research on preschool prevention programs for children". *Prevention & Treatment*. 6 (1). Disponible en: https://www.researchgate.net/publication/232570779_A_Meta-Analysis_of_Longitudinal_Research_on_Preschool_Prevention_Programs_for_Children
- NICHD Early Child Care Research Network. (1996). "Characteristics of Infant Child Care: Factors Contributing to Positive Caregiving." *Early Childhood Research Quarterly*. 11: 269–306.
- NICHD Early Child Care Research Network. (1999). "Child Outcomes when Child Care Center Classes meet Recommended Standards for Quality." *American Journal of Public Health*. 89(7):1072-1077.
- NICHD Early Child Care Research Network. (2000a). "Characteristics and quality of child care for toddlers and preschoolers". *Applied Developmental Science*. 4: 116-135.
- NICHD Early Child Care Research Network. (2000b). "The relation of child care to cognitive and language development". *Child Development*. 71: 960-980.
- NICHD Early Child Care Research Network (2002a). "Early child care and children's development prior to school entry: Results from the NICHD Study of Early Child Care". *American Education Research Journal*. 39:133–164.
- NICHD Early Child Care Research Network. (2002b). "The relation of global first- grade classroom environment to structural classroom features and teacher and student behaviors". *Elementary School Journal*. 102(5): 367–387.
- NICHD Early Child Care Research Network. (2009). "Study of Early Child Care and Youth Development Phase I Instrument Document Observational Record of the Caregiving Environment (ORCE): Behavior Scales, Qualitative Scales, and Observed Structural Variables for 6, 15, 24, & 36 months".
- Office of Head Start. (2008). Head Start. Recuperado el 10 de noviembre de 2009 de: <http://www.acf.hhs.gov/programs/ohs/>
- Peisner-Feinberg, E., Burchinal, R., Clifford, R., Culkin, M., Howes, C., Kagan, S. y Yazejian, N. (2001). "The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories through Second Grade". *Child Development*. 72(5): 1534-1553.
- Peña, E. (2007). "Lost in translation: Methodological considerations in cross-cultural research". *Child Development*. 78(4): 1255-1264.
- Phillips, D., editor. (1987). *Quality in Child Care: What Does Research Tell Us?*. Washington, DC: National Association for the Education of Young Children.
- Phillipsen, L. C., Burchinal, M. R., Howes, C. y Cryer, D. (1997). "The prediction of process quality from structural features of child care". *Early Childhood Research Quarterly*. 12: 281–303.
- Pianta, R. (2003). "Experiences in p-3 classrooms: The implications of observational research or re-designing early education". *Foundation for Child Development*, New York.
- Pianta, R. (2011). *Teaching Children Well: New Evidence-Based Approaches to Teacher Professional Development and Training*. Center for American Progress.
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., y Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*. 9(3): 144–159.

- Pianta, R. C., La Paro, K. M., y Hamre, B. K. (2008). Classroom assessment scoring system, manual, K-3. Baltimore, MD: Brookes Publishing.
- PREAL. (2006a). *Quantity Without Quality: A Report Card on Latin America*. Recuperado el 11 de marzo de 2012 de: <http://elibrary.kiu.ac.ug:8080/jspui/bitstream/1/1465/1/Quantity%20without%20Quality%20-%20Education%20in%20Latin%20America.pdf>
- Provence, S. y Apfel, N. (2001). "Infant-toddler and Family Instrument (ITFI)". Paul H. Brookes Publishing Co. Baltimore.
- Raikes, H. y Edwards, C. (2009). Extending the dance in infant and toddler caregiving. Baltimore, MD: Brookes Publishing Company.
- Reyes Ugalde, J. y Schodt, S. (2011). "Manual de registro de video UBC 2011" Documento de trabajo.
- Rolnick, A. y Grunewald, R. (2007). "Early Intervention on a Large Scale". Recuperado el 28 de marzo de 2012 de: <https://www.minneapolisfed.org/publications/special-studies/early-childhood-development/early-intervention-on-a-large-scale>
- Rosero, J. y Oosterbeck, H. (2011). "Trade-offs between different early childhood interventions: Evidence from Ecuador". IADB Unpublished Manuscript.
- Schweinart, L., Montie, J., Xiang, Z., Barnett, W., Belfield, C. y Nores, M. (2005). "Lifetime effects: The High/Scope Perry Preschool study through age 40". Monographs of the High/Scope Educational Research Foundation, 14. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Schodt, Sara (2012). Nota conceptual del proyecto de diagnóstico de calidad de los CIBV.
- Shonkoff, J. y Phillips, D., editores. (2000). *From Neurons to Neighbourhoods: The Science of Early Childhood Development*. National Academy Press, Washington, DC.
- Teachstone. (2012a). The Head Start Professional Development and CLASS Initiative: Ensuring Strong Linkages between Program Monitoring and Professional Development. Recuperado el 10 de abril de 2012 de: <http://www.vbgrowsmart.com/providers/Documents/CLASSImplementationGuide.pdf>
- Teachstone. (2012b). Virginia's Star Quality Initiative: Careful Planning and Infrastructure Development. Recuperado el 10 de abril de 2012 de: <http://www.vbgrowsmart.com/providers/Documents/CLASSImplementationGuide.pdf>
- Teachstone. (2012c). About the CLASS. Recuperado el 11 abril de 2012 de: www.teachstone.com
- Thomason, A. y La Paro, K. (2009). "Measuring the quality of teacher-child interactions in toddler child care". *Early Education and Development*. 20(2):285-304.
- Urzúa, S. y Veramendi, G. (2011). "The impact of Out-of-Home Childcare Centers on Early Childhood Development". Inter-American Development Bank. Unpublished manuscript.
- Vandell, D. y Wolfe, B. (2000). "Child care quality: Does it matter and does it need to be improved?" Reported prepared for the US Department of Health and Human Services, Office for Planning and Evaluation. *Child Development*. 63: 938-949.
- Van de Vijver, F., y Tanzer, N. (2004). "Bias and equivalence in cross-cultural assessment: An overview". *Revue européenne de psychologie appliquée*. Elsevier. 54: 119-13.
- Van Widenfelt, B., Treffers, P., de Beurs, E., Siebelink, B. y Koudijs, E. (2005). "Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families". *Clinical Child and Family Psychology Review*. 8(2):135-147.

- Verdisco, A. y Pérez Alfaro, M. (2010). "Measuring education quality in Brazil". Inter-American Development Bank, disponible en: <https://publications.iadb.org/handle/11319/3100?locale-attribute=en>
- Vernon-Feagans, L., Emanuel, D. y Blood, I. (1997). "The Effect of Otitis Media and Quality Daycare on Children's Language Development." *Journal of Applied Developmental Psychology*. 18: 395-409.
- Volling, B. y Feagans, L. (1995). "Infant day care and children's social competence". *Infant Behavior and Development*. 18:177-188.
- Walker, S., Chang, S., Vera-Hernández, M. y Grantham-McGregor, S. (2011). "Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior". *Pediatrics*. 127(5): 849-857 (doi: 10.1542/peds.2010-2231): <http://pediatrics.aappublications.org/content/127/5/849>
- Walker, S., Chang, S., Grantham-McGregor, S., Osmond, C. y López Boo, F. (2012). "Do early childhood experiences affect development in the next generation: Preliminary findings from the Jamaica intergenerational study". Trabajo de investigación presentado en "Early Childhood Development and Human Capital Accumulation" conferencia en UCL, Londres, 25-26 junio de 2012.
- Weber, K y Kariger, P. (2011). Presentación "Instrumentos para la evaluación de calidad de centros y desarrollo infantil", en IDB, diciembre de 2011, Nicaragua.
- Wilkes, D. (1989). "Administration, classroom program, sponsorship: Are these indices of quality care in day care centers." (Doctoral Dissertation, Georgia State University, 1989). *Dissertation Abstracts International 50*, AAI8922912.
- Yamauchi, C. y Leigh, A. (2011). Which children benefit from non-parental care?, *Economics of Education Review*. 30(6): 1468-1490, ISSN 0272-7757.
- Yoshikawa, H., Barata, M., Rolla, A., da Silva, C., Ayoub, C., Arbour, M., y Snow, C. (2008). "Un Buen Comienzo, an initiative to improve preschool education in Chile: Data from the initial year of implementation". Santiago Chile: Unicef.
- Zaslow, M., Halle, T., Martin, L., Cabrera, N., Calkins, J., Pitzer L., y Geyelin, N. (2006). "Child Outcome Measures in the Study of Child Care Quality". *Eval Rev*; 30; 577.
- Zill, N., Resnick, G., Kim, K., O'Donnell, K., Sorongon, A., McKey, R. H. y otros (2003). "Head Start FACES 2000: A whole-child perspective on program performance". Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth, and Families.
- Zins, J., Bloodworth, M., Weissberg, R., y Walberg, H. (2007). "The scientific base linking social and emotional learning to school success". *Journal of Educational and Psychological Consultation*, 17: 191-210.

Anexo 1. Validez del instrumento Pre K CLASS (basado en Pianta et al., 2005)

Validez del criterio

Los dominios de apoyo emocional, organización del aula y apoyo educativo en CLASS se correlacionan con informes de los maestros en cuanto a depresión y actitudes del adulto. En concreto, las aulas con menores puntuaciones en estas dimensiones reportaron que sus maestros presentaban mayores niveles de depresión, mientras que aquellos con puntuaciones más bajas en la organización del aula y el apoyo a la instrucción reportaron un mayor número de actitudes centradas en los adultos y no en los niños.

Validez concurrente

En las comparaciones de CLASS con la Escala de evaluación del ambiente en la primera infancia (ECERS-R), las aulas con un puntaje CLASS superior presentaban también mayores puntajes en el ECERS-R de interacciones (con rango de correlaciones de 0,45 a 0,63). Las correlaciones entre el puntaje CLASS total y el factor de muebles y materiales de la escala ECERS-R fue sólo moderada e iba desde 0,33 hasta 0,36 (Pianta et al., 2005).

No es sorprendente que la correlación entre el dominio de apoyo educativo del CLASS y el tiempo invertido en alfabetización y matemáticas según la medida del “Snapshot” (instrumento de observación que mide la calidad y variedad de actividades que el profesor les ofrece a los alumnos en la clase, evaluando el porcentaje de tiempo dedicado a diversas actividades; se usa para indicar en qué medida el profesor implementa el currículo [ver Pianta et al., 2005]) fueron bajas (pero significativa) ya que el CLASS mide la calidad más que la cantidad de las actividades.

Validez predictiva

Los resultados de un estudio de varios estados de Estados Unidos proporciona evidencia de que la calidad de la enseñanza, según el CLASS, está asociada con el rendimiento de los niños al final de la edad preescolar, así como con mejoras en el rendimiento de todo el año de preescolar (Howes et al., 1992). La escala de CLASS en la dimensión de apoyo emocional se asoció con un crecimiento en las puntuaciones de los niños en el lenguaje expresivo y receptivo, así como con disminuciones en los problemas de conducta reportados por el maestro (Howes et al., en imprenta).

Validez de contenido

Las dimensiones de CLASS se basan en interacciones observadas entre profesores y estudiantes en las aulas. Las dimensiones se identificaron a partir de una extensa revisión de los constructos validados por otros instrumentos de observación de clases, tales como los que se utilizan en la literatura del cuidado de niños, la investigación sobre la educación primaria o las prácticas efectivas de enseñanza. Esta revisión de la literatura se complementó con consultas a través de grupos focales y con pruebas piloto.

ISBN: 978-1-59782-221-3



9 781597 822213

9 0 0 0 0

